



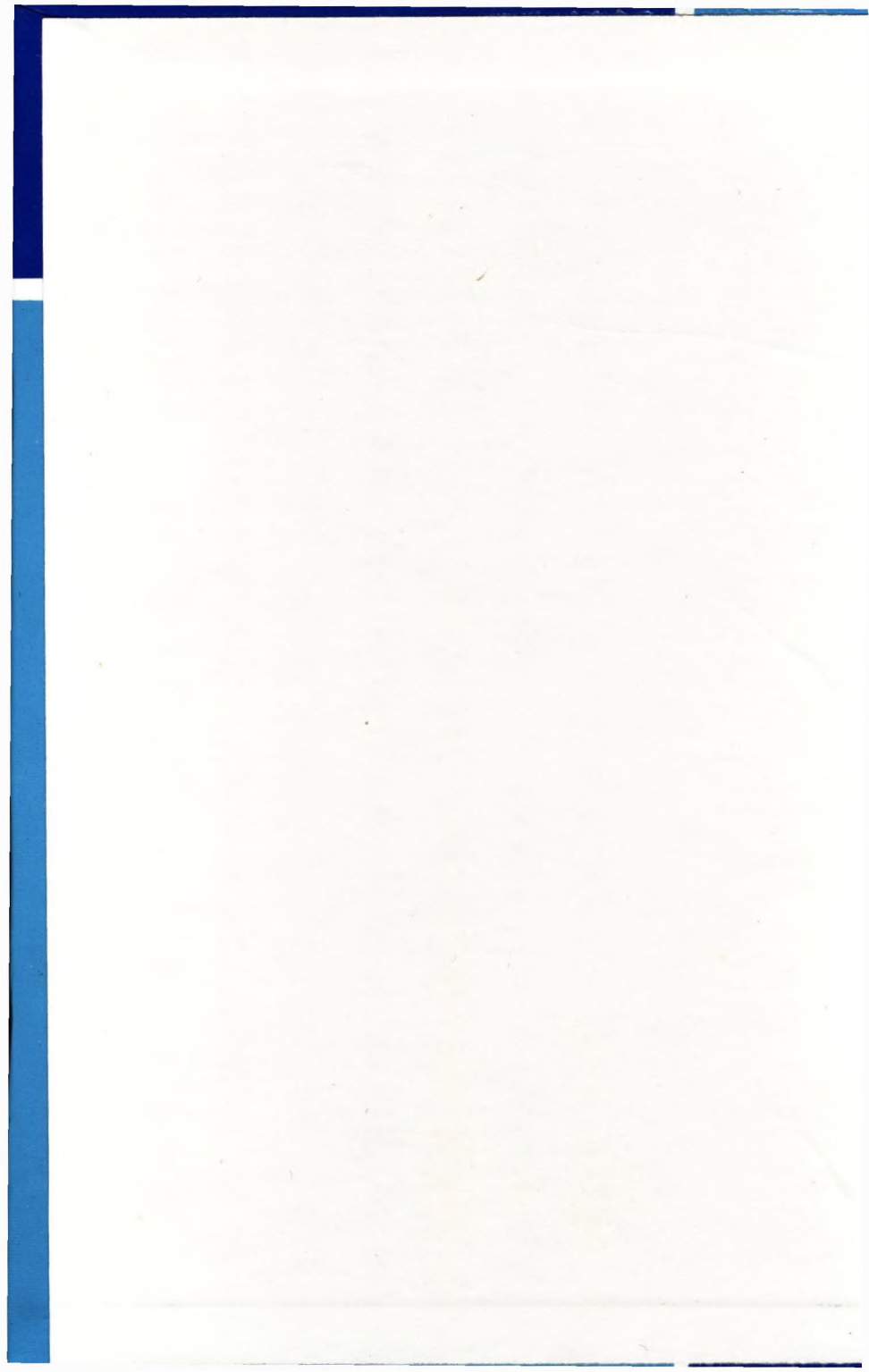
**Az elektronika
újabb eredményei**

MŰSZAKI TUDOMÁNYOK

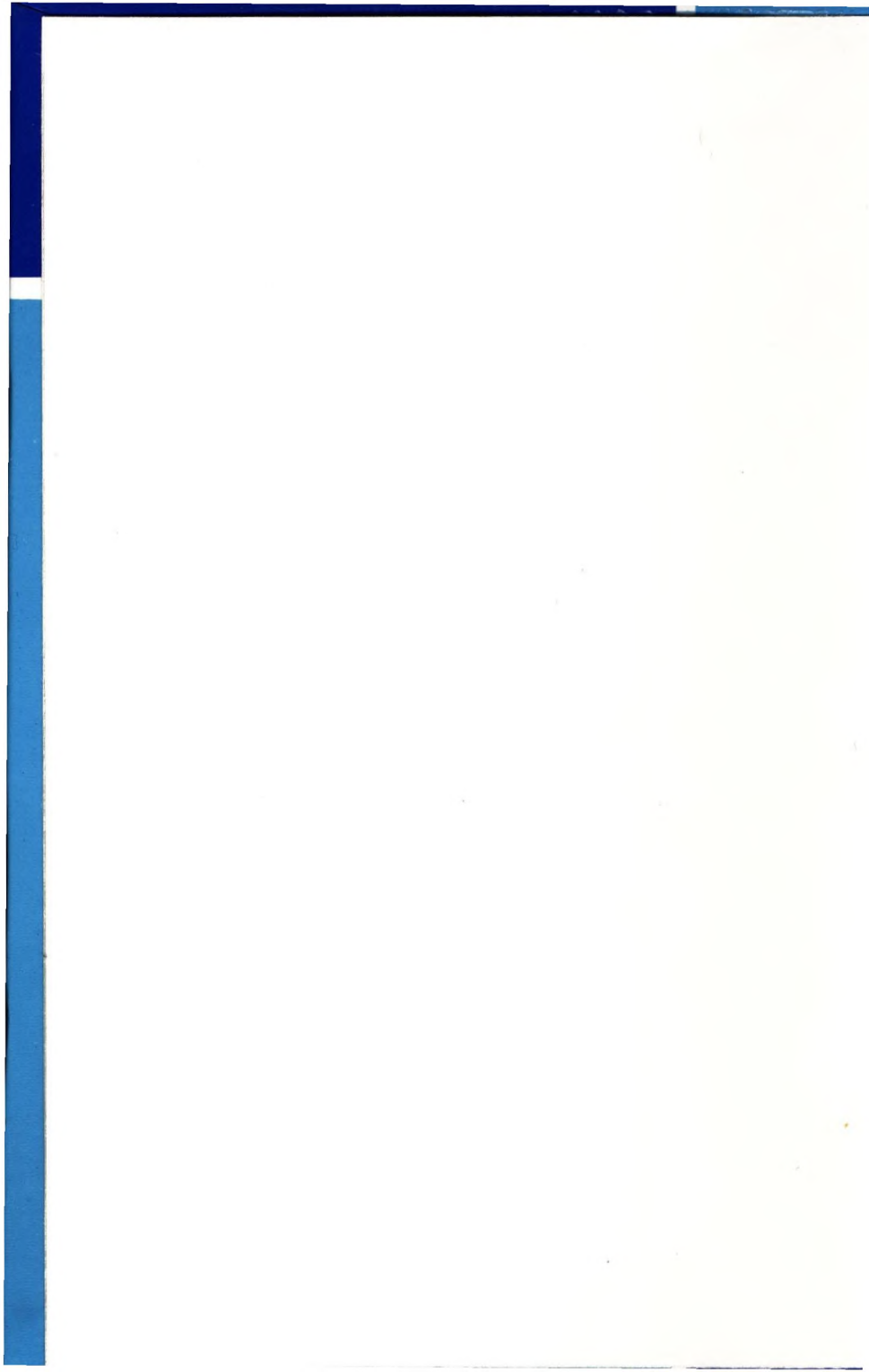
ÁLLÓ GÉZA—HEGEDŰS GY. CSABA—
KELEMEN DEZSŐ—SZABÓ JÓZSEF

**A DIGITÁLIS
KÉPFELDOLGOZÁS
ALAPPROBLÉMÁI**

AKADÉMIAI KIADÓ · BUDAPEST







**MŰSZAKI TUDOMÁNYOK
AZ ELEKTRONIKA ÚJABB EREDMÉNYEI**

MŰSZAKI TUDOMÁNYOK AZ ELEKTRONIKA ÚJABB EREDMÉNYEI

4. kötet

Szerkeszti

PAP LÁSZLÓ

A szerkesztőbizottság

elnöke: GÉHER KÁROLY

titkára: PAP LÁSZLÓ

tagjai: CSIBI SÁNDOR

CSURGAY ÁRPÁD

GYULAI JÓZSEF

SOMLÓ JÁNOS

TÓFALVI GYULA

TUSCHÁK RÓBERT

VALKÓ I. PÉTER



AKADÉMIAI KIADÓ · BUDAPEST 1989

ÁLLÓ GÉZA—HEGEDŰS GY. CSABA—
KELEMEN DEZSŐ—SZABÓ JÓZSEF

A DIGITÁLIS KÉPFELDOLGOZÁS ALAPPROBLÉMÁI



AKADÉMIAI KIADÓ · BUDAPEST 1989

FELMÉRÉS A
KÉPZŐ-NEVELÉSI
RÉSZLETBŐL

A BIZALOM
KÉPZŐ-NEVELÉSI
RÉSZLETBŐL

ISBN 963 05 4897 6

Kiadja az Akadémiai Kiadó, Budapest

© Álló Géza, 1989

Minden jog fenntartva, beleértve a sokszorosítás, a nyilvános előadás, a rádió- és televízióadás valamint a fordítás jogát, az egyes fejezeteket illetően

Printed in Hungary

ELŐSZÓ

A 80-as évek elején a számítógépes képfeldolgozás egyes területein hazánkban is tuljutottunk a kísérleti kutatások szintjén. Ennek eredményeképpen egyrészt ma már számos, általános célú kisgépre (TPA-1140, TPA-1148, R-11, SZM-4), illetve hazai vagy import mikro- és személyi számítógépre (HT-680X, IBM-PC/XT/AT, PROPER-16, ROEOTRON-6270, PERICOLOR-2000 stb.) kidolgozott digitális képfeldolgozó rendszer működik üzemszerűen.

Ezeken kívül számos célrendszer készült egy-egy konkrét feladat megoldására pl. mikroszkópi, tomográf-, hő-kamera-felvételek elemzésére; ürfelvételek feldolgozására mezőgazdasági, meteorológiai vagy egyéb célból stb.). Másrészt megjelentek és az utóbbi időben széles körben elterjedtek a legkülönbözőbb grafikus rendszerek.

A potenciális vagy eseti felhasználók köre jóval nagyobb; náluk általában az okoz gondot, hogy a számítógépes módszerek egyelőre meglehetősen költségigényesek. Ennek két fő oka van:

- egyrészt a hazai fejlesztésű hardvereszközök ára *aránytalanul magas*, s ráadásul még a képfeldolgozáshoz speciális eszközök is szükségesek;
- másrészt az eddig kidolgozott képfeldolgozási programok többnyire *nem vihetők át* egyik gépről a másikra, emiatt minden alkalmazói rendszert gyakorlatilag nulla szintről kell kifejleszteni.

A második probléma ismét két okra vezethető vissza:

- *Nincs* semmilyen *képfeldolgozási programnyelv*.

A jelenlegi eszközeink műveleti sebessége és tárcapacitása mellett elfogadható futási időket csak assembly nyelvű programokkal lehet elérni.

- A képfeldolgozásnak ez idő szerint *nincs egységes elmélete*. Egy-egy feladatra egyedi (ad hoc) megoldások léteznek, de csak kísérleti úton lehet eldönteni, hogy egy konkrét alkalmazásban melyiket kell választani.

Ezen utóbbi gondon szeretnénk némileg enyhíteni tanulmánykötetünkkel, amelyben az utóbbi években elért eredményeket tekintjük át részben az irodalom, részben saját tapasztalataink alapján. Az olvasóról feltételezzük, hogy *rendelkezik* azokkal az alapismeretekkel, amelyek a "Bevezetés a számítógépes képfeldolgozásba" c. korábbi munkánkban találhatóak [A-1]. Ismereteit azonban máshol is szerezhette, ezért - a könnyebbség és az egységes szóhasználat kedvéért - az 1. fejezetben összefoglaljuk a legfontosabb *alapismereteket*, a függelékben pedig áttekintjük a leggyakrabban használt *jelöléseket* és *matematikai módszereket*.

Teljes körű átfogó munka kidolgozását az adott kérések sajnos nem tették lehetővé, de még az óriási (tulnyomóan angol nyelvű) irodalmi anyag áttekintő ismertetésére sem vállalkozhattunk. Ehelyett hasznosabban itéltük, ha inkább csak néhány - nagyrészt általunk is művelt - a gyakorlatban fontos témakört ismertetünk nagyobb mélységben. Ez az elképzelés jobban illik a sorozat jellegéhez is. A feldolgozott témakörök:

- képjavítások (2. fejezet),
- geometriai korrekciók (3. fejezet),
- szegmentálás (4. fejezet) és
- képosztályozás (5. fejezet).

Munkánkat *rendszerzésnek* is szántuk, s ezt kettős értelemben:

- Egyrészt - az irodalmi adatok és saját eredményeink alapján - az 1.1.2 alpontban megadtunk egy modellt, amelyben összefoglaltuk a számítógépes képfeldolgozás főbb folyamatait és kapcsolataikat; a kiindulási és az eredményképeket. A tárgyalás során kijelöljük témaköreink helyét ebben a modellben.
- Másrészt az egyes témakörökön belül kíséreltünk meg rendszerzést, a felhasznált *módszerek* és a lehetséges *alkalmazási területek* szerint.

A tárgyalás mélységét a rendelkezésünkre álló terjedelemlhez kellett igazítanunk. Az ezen túlmenő részletek iránt érdeklődő olvasóknak kívántunk segítséget nyújtani az irodalomjegyzék összeállításával, amelyet a könyv végén közlünk.

Sajnos, a korlátozott terjedelem miatt nem kerülhetett sor fontos képosztályozási módszerek, elsősorban a *texturaelemzés* (texture analysis), valamint a *mintaillesztés* (template matching) és a *szintaktikus alakfelismerés* (syntactic pattern recognition) érdemi tárgyalására. Egyáltalán nem foglalkozunk továbbá a digitális képfeldolgozás legmagasabb szintjének, a *képfelismerésnek* (scene analysis) problémaköreivel sem; ez utóbbiak gyakorlati vizsgálatára és megoldására nálunk egyelőre csak laboratóriumi keretek között van lehetőség.

Az anyagban törekedtünk a magyar nyelvű szakkifejezések használatára, amihez olvasóink szives támogatását és javaslatait is kérjük. Megadjuk azonban az angol nyelvűeket is, hogy segítsük a tájékozódást az irodalomban.

A Számítástechnikai Kutatóintézet és Innovációs Központ (SZKI) programjában mintegy 10 éve jelent meg a képfeldolgozás, dr. Dénes József kezdeményezésére. Szervező munkája sokban hozzájárult ahhoz, hogy ez a tevé-

kenység - az *Intézet* vezetésének erkölcsi és anyagi támogatásával - tuljutott a kezdeti nehézségeken és jelentős sikerekhez vezetett.

Tanulmánykötetünk megírása során elsődlegesen két nagyméretű, általános célu, kispépes képfeldolgozó rendszerre támaszkodtunk, amelyek - a szerzőket is megába foglaló - közel 30 fős, lelkes kutató-fejlesztő kollektiva sokéves munkájának eredményeként jöttek létre:

- a "Modular Image Processing" (MIP) elnevezésű rendszer [1-31] az *Országos Műszaki Fejlesztési Bizottság* megrendelésére készült, TPA-1148 vagy ezzel kompatibilis számítógépekre; elsősorban ürfelvétel feldolgozására szolgál;
- a "VIKING" rendszer [1-48] kifejlesztését a *VIDEOTON Elektronikai Vállalat* finanszírozta, R-11 számítógépre; alapvetően természetes képek (légi-, kamera- és mikroszkópi felvételek) gépi kiértékelése céljából.

Ezuton köszönjük meg *minden munkatársunk* közreműködését, akik részt vettek a két rendszer létrehozásában; külön megemlítve *Bodrogi Hedvig, Feró László és Staszny Gábor* kiemelkedő teljesítményét.

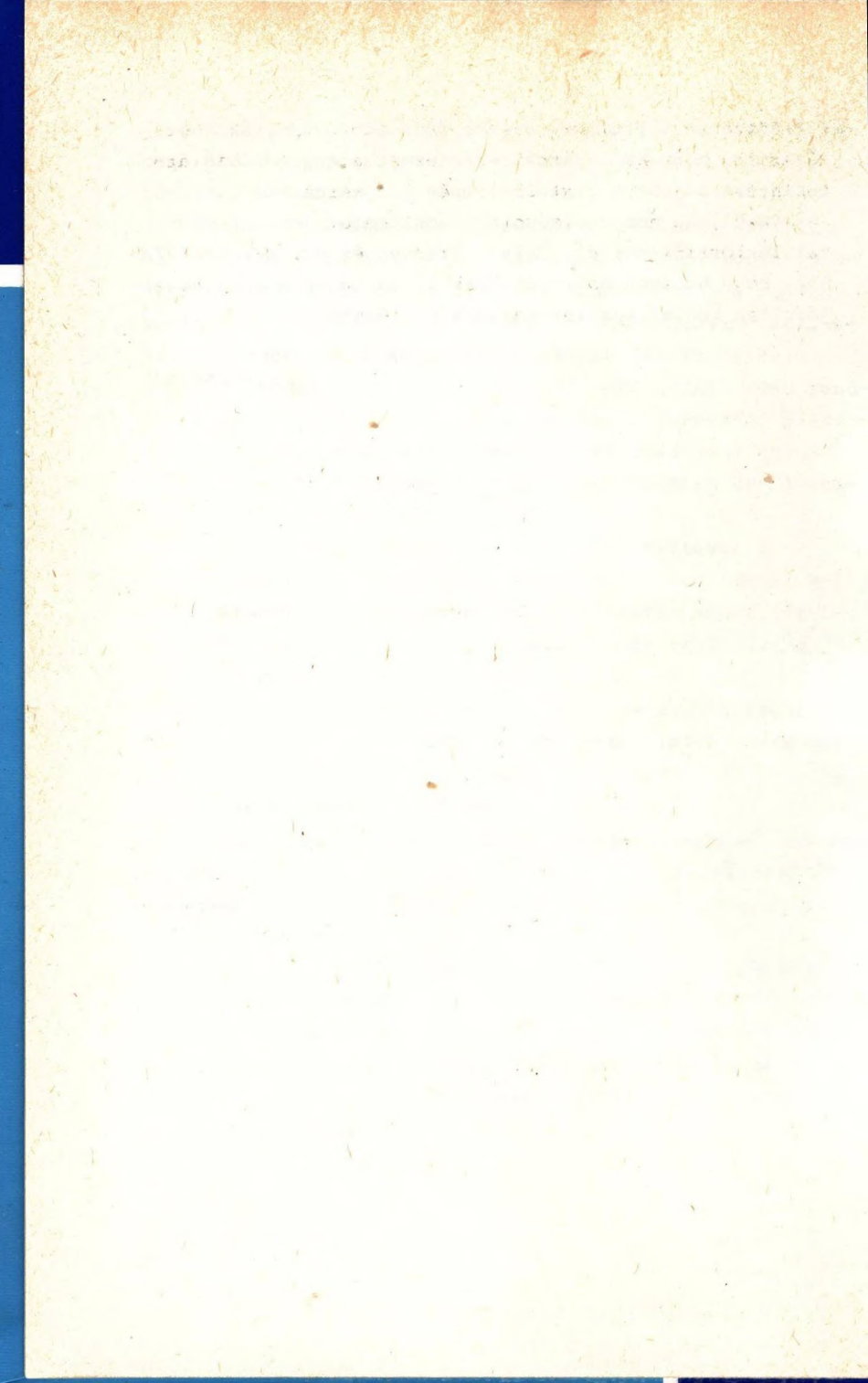
Ilyen nagy rendszerek kidolgozását természetesen sok kisebb előzi meg. Az előkészítés során kísérletező és utkereső munkával, ötleteivel *Kovács Györgyné* nyújtott sok segítséget.

Köszönetünket fejezzük ki a *BME Műmőki Továbbképző Intézet* vezetőségének, hogy az ott megtartott előadásain anyagát jegyzet formájában is megjelentethettük [A-1] és ebben a könyvben is felhasználhattuk. Külön is megköszönjük *dr. Főglein Jánosnak*, hogy anyagára jelentős mértékben támaszkodhattunk.

Köszönet illeti *dr. Csetverikov Dimitrijt*, aki hozzájárult, hogy kandidátusi értekezése alapján rövid áttekintést adjunk a texturaelemzés alapkérdéseiről.

Végül, de nem utolsósorban köszönetet mondunk szakmai lektorainknak: *dr. Gulyás Ottónak* és *dr. Mérés Lászlónak*, hogy hasznos észrevételeikkel és tanácsaikkal hozzájárultak könyvünk színvonalának emeléséhez.

A szerzők



TARTALOMJEGYZÉK

1. <u>ALAPISMERETEK</u>	19
1.1 Bevezetés (ÁG)	19
1.1.1 Történeti áttekintés	19
1.1.2 A számítógépes képfeldolgozás modellje	21
1.1.2.1 Számítógépes grafika	23
1.1.2.2 Digitális képfeldolgozás	25
1.1.3 A látáselmélet alapjai	33
1.1.3.1 Az emberi látás agyi mechanizmusai	33
1.1.3.2 Szürkelátás	39
1.1.3.3 Szinlátás	42
1.1.3.4 Alakzatlátás	49
1.2 Digitális kép létrehozása és megjeleni- tése (ÁG)	53
1.2.1 Mintavételezés	55
1.2.2 Kvantálás	61
1.2.3 Visszaállítás (rekonstrukció)	65
1.3 A képfeldolgozás sajátos eszközei (SZJ)	71
1.3.1 Számítógépek	72
1.3.2 Speciális háttértárak	74
1.3.3 A képmű	75
1.3.4 Képfelvevő eszközök	78
1.3.5 Képmegjelenítő és képrögzítő eszközök	82

1.3.6	Interaktív eszközök	87
1.3.7	Programeszközök	89
2.	<u>KÉPJAVÍTÁSOK (SZJ)</u>	91
2.1	Elvi áttekintés	93
2.2	Világosságkód-transzformációk	95
2.2.1	A hisztogram	95
2.2.2	Skálázások	98
2.2.3	Képvágások	104
2.2.4	Hisztogramtranszformációk	107
2.3	Zajelnyomás	112
2.3.1	Simitás a frekvenciatartományban	112
2.3.2	Simitás a képtartományban	119
2.4	Élkiemelés	125
2.4.1	Élkiemelés a frekvenciatartományban	126
2.4.2	Élkiemelés a képtartományban	131
2.5	Többcsíves képek javítása	134
2.5.1	Szinkorrekciók	135
2.5.2	Képscívek közötti műveletek	140
3.	<u>GEOMETRIAI KORREKCIÓK (HCS)</u>	143
3.1	Elvi áttekintés	144
3.2	Koordinátatranszformációk	146
3.2.1	Lineáris koordinátatranszformációk	147
3.2.1.1	Elemi koordinátatranszformációk	148
3.2.2	Homogén koordináták	151
3.2.2.1	Eltolás és perspektív transzformáció	152
3.2.3	Nemlineáris koordinátatranszformációk	154
3.3	A gyakorlati megvalósítás problémái	155
3.3.1	A hatékonyságnövelés elvi lehetőségei	156

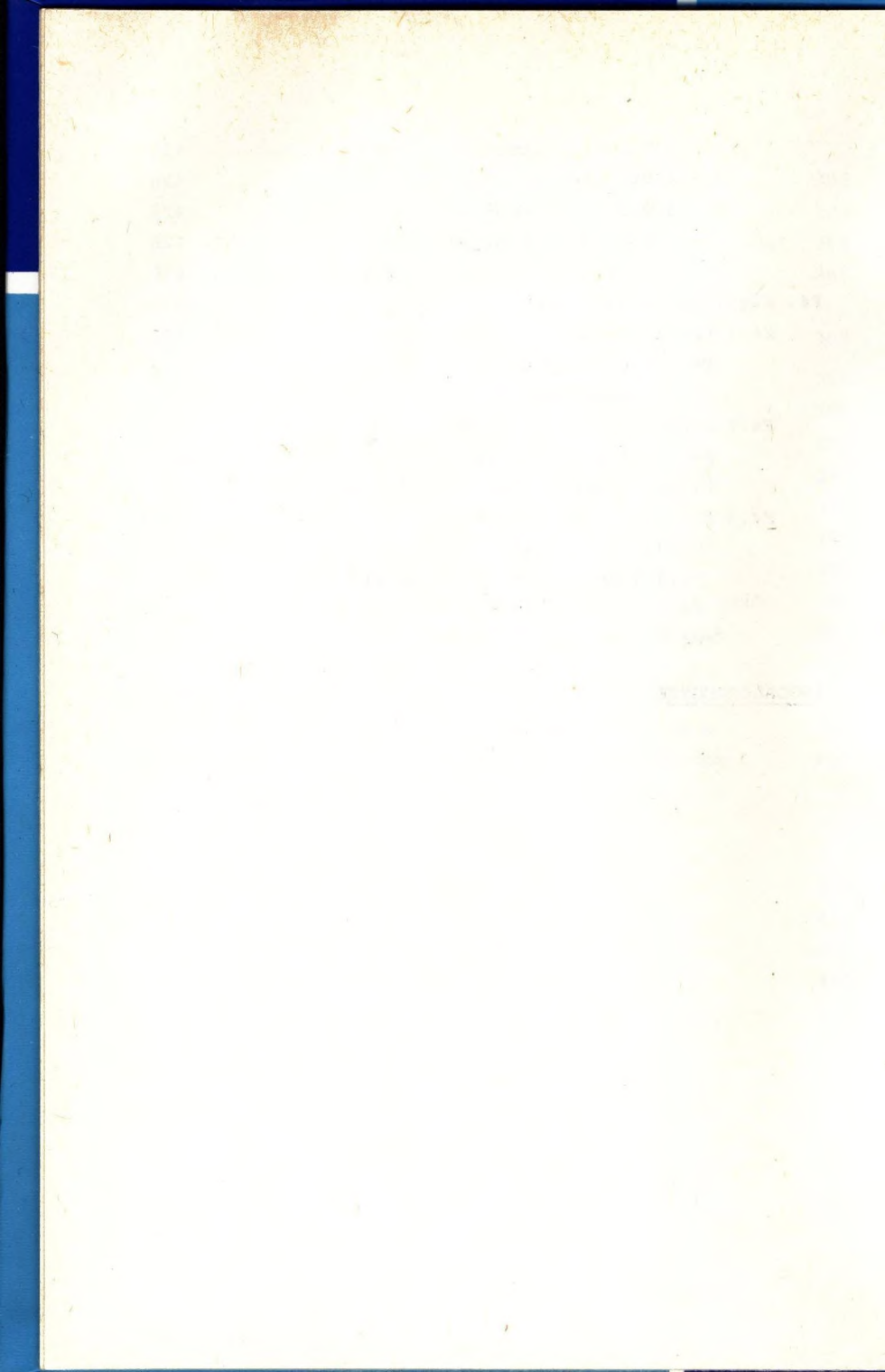
3.3.2	A korrekciós összefüggés értelmezése	158
3.3.3	A geometriai korrekció szervezése	159
3.3.4	A korrekciós adathalmaz szűkítése	160
3.3.5	Felületelemes közelítés	162
3.3.6	Diszkrét egyenes szakasz létrehozása	165
3.3.6.1	Rácsmetszéses módszer	166
3.3.6.2	Rekurzív megvalósítás	168
3.3.6.3	Számolás egész értékű ordinátákkal	169
3.3.6.4	Adaptív megvalósítás	171
3.3.6.5	Egyenesgenerálás fixpontos műveletekkel	171
3.3.6.6	Szakaszok ismétlésének módszere	174
3.3.6.7	Egyéb módszerek	175
3.3.7	Felületelem transzformációja	176
3.3.8	Ujra-mintavételezés	182
3.3.9	Adaptív geometriai korrekció	186
3.4	A korrekciós összefüggés meghatározása	187
3.5	Gyakorlati megoldások	195
3.5.1	Irodalmi áttekintés	196
3.5.2	Megvalósítási eredmények	200
4.	<u>SZEGMENTÁLÁS (ÁG)</u>	205
4.1	Elvi áttekintés	206
4.2	Foltkeresés	211
4.2.1	Globális foltkereső eljárások	212
4.2.1.1	Hisztogramparticionálás	213
4.2.1.2	Közelítés normális sűrűségfüggvényekkel	216
4.2.1.3	Küszöbbeclés entrópiafüggvényyel	220
4.2.1.4	Szegmentálás az eloszlásfüggvény alapján	222

4.2.1.5	Iteratív globális eljárások	228
4.2.1.6	Foltkeresés többsávós képeken	229
4.2.2	Lokális foltkereső eljárások	235
4.2.2.1	Összefüzési módszerek	236
4.2.2.2	Területnövesztés	237
4.2.2.3	Szétvágás és egyesítés	239
4.2.2.4	A ferde síkok ("palatető") módszere	240
4.2.3	Hibrid módszerek	243
4.2.3.1	Foltkeresés a gradiensmező felhasználásával	244
4.2.3.2	Sikklaszterezés	248
4.3	Foltelemzés	250
4.3.1	Foltszétválasztás	251
4.3.1.1	Hámozás (erózió)	251
4.3.1.2	Szintvonalas vágás	254
4.3.2	Mintaillesztés	256
4.3.2.1	Numerikus eljárások	257
4.3.2.2	Grafikus eljárások	259
4.3.3	Terület szerinti osztályozás	261
4.3.3.1	Foltszámlálás soronkénti növesztéssel	261
4.3.3.2	Foltszámlálás szomszédság- elemzéssel	265
4.4	Élkeresés	267
4.4.1	Élkitűzés	268
4.4.1.1	Gradiensmódszerek	271
4.4.1.2	Iránykereső módszerek	277
4.4.1.3	Vonalkereső módszerek	284
4.4.1.4	Élkijelölés texturaelemzéssel	290
4.4.2	Élkorrekció	293
4.4.2.1	Analitikus módszerek	294
4.4.2.2	Vonalvékonyítás	303
4.4.2.3	Relaxációs módszerek	310
4.4.2.4	Heurisztikus módszerek	316

5. <u>KÉPOSZTÁLYOZÁS</u>	~ 324
5.1 Elvi áttekintés (ÁG)	325
5.2 Statisztikus alakfelismerés (KD)	328
5.2.1 Az osztályozás előkészítése	329
5.2.1.1 A sajáttságvektorok fontosabb típusai	329
5.2.1.2 Dimenziócsökkentés	330
5.2.1.3 A döntési feladat megfogalmazása	337
5.2.2 Bayes-döntéselmélet	339
5.2.2.1 Bayes-tipusu döntések	339
5.2.2.2 A normális eloszlás esete	343
5.2.2.3 Paraméterbecslés	345
5.2.3 Lineáris diszkriminánsfüggvények	347
5.2.3.1 Döntés diszkriminánsfüggvények alapján	347
5.2.3.2 Lineáris szeparálás	349
5.2.3.3 Lineáris diszkriminánsfüggvények előállítása	351
5.2.4 Távolságmérő módszerek	357
5.2.4.1 Legközelebbi szomszéd módszerek	358
5.2.4.2 A dobozmódszer	360
5.2.5 A potenciálfüggvény-módszer	361
5.2.6 Klaszterezés	362
5.2.6.1 Hierarchikus módszerek	363
5.2.6.2 Klaszterezés célfüggvény alapján	365
5.2.6.3 Iteratív módszerek	368
5.2.7 A környezeti információk figyelembevétele	375
5.3 Szintaktikus alakfelismerés (KD)	377
5.4 Texturaelemzés (ÁG)	379

5.4.1	Texturális sajáttságvektorok előállítása	380
5.4.2	Homogén texturák vizsgálata	383
5.4.2.1	A texturák alaptulajdonságai	384
5.4.2.2	Elemzési módszerek	385
<u>FÜGGELÉK</u>		388
F1.	Mátrixalgebra (HCS)	390
F1.1	A mátrix fogalma és főbb típusai	390
F1.2	Műveletek mátrixokkal	393
F1.2.1	Összeadás, kivonás és szorzás	393
F1.2.2	Osztás négyzetes mátrixszal	394
F2.	Valószínűségszámítás (KD)	397
F2.1	Alapfogalmak és alaptételek	397
F2.1.1	Eseménytér és valószínűségi mező	397
F2.1.2	Valószínűségszámítási tételek	402
F2.2	Valószínűségi változók	404
F2.2.1	Nevezetes eloszlások	406
F2.2.1.1	Diszkrét eloszlások	406
F2.2.1.2	Folytonos eloszlások	407
F2.2.2	Valószínűségi változók teljes függetlensége	410
F2.3	Statisztikai jellemzők	411
F2.3.1	Momentumok	411
F2.3.2	Medián	413
F2.3.3	Korrelációs együttható	413
F3.	Diszkrét geometria (HCS)	415
F3.1	Topológiai jellemzők	416
F3.1.1	Szomszédság	416
F3.1.2	Utvonal	417
F3.1.3	Összefüggőség	419
F3.2	Geometriai jellemzők	420
F3.2.1	Diszkrét metrikák	420

F3.2.1.1	Diszkrét normák	425
F3.2.2	Diszkrét egyenesek	428
F3.2.3	Diszkrét görbék	428
F3.2.3.1	Görbejellemzők	428
F3.2.3.2	Diszkrét görbék előállítása	430
F4.	Függvénykalkulus (ÁG)	435
F4.1	Lineáris függvényoperátorok	436
F4.1.1	A Dirac-delta	436
F4.1.2	Konvolúció	439
F4.2	Differenciáloperátorok	442
F4.2.1	A gradiens és a Laplace-operátor	442
F4.2.2	Diszkrét differenciaoperátorok	445
F4.3	Függvénytranszformációk	449
F4.3.1	Ortogonalis függvényrendszerek	450
F4.3.2	Fourier-transzformáció	452
F4.3.3	Hadamard-transzformáció	455
F4.3.4	Haar-transzformáció	460
	<u>IRODALOMJEGYZÉK</u>	461



1. ALAPISMERETEK

Ezt a fejezetet az olvasás megkönnyítése, az egységes szóhasználat és a rendszerezés kedvéért iktatjuk be. Az 1. pontot rövid történeti áttekintéssel kezdjük, majd körülhatároljuk témaköreinket. Tekintve, hogy a számítógépes képfeldolgozás képanyagát igen nagy részben *természetes képek* (valóságos tárgyak optikai vagy elektronoptikai úton nyert síkbeli leképzései) alkotják, és sokszor van szükség emberi közreműködésre, úgy gondoljuk, nem lesz érdektelen, ha röviden összefoglaljuk az emberi látórendszer működését is. Mivel könyvünkben kizárólag *síkbeli, digitális* képek feldolgozásával foglalkozunk, a 2. pontban felelevenítjük a létrehozásukkal és a megjelenítésükkel kapcsolatos tudnivalókat. Végül a 3. pontban röviden ismertetjük a képfeldolgozás speciális eszköztárát.

1.1 BEVEZETÉS

1.1.1 TÖRTÉNETI ÁTTEKINTÉS

A képi (vizuális) információ *számítógépes* feldolgozásának mintegy 30 éves multja van. A problémák az 50-es évek végén magas szinten művelt *optikai képfeldolgozás* területein merültek fel. Már korábban is nyilvánvaló volt ugyanis, hogy analóg és kézi módszerekkel bizonyos esetekben (pl. a rádiócsillagászati, atomfizikai vagy radar-

felvételek kiértékelése) egyszerűen nem lehet a feladatokat megoldani; részben a pontossági igények, részben a képanyag nagy mennyisége miatt.

A gépi feldolgozáshoz a képeket *digitalizálni* kell, ezt már a tv-technikából tudták. Ott azonban csak a képet bontják fel képpontokra, az egyes pontokhoz tartozó amplitudót nem kvantálják. Bár a mintavételezés elvi kérdéseire már a 30-as évek elején születtek megoldások (lásd pl. [1-29]), az információelmélet egzakt megalapozását Shannon csak 1949-ben végezte el [1-40].

Az ötvenes évek elején már adva voltak a műszaki feltételek a - mai szóhasználat - analóg-digitális (A/D) átalakítók kifejlesztéséhez, amelyek az amplitudót is kvantálják (lásd 1.2.2 alpont), s így a mai értelemben vett *digitális képet* hoznak létre. (Optimális kvantálásról olvashatunk pl. [1-41]-ben, ezt azonban nyilvánvalóan sokéves kutatási munka előzte meg.) Az ötvenes évek második felére a számítógépek is elérték azt a fejlettségi szintet (művelési sebességben, tárkapacitásban, megbízhatóságban), amely lehetővé tette a nagy műveletigényű képfeldolgozási feladatok megoldását. Az ilyen irányú munka már az 50-es évek közepén megkezdődött, az első cikkek egyike pl. egy szűrési eljárásról számol be [1-6]. Vagyis elmondhatjuk, hogy a képfeldolgozás a számítógépek első alkalmazási területei közé tartozik.

Az azóta eltelt, közel 30 esztendő 3 szakaszra lehet osztani (lásd pl. [A-1], 9-10. old.):

1. Az első időszakban - amely kb. a 60-as évek végéig tartott - elsősorban a felmerülő *konkrét feladatok* megoldása volt a cél. Rendszerint *analóg* jelfeldolgozási módszereket adaptáltak, vagy *heurisztikus* eljárásokat fejleszttek ki. Általánosításra alig törekedtek, alkalmazhatóságuk feltételeit és érvényességi körüket ritkán vizsgálták.

2. A 70-es évtizedet - ezzel szemben - a *matematikai megalapozottság igénye* jellemzi. Számos feladatra dolgoztak ki egzakt vagy matematikailag precíz közelítő eljárásokat. Megkezdődött a rendszerező munka; az elért eredményeket összefoglaló művek sora ma is tart. (Lásd pl. [A-2]-[A-10], [A-12]-[A-17], [F-1] stb.) Az évtized második felétől megjelentek az első - alacsony szintűnek nevezett - *univerzális látórendszerek*, amelyeknek kialakulásához a hardver eszközök nagy ütemű fejlődése is hozzájárult. Hosszu ideig tartotta magát az a nézet, hogy ezt a problémát próbálgatással is meg lehet oldani, az elért eredmények heurisztikus felhasználásával.
3. A harmadik szakaszt a 70-es évtized utolsó éveitől számíthatjuk. Addigra nyilvánvalóvá vált, hogy a *látás automatizálásához* elengedhetetlen az elméleti alapok kellő kiépítése. Az akkoriban kezdődött kutatómunka jelenleg is folyik, és újabb lendületet kapott az 5. generációs rendszertervek megjelenése óta. Ennek ellenére a képfeldolgozás *egységes és általános elmélete még nem született meg.*

1.1.2 A SZÁMITÓGÉPES KÉPFELDOLGOZÁS MODELLE

A számítógéppel megoldott képfeldolgozási feladatok skálája nagyon széles, együttes tárgyalásuk reménytelen volna. Ezért a továbbiakban *nem* foglalkozunk sem optikai, sem más analóg módszerekkel; sem olyanokkal, amelyekben a térbeliségre vonatkozó információkat is felhasználnak (sztereo- vagy kislátószerű (range) felvételek kiértékelése stb.).

Számítógépes képfeldolgozáson kizárólag kétdimenziós (síkbeli), digitális képek (lásd 1.2 pont) létrehozását, illetve feldolgozását értjük.

A kép vizuális információtartalmát *jelenetnek*, ebben a lényeges információt hordozó képpontokat *értékes*, a többi *háttérpontnak* nevezzük. (Hangsúlyozzuk, hogy ez a megkülönböztetés a feladattól függően változhat. Jellemző példaként megemlítjük, hogy pl. egy ürfelvételen egy vízfelületi pont vizgzaldalkodási kiértékelés esetén értékes, míg mezőgazdasági szempontból háttérpontnak fog minősülni.)

A lényeges képi információt nem az egyes értékes képpontok, hanem bizonyos csoportjaik, az *objektumok* hordozzák. Ezekhez olyan jellemző tulajdonságokat: *sajátságokat* (features) rendelhetünk, amelyekkel a képpontok önmagukban nem rendelkeznek. Az objektumokat sajátágaik szerint *osztályokba* sorolhatjuk. (A háttérpontok összességét háttérosztálynak vagy egyszerűen csak háttérnek hívjuk.)

Az objektumok kapcsolatait, egymáshoz való viszonyukat a kép *strukturája* fejezi ki. Ennek feltárásakor ismét csak olyan magasabb szintű információkhoz jutunk, amelyeket az objektumok önmagukban nem hordoznak.

A számítógépes képfeldolgozás két fő ága a *számítógépes grafika* és a *digitális képfeldolgozás*. Ezek célkitűzéseikben és módszereikben is lényegesen eltérnek egymástól; bár azonos vagy hasonló eljárásokat is használnak, és a képleírásban azonos elveket követnek. Mindkét ágon jól megkülönböztethetjük a feldolgozás *3 szintjét* (lásd pl. [1-2], [1-26], [1-28]), amelyek között azonban mégsem lehet éles határvonalakat húzni:

1. A *fizikai szinten* (signal level) a képet *képpontok* halmazaként kezeljük. A képi információt az egyes képpontokhoz rendelt *számkódok* (világosság-, il-

letve szinkódok, lásd 1.2.2 alpont) hordozzák.

A feldolgozás során a *bemenő képből kimenő képet* állítunk elő.

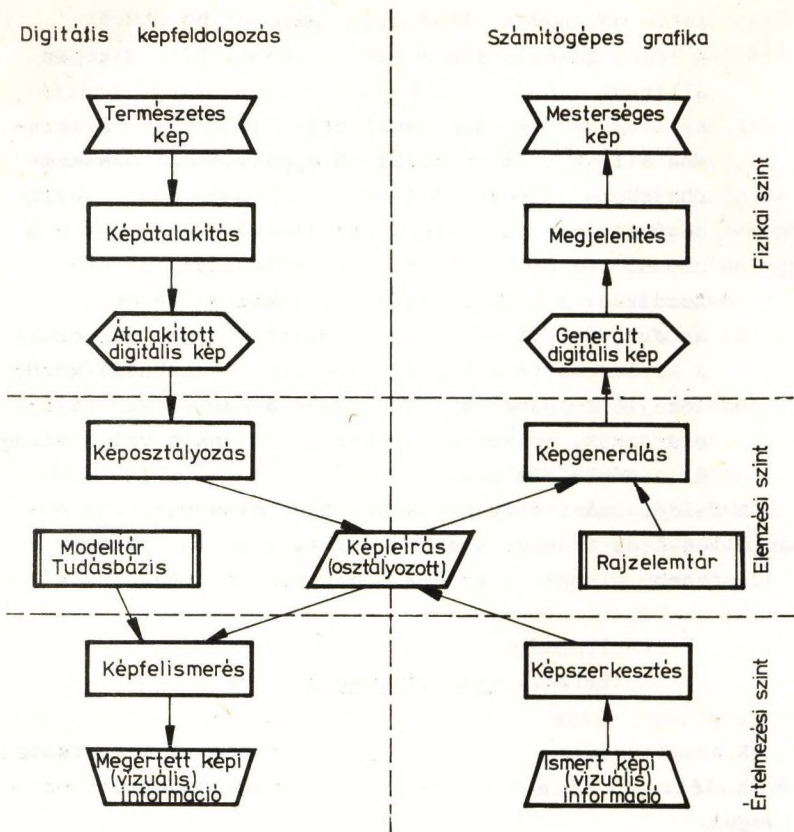
2. Az *elemzési szinten* (evaluation level) az *objektumok* állnak a feldolgozás középpontjában. Számszerűsíthető lényeges jellemzőik, a sajátosságok segítségével *leírjuk* a képet. Az elemzés célja, hogy a bemenő képből elkészítsük a képleírást, illetve fordítva: a bemenő leírásból a kimenő képet.
3. Az *értelmezési szinthez* (semantic level) tartoznak a képleírás és a kép felismerése (megértése) között lezajló folyamatok. Ezek során a kép strukturáját elemezzük, és kapcsolatokat teremtünk a valós világ és a képleírás között.

A feldolgozási folyamat a két ágon bizonyos értelemben ellentétes irányú, s ez a képleírásban is tükröződik. Részletesebb elemzését az 1-1. ábra alapján végezzük el.

1.1.2.1 Számítógépes grafika

A *számítógépes grafika* (computer graphics) mesterséges képek létrehozására, *szintézisére* (image construction) irányul.

1. A folyamat ismert vagy elképzelt képből (esetleg egy természetes kép absztrakciójából) indul az *értelmezési szinten*. A konstruktőr - előzetes ismeretei alapján - *megszerkeszti* a képet, vagyis meghatározza az alkotóelemeit és kapcsolataikat, majd ezekkel *leírja a képet* (értelmezés → leírás leképzés). A leírás szintaktikus jellegű, készülhet akár kézi, akár gépi módszerekkel. Tulajdonképpen egy grafikus nyelven írt programnak lehet tekinteni, amelynek alapján egy *vonalas kép* megrajzolható.



1-1. ábra: A számítógépes képfeldolgozás modellje

(A vonalas kép egyenes és/vagy görbe vonalakkól és ezekkel határolt idomokból áll, amelyek természetesen színesek is lehetnek. Megfelelő vonalkázással sík vagy térbeli felületeket is lehet ábrázolni.)

2. Az *elemzési szinten* a számítógép - egy *rajzelem-tár* felhasználásával, amely az objektumok megjelenítési szabályait tartalmazza - *előállítja* (generálja) a digitális képet. Ez raszteres megjelenítés esetén képpontsorozatokból, vektoros megjelenítés esetén vonaldarabokból áll.
3. A *fizikai szinten* állítjuk elő a digitális képből a megjelenítéshez szükséges analóg jeleket (digitális-analóg (D/A) átalakítókkal). A megjelenített képet a konstruktőr vizuálisan ellenőrzi, s ha nem felel meg az eredeti elképzeléseknek, szükség szerint javítja a képleirást.

A grafikus rendszerek az ember-gép kapcsolatban első-sorban *kimeneti* funkciót valósítanak meg. Feltehető azonban, hogy a jövőben széles körben el fognak terjedni a többé-kevésbé korlátozott látási funkcióval is bíró rendszerek, amelyeknek bemenő adatai mesterséges képek is lehetnek. (Pl. kapcsolási rajzból kiinduló nyomtatott áramkörti kártyatervező rendszer stb.)

A számítógépes képfeldolgozásnak ezzel az ágával a továbbiakban nem foglalkozunk (részletesebben lásd pl. [A-11]).

1.1.2.2 Digitális képfeldolgozás

A *digitális képfeldolgozás* során arra törekszünk, hogy a *természetes képek* (vagy más képi bemenő adatok) elemzése révén *fokozatosan megértsük*, azaz helyesen *értelmezzük* a képben foglalt vizuális információkat; másszóval: felismerjük, hogy mit ábrázol a kép. A távlati végső *cél* a *látás automatizálása* (azaz a számítógépes látás (computer vision) megvalósítása), olyan "*látórobotok*"

létrehozása, amelyek a képi információ feldolgozásával bonyolult vezérlési és egyéb feladatokat tudnak automatikusan megoldani ([1-2], [1-3]).

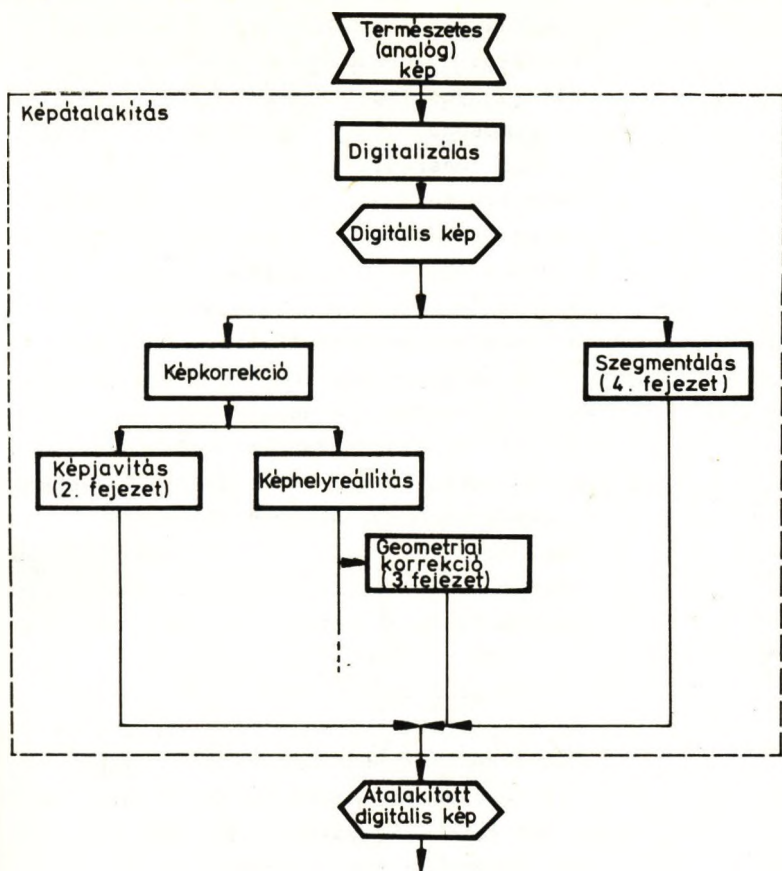
1. A folyamat fizikai szintjéhez tartozó módszereket és eljárásokat *képátalakításoknak* nevezzük. (A korábban szokásos "szűkebb értelemben vett számítógépes képfeldolgozás" elnevezés bonyolult és félreérthető, ezért nem használjuk.)

A *kiindulás* valamilyen természetes analóg kép, amelyet először is digitalizálni kell (lásd 1.2 pont). A *cél* az, hogy a vizuális vagy gépi kiértékeléshez, illetve a további feldolgozáshoz az *eredetnél kedvezőbb* tulajdonságú képet nyerjünk. (Az utóbbi esetben gyakran *előfeldolgozásról* beszélünk.) A célt a lényeges képinformációkat megőrző adekvát átalakításokkal érjük el (kép → kép leképezések).

Az eljárásokat *globálisnak*, illetve *lokálisnak* nevezzük attól függően, hogy a feldolgozás során egyidejűleg az összes képpont, illetve az egyes képpontoknak egy meghatározott környezetéhez tartozó képpontok kódját értékeljük-e ki.

A szomszédos vagy közeli képpontok kódja *nem* változik *véletlenszerűen*, sőt ellenkezőleg: eloszlásuk lényeges képi információkat közvetít, kódolt formában. Ezt a kapcsolatot a továbbiakban *környezetfüggőségnek* (context) nevezzük, és élesen megkülönböztetjük a geometriai *összefüggéstől* (connectedness). (A szomszédosság és az összefüggőség definícióját lásd az F3.1 pontban). Az utóbbi időben a környezetfüggőség vizsgálata egyre több lokális eljárásban kap fontos szerepet.

A képátalakítások két fő területe a következő (lásd 1-2. ábra):



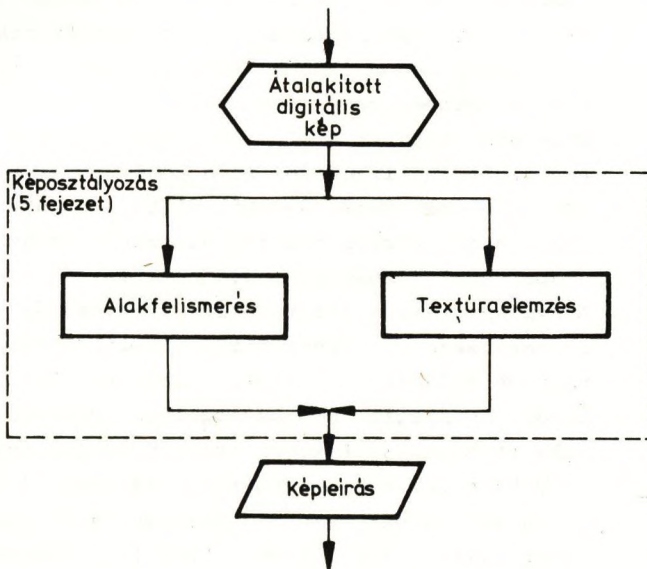
1-2. ábra. A képátalakítások területei

- a) A *képkorrekciók* célja egyrészt kijavítani a leképzési hibákat (vagy legalábbis minimálisra csökkenteni a zavaró hatásukat), másrészt kiemelni a lényeges képi információtartalmat. A korrekciós eljárásokat - célkitűzésük és módszereik szerint - 2 fő csoportba sorolhatjuk:

- A *képjavitási* (enhancement) eljárások egyrészt a leképezés során óhatatlanul elszegényedő kontrasztosság fokozására és az elmosódottság ("életlenség") csökkentésére, másrészt a "zajok" (pontoszerű vagy kis területre kiterjedő képhibák) megszüntetésére (kiszűrésére) irányulnak. Részletesen a 2. fejezetben tárgyaljuk.
 - A *képhelyreállítás* (restoration) során olyan hibátlan kép előállítására törekszünk, amelyet egy ideális (hibátlan) leképező rendszer produkált volna a kiinduló képből. Ez nagy számításigényű, bonyolult modellezési és transzformációs lépésekkel valósítható meg. Nehézkességük miatt a helyreállítási módszerek gyakorlati jelentősége kicsi, ezért általában nem foglalkozunk velük. Kivételt képez a *geometriai korrekció* (geometrical correction), amelynek célja a legkülönbözőbb okokból fellépő geometriai torzulások "visszatorzítása". Ezt a témakört fontossága és az előzőektől lényegesen eltérő módszerei miatt érdemes külön tárgyalni. A kérdéskörrel a 3. fejezetben foglalkozunk.
- b) A *szegmentálás* az értékes és a háttérpontok geometriai szétválasztását, vagyis az *objektumok elkülönítését* jelenti. Lényegében osztályozást végzünk, aminek egyik kritériuma mindig az *összefüggőség*. Az osztályozáshoz figyelembe vett sajátosság legtöbbször a világozottság-, illetve színkód; ritkábban valamilyen ezekből származtatott egyszerű geometriai, texturális vagy statisztikai jellemző

(lásd pl. [1-26]). Attól függően, hogy az osztályozáshoz hasonlósági vagy különbözőségi mértékeket használtunk-e, a szegmentálással nyert objektumok a *foltok*, illetve az *élek*. Részletes vizsgálatuk a 4. fejezet tárgya.

2. A folyamat elemzési szintjén végzett feldolgozást *képosztályozásnak* nevezzük, amelynek két fő területe az alakfelismerés és a textúraelemzés (lásd 1-3. ábra). A lényeges mozzanat mindkét területen az *objektumok sajátosságai* alapján végbemenő osztályozás. Ha egy objektumhoz n sajátosságot számítunk ki vagy mérünk meg, ezeket tekinthetjük egy n -dimenziós *sajátosságvektor* összetevőinek, amelyet viszont az ugyancsak n -dimenziós *sajátosságter*



1-3. ábra. A képosztályozás területei

(\mathbb{R}^n) helyvektoraként kezelhetünk. Végül is a képet az osztálybasorolt objektumokkal, a köztük felismert relációkkal és az osztályok jellemzőivel írjuk le (kép \rightarrow leírás leképezősek).

a) *Alakfelismerésről* (pattern recognition) beszélünk, ha az objektumokat a kép *makrostruktúrájából* származtattuk le, és az osztályozáshoz figyelembe vett sajátosságok is a makroszerkezettel kapcsolatosak. Ezen belül részletesen csak a *statisztikus alakfelismeréssel* foglalkozunk (lásd 5.2 pont), amikor is az osztályozáshoz a matematikai statisztika és a valószínűségszámítás módszereit használjuk. A *szintaktikus alakfelismerés* során a képet egyszerűbb felépítésű részekre, majd ezeket még egyszerűbbekre bontjuk, és így tovább; mindaddig, amíg tovább nem bontandó, illetve nem bontható *képelemekhez* nem jutunk. Ezek sajátosságvektorai vagy ismertek, vagy elemi uton előállíthatók. Ezek után a képet a képelemek hierarchikus rendszereként írjuk le, és a sajátosságvektorokat az elemek sajátosságaiból állítjuk össze. Ezzel a témakörrel nem foglalkozunk, de az alapelveket vázoljuk az 5.3 pontban.

b) A *texturaelemzés* (texture analysis) a kép *mikroszerkezetének* vizsgálatát jelenti. A texturát nem teljesen egzaktul, de képfeldolgozási célokra kielégítő pontossággal úgy definiáljuk, mint statisztikusan ismétlődő, (azonos vagy hasonló) *vonal-* vagy *területelemekből* felépülő, szabályos vagy véletlenszerű (non-figuratív) *mintát* ([1-16], [5-5]). Lényeges, hogy a texturaelemek elhanyagolhatóan kicsik ahhoz az alakzathoz képest, amelyet alkotnak;

s önmagukban nem észlelhetők (sokszor a lát-hatóság határa alatt is vannak). Így a lát-ványt nem "egyéni" tulajdonságaik, hanem *strukturájuk*, azaz rendezettségük és egymás-hoz való viszonyuk határozza meg. (Az elneve-zés a szövetminta analógiája: ahogyan egy anyag összeáll az - esetleg periódikusan - ismétlődő mintaelemekből, úgy áll össze a kép a texturaelemekből.) A texturaelemzés "objek-tumai" a *texturaelemek* (ha egyáltalán elkülö-níthetők); a felhasznált sajáttságvektorok pe-dig természetesen a mikrostrukturát jellemzik. A témakör érdemi tárgyalása tulmegy könyvünk keretein; az alapelveket azonban áttekintjük az 5.4 pontban. Ezen kívül megemlítünk néhány, a texturális jellemzőkön alapuló élkitüzési eljárást is (lásd 4.4.1.4 alpont).

3. A digitális képfeldolgozás legmagasabb (értelme-zési) szintje a *képfelismerés* (scene analysis). Ennek során a képleírás alapján azonosítjuk az objektumokat a valóságos tárgyakkal, és egymás-hoz viszonyított helyzetük, méreteik stb. alap-ján *felismerjük*, hogy mit ábrázol a vizsgált kép. A legnagyobb nehézséget az okozza, hogy a tárgyak általában térbeliek, s a térbeliségre vonatkozó információ - a kiinduló feltevéseink szerint - hiányzik a digitális képből. Ezért a képfelisme-réshez a világra vonatkozó további ismeretekre - *tudásbázisra* - van szükség. Ennek felhaszná-lásával vagyunk képesek megérteni a képi (vizuá-lis) információt. A legújabb kutatások szerint, a kitüözött célt legjobban a *modellbázisú* látórend-szerekkel lehet megközelíteni. Ezek egy modell-tárban tartalmazzák a "lehetséges világ" tárgyai-

nak leírását. A képfelismerés első lépése ekkor az *aktuális modell* megalkotása az alakfelismerésből kapott leírás alapján. A "felismerési" lépés lényegében abból áll, hogy megkeresik a modelltárban az aktuálishoz legközelebbi modell(ke)t, s az objektumo(ka)t a megfelelő tárggyal (tárgyakkal) azonosítják. A térbeli távolságok és méretek megbecslése azonban további nehézségeket okoz. Néhány kísérleti rendszer már működik [1-5], nálunk azonban csak laboratóriumi szintű kísérletek folynak. Ezért a képfelismeréssel a továbbiakban nem foglalkozunk.

Az 1-1. ábrán vázolt modell alapján működő és az emberit kielégítően megközelítő *univerzális látórendszert* ez ideig még *nem sikerült* létrehozni. Ez a képfeldolgozás egységes elméletének hiányán kívül két fő okra vezethető vissza:

- egyrészt még az emberi látórendszer működését sem ismerjük eléggé, így még abban sem lehetünk biztosak, hogy a megoldást ennek modellezésével kell keresnünk;
- másrészt a jelenlegi hardvereszközök sebessége és tárkapacitása - bár az utóbbi években jelentős fejlődést tapasztalhattunk - még mindig nem elég nagy.

Emiatt valószínű, hogy a jövőben a fenti modellt is tovább kell majd fejleszteni. Könyvünkben mégis ezt vesszük alapul, mivel alkalmas a jelenlegi ismeretek kielégítő rendszerezésére.

1.1.3 A LÁTÁSELMÉLET ALAPJAI

Az ember és a környezete közötti kapcsolatban a vizuális információ a legfontosabb, minden mást felülmúló tömörsége miatt. Alapvető jellemzői a *világosság*, illetve a *szín*, az *alak*, a *mélység* (térbeliség) és a *mozgás* (időbeliség).

A látási folyamat problémái nem tartoznak könyvünk tárgyához, nagy részük ez idő szerint még megközelítően sincs megoldva. (Rövid összefoglalás található [1-47]-ben.) Az alábbiakban csak a képfeldolgozás szempontjából legfontosabb kérdéseket tekintjük át vázlatosan.

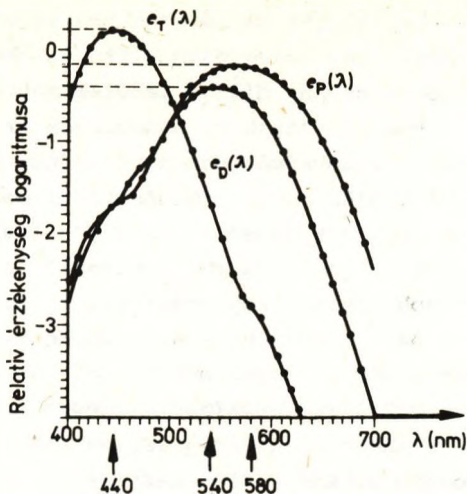
1.1.3.1 Az emberi látás agyi mechanizmusai

Az emberi látórendszer "bemeneti készüléke" a szem. Az elektromágneses sugárzásból - megfelelő fénysűrűség esetén - kicsit több, mint egy oktávnyit, a 380 nm-780 nm hullámhossztartományt (illetve a $7,89 \cdot 10^{14}$ - $3,85 \cdot 10^{14}$ Hz frekvenciatartományt) érzékeli. A látható spektrumtartományba eső elektromágneses sugárzást nevezzük *fénynek*. A *fénysűrűség* az egységnyi felületű fényforrás által a szemünk irányában az egységnyi térszögbe kisugárzott fényenergiát jelenti. Csökkenő fénysűrűség esetén a látható tartomány fokozatosan beszűkül, egészen 500-700 nm-ig.

Mint ismeretes, a *szemlencse* a látótér éles képét a recehártyában (retina) levő *látógödörben* (fovea) állítja elő, a fényrekesz szerepét betöltő *pupillán* át belépő fénysugarakból. A szemlencse középpontját a látógödörével összekötő egyenest *szemtengelynek* nevezzük, ez egyben a *látókup* tengelye is, amelynek csucsa az előbbiben van. A *látótér* nagyságát a hozzá tartozó látókup nyílásszögével szokás megadni.

A recehártyában kétféle *fényérzékelő* (receptor) található:

1. A *pálcikák* (rods) a 2° - 10° -os látótérben kisebb számban fordulnak elő, ezen kívül viszont ezek a domináns receptorok. Tulnyomó többségük a 10° - 20° látótérben helyezkedik el; számuk kb. 100-120 millió. Fontos szerepük van a kis fénysűrűségű fényforrások észlelésében (szkotopos látás): igen kis energiák és változások érzékelésére képesek. Mivel *spektrális érzékenységük* (vagyis az egységni energiájú, különböző hullámhosszu fényhatásokra adott ingerválaszuk) az érzékelési tartományban gyakorlatilag *állandó*, "sötétben" nem tudunk színeket megkülönböztetni.
2. A *csapok* (cones) főleg a 0° - 2° -os látótérben találhatóak, 2° - 10° -ig ezek a domináns receptorok, ezen kívül előfordulásuk és hatásuk elenyésző; számuk 6-6,5 millió. Nagyobb fénysűrűségek esetében lépnek működésbe (fotopos látás). Kvalitatively tudjuk, hogy spektrális érzékenységük lényegesen különböző, és erősen függ a fény hullámhosszától; a pontos összefüggések meghatározása azonban súlyos elvi nehézségekbe ütközik. A nagy számu kísérleti mérés alapján 3 típusukat sikerült megkülönböztetni, ezeket P- (protonopen), D- (deuteronopen), illetve T- (tritonopen) csapnak nevezik. Mindegyiknek van fényre aktiválódó *ingeradó* és fényre lezáródó (vagyis a fényhatás megszűnésekor aktiválódó) *gátláskeltő* változata. Feltételezett érzékenységi görbéiket az 1-4. ábrán adjuk meg [1-42]. (A függőleges tengelyen az érzékenység logaritmusát mértük fel. Az ábráról leolvashatjuk, hogy a P-típusu csapok érzékenysége 580 nm (= vörössárga, ezeket legtöbbször mégis vörösérzékelőknek nevezik), a D-típusua-



1-4. ábra. A csapok spektrális érzékenységi görbéi

ké 540 (= zöld), a T-típusuaké 440 nm (= kék) körül a legnagyobb.)

A szembe érkező fény hatására fotokémiai változások mennek végbe a fényérzékelőkben levő speciális folyadékban, a látóbiborban. Az így keletkező ingereket, illetve gátlásokat a szemben levő látóneuronok veszik át. (A neuron az idegszövet legkisebb funkcionális egysége. Működését olyan "fekete doboz"-zal modellezhetjük, amelynek több bemenete és egy kimenete van. A kimenő jel a bemenő jelek *lineáris kombinációja*; ebben a pozitív súlyukat nevezzük *ingernek*, a negatív súlyukat *gátlásnak*. A bemenő jelek rendszerint akkor hatásosak, ha értékük nagyobb egy küszöbszintnél.)

A legújabb kutatások szerint [1-22] az emberi látórendszer különböző absztrakciós képességű "feldolgozó" egységek legalább 4 szintű hierarchikus rendszere, amelyben a látóneuronok képviselik a legalsó szintet. Számuk

mindkét szemben kb. 10^6 , de eloszlásuk nem egyenletes; sűrűségük a foveában kb. megegyezik a fényérzékelőkével, és itt kb. 35-ször nagyobb, mint a látótér szélein.

Kísérletileg kimutatták, hogy a neuronok érzékelési területe - amelyet végeredményben a hozzájuk közvetlenül kapcsolódó fényérzékelők elhelyezkedése határoz meg - közel kör alakú, és egy középső *körlapra*, valamint az ezt koncentrikusan körülvevő *körgyűrűre* bontható. A fényérzékelők és a neuronok közötti kapcsolatokat pontosan nem ismerjük; az azonban biztos, hogy a körlaphoz, illetve a körgyűrűkhöz kapcsolódók *inverz* működésűek. Az ingerek, illetve a gátlások 2-4 *szinapszison* (= neuron-neuron kapcsolat) és 4 féle más típusu retinasejten keresztül érkeznek a látóneuronokhoz. Ennek következtében ezek már nem az egyes képrészletek fénysűrűségét, hanem egy képpont és a közvetlen környezete között - a szinben, fénysűrűségben vagy ezek időbeli eredőjében - mutatkozó különbséget, a *kontrasztot* érzékelik. A kontrasztérzékelést fokozzák a körgyűrűben érvényesülő "oldalsó" gátlások, amelyeket ilyen értelemben optimalizáló mechanizmusnak tekinthetünk.

A látóneuronok kimenőjeleit vezető idegrostok (*axonok*) két idegducban (ganglion) egyesülve jutnak el a második hierarchiaszintet realizáló bal, illetve jobb oldali *térdestestbe*. (Érdeemes felfigyelni rá, hogy az ingervezetés impulzus-kódmodulációval (PCM) történik, ami a jelenleg ismert legnagyobb megbízhatóságu hibavédelmet nyújtja!) A térdestestek neuronjai *6 rétegben* helyezkednek el, amelyek felváltva a bal, illetve a jobb szem azonos oldali látóteréből kapják az ingereket. (A szemlencse fordított képet ad, ezért a bal oldali térdestestbe a jobb oldali látóteréből származó információk jutnak el és fordítva.) A neuronok érzékelési területe itt is "körlap-körgyűrű" típusú, és az egyes rétegekben a felületre me-

rölelesen egymás alatt levőké gyakorlatilag a látótér azonos pontjára vonatkozik. Valószínűleg itt történik meg a *mélységi információk* elsődleges elemzése.

Az axonok egy része az *agytörzsbe* fut; ez vezérli a látással összefüggő motorikus folyamatokat. (Pl. a pupilla átmérőjének változtatása a beeső fényenergia függvényében; a szemtengelyek konvergenciájának beállítása a képtávolságtól függően stb.)

A térdestestekben képzett szinapsziszokon át az ingerületek a harmadik hierarchiaszintre: az azonos oldali *elsődleges agykérgi látómezőbe* (visual cortex), pontosabban ennek 4. rétegébe jutnak. A réteg alsó, nagyobbik részébe az ellenkező, a felső kisebb részébe az azonos oldali szemből kerülnek az információk. (Ismeretes, hogy az agykéreg jól felosztható különböző funkcionális mezőkre (látó-, halló, mozgató- stb. "központokra"); ezek számát 50-100 közöttire becsülik. Az agykéreg *Brodmann-féle térképén* az elsődleges látómező a 17-es, a magasabbrendű látással kapcsolatos további két mező a 18-as és 19-es számot viseli.)

A 4. réteg neuronjainak érzékelési területe még mindig "kőrlap-körgyűrű" típusu; a többi rétegben azonban olyan "egyszerű", illetve "összetett" neuronokat találunk, amelyek meghatározott irányu és elég keskeny *vonaldarabok* érzékelésére képesek! Ezt úgy lehet elképzelni, hogy egy-egy ilyen neuronnal a 4. réteg több olyan neuronja alkot szinapszist, amelyeknek érzékelési területe azonos típusu, és egy vonal mentén helyezkedik el.

Az agykérgi látómező egymás alatti rétegeiben (a felületre merőlegesen) található neuronok itt is egy szűk látótér szomszédos pontjaihoz kapcsolódnak. A felülettel párhuzamosan haladva azt tapasztaljuk, hogy kb. 1 mm-enként kb. 10^0 -kal változik az egyszerű neuronok irányérzékenysége; erre merőleges irányban haladva pedig felváltva

a bal, illetve a jobb szemmel kapcsolatban álló neuronokat találunk. A kb. 180°-os szögterületet lefedő neuroncsoportok periodikusan, de eddig fel nem derített szabályok szerint helyezkednek el. Az összetett neuronok hasonló irányérzékenységet mutatnak, de kevésbé érzékelik a vonaldarabok pontos helyzetét, így képesek az érzékelési területükön áthaladó, meghatározott irányú mozgó vonaldarabok felismerésére (előnyben részesítve a vonaldarabra merőleges mozgásirány egvik értelmét).

Az összetett neuronok között fordulnak elő első ízben olyanok is, amelyekhez mindkét szemből jut el információ (binokulárisak). Ezeknek is több típusuk van aszerint, hogy a két szem közül melyik a domináns és milyen mértékben.

A 6. rétegben levő neuronok visszacsatolást képeznek az oldalsó térdestestekhez. A 2. és a 3. rétegből a 18-as és a 19-es agykérgi látómezőbe, az 5. rétegből pedig a belső agyi látóközpontba jutnak a részlegesen már feldolgozott információk. Egyelőre feltételezik, de kellően nem bizonyított, hogy ezek alkotják a 4. hierarchiaszintet, ahol a látással kapcsolatos legmagasabbrendű képzeink kialakulnak. Az ezekben rejlő vizuális információkat agyunk - az említett 4 fő jellemző szerint - a korábban megszerzett ismeretei ("képi adatbázisa" és "vizuális tudásbázisa") alapján dolgozza fel.

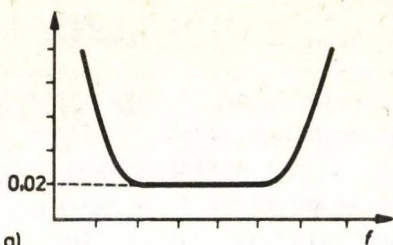
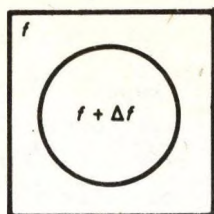
A továbbiakban csak a fotopós látással foglalkozunk, mert a képfeldolgozás szempontjából csak ennek van jelentősége. Részletesebben megvizsgáljuk, hogy hogyan alakulnak ki a kontrasztra, a színre és az alakzatokra vonatkozó képzeink. A mélységi és a mozgási képzetekkel nem foglalkozunk, mivel könyvünk anyagához csak közvetve kapcsolódnak (lásd pl. [1-15], [1-50]).

1.1.3.2 Szürkelátás

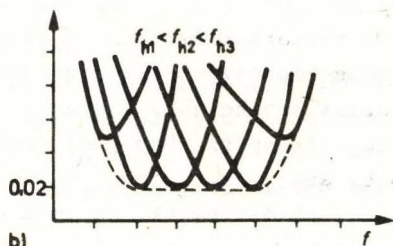
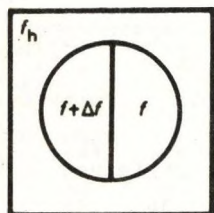
A fekete-fehér (a továbbiakban: ff) képek különböző fénysűrűségű képrészekből állnak. Mint említettük (lásd 1.1.3.1 alpont) a fénysűrűség a fényforrás által kisugárzott fényenergiával függ össze, és így objektíven mérhető. A szemünk viszont *csak szubjektív* érzékelésre képes: a nagyobb fénysűrűségű fényforrást *világosabbnak* látjuk. Világosságérzetünk azonban *relatív*; tulajdonképpen a szomszédos képrészek fénysűrűségének különbségét, a *kontrasztot* érzékeljük. A szem kontrasztérzékenysége *anizotrop* (iránytól függő) és *nemlineáris* (exponenciális). Az utóbbi azt jelenti, hogy nem abszolút fénysűrűségkülönbségeket észlelünk, hanem a változásnak az eredeti fénysűrűséghez viszonyított *arányát*.

A még észlelhető legkisebb Δf fénysűrűség-különbséget *kontrasztküszöbnek* nevezzük. Ez a mondottak szerint függ a fénysűrűségtől, de a $\Delta f/f$ *kontrasztküszöb-arány* (= Weber-féle tört) széles fénysűrűség-tartományban állandó és értéke kb. 0,02 (lásd 1-5/a ábra; [1-19]). Az állandó tartomány lényegesen beszűkül, ha a kontrasztküszöb-arányt különböző fénysűrűségű háttérben vizsgáljuk. (Lásd 1-5/b ábra; figyeljük meg, hogy a görbesereg *burkolója* az 1-5/a ábrán látható görbe! Ezzel magyarázható az a jelenség, hogy azonos szürkeségű képrészleteket világos háttérben sötétebbnek látunk, mint sötétebb háttérben (lásd 1. kép; a fekete-fehér képek a mellékletben található, a 40. oldaltól kezdődően).

A kontraszttal kapcsolatos jellemző a *képélesség*, ami a különböző világosságú képrészek közötti átmeneti tartomány (= határvonal) szélességével fordítottan arányos. (Minél szélesebb ez a tartomány, annál *elmosódottabbnak* érzékeljük a képet.) Érdekes jelenség viszont,



a)

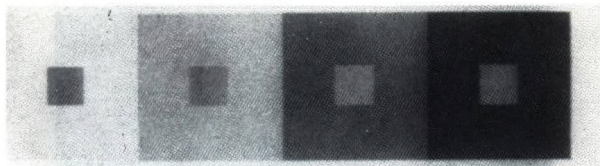


b)

1-5. ábra. A kontraszküszöb és a háttér-megvilágítás összefüggése

hogy az ugrásszerű kontrasztváltozásra szemünk "tullövés"-szerű reakcióval válaszol, ami a kontraszküszöb *lokális megváltozását* jelenti. (Az ugrás világosabb oldalán csökken, a sötétebb oldalon nő a kontraszküszöb.) Ezzel magyarázható a határvonalak automatikus kiegészítése az 1-6/a-c) ábrán ("fehér a fehérben")..

A ff képek további fontos jellemzője a *síkfrekvencia* (spatial frequency), amely a képsík egy tetszőleges irányában az egységnyi hosszra eső periodikus fénysűrűségváltozások számát adja meg. (Nem tévesztendő össze a frekvenciának az *időbeli* változásokra vonatkozó szokásos és általunk is megtartott értelmezésével.) Az *észlelhető* síkfrekvencia függ az 1^o-os látómezőbe eső periódusok számától: ff képek esetében legjobban észleljük az 5-10 periódus/fok értékeket (lásd pl. 1-7. ábra; [1-10]).



1. kép. Egyidejű kontraszthatások



a)

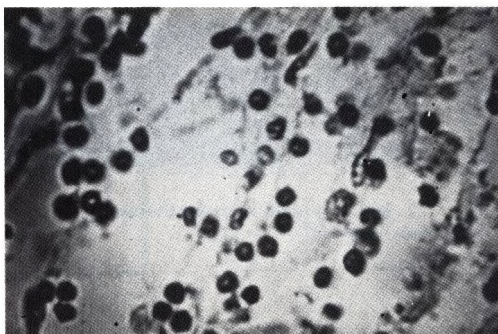


b)

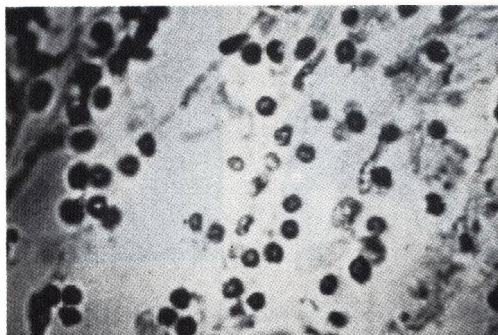


c)

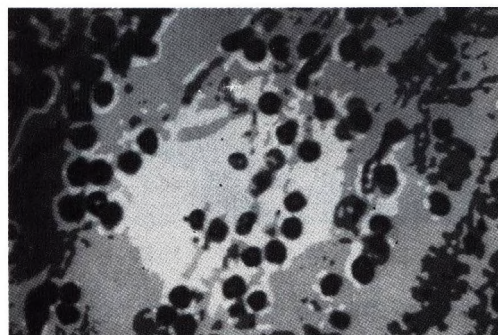
2. kép. A síkbeli és a gradációs felbontás hatása alacsony síkfrekvencia-spektrumú képen:
a) $384 \times 288 \times 6$ bitre, b) $192 \times 144 \times 6$ bitre, c) $384 \times 288 \times 2$ bitre digitalizált kép



a)

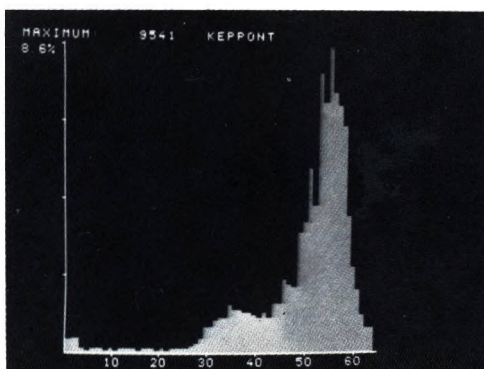


b)

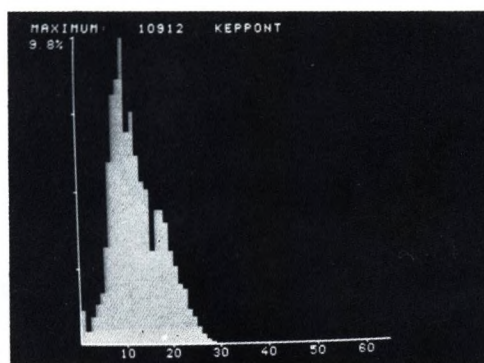


c)

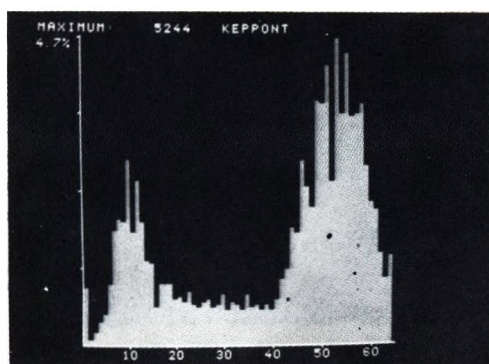
3. kép. A síkbeli és a gradációs felbontás hatása magas síkfrekvencia-spektrumú képre:
a) $384 \times 288 \times 6$ bitre, b) $192 \times 144 \times 6$ bitre, c) $384 \times 288 \times 2$ bitre digitalizált kép



a)

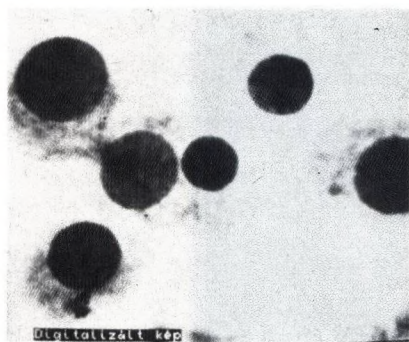


b)

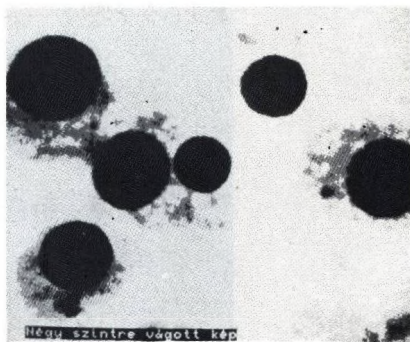


c)

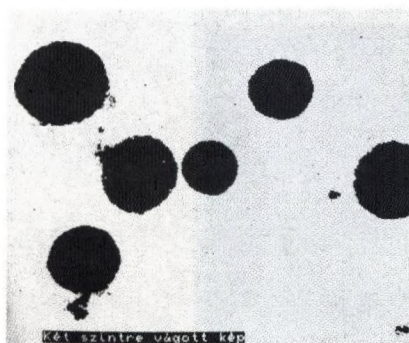
4. kép. Jellegzetes hisztogramok:
a) világos tónusú, b) sötét tónusú,
c) kétszintű kép hisztogramja



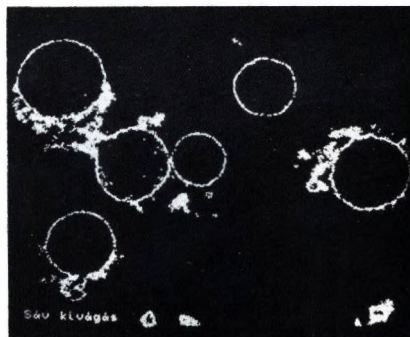
a)



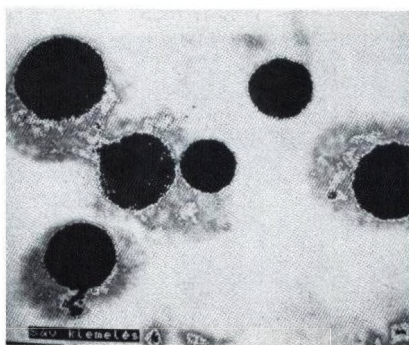
b)



c)



d)



e)

5. kép. Példák szintrevágásra:

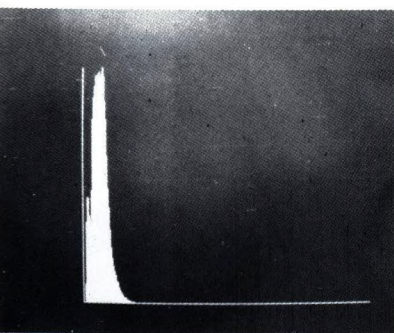
a) mikroszkópi felvétel, b) négy szintre vágás, c) két szintre vágás,
d) sáv kivágás, e) sávkiemelés



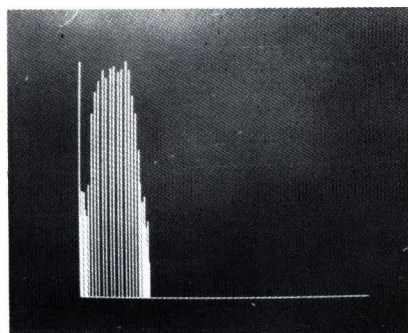
a)



b)



c)

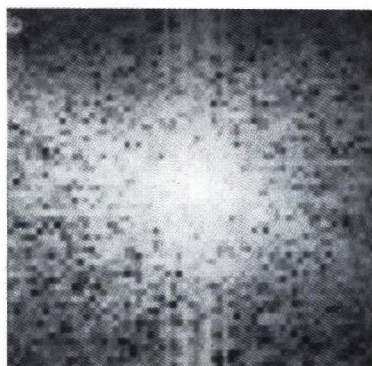


d)

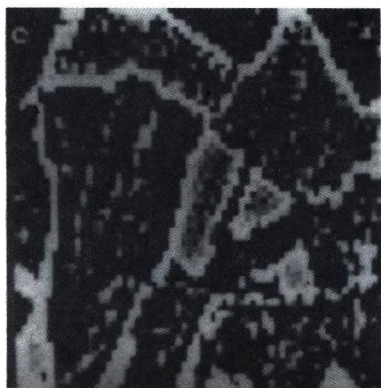
6. kép. Hisztogramkiegyenlítés hatása:
a) eredeti LANDSAT felvétel (MSS5-sáv), b) hisztogramkiegyenlítés után,
c) eredeti kép hisztogramja, d) kiegyenlített hisztogram



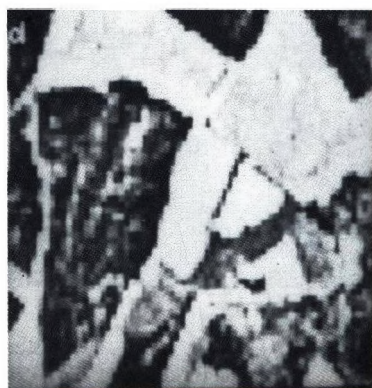
a)



b)

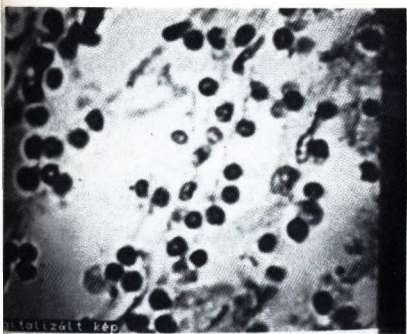


c)

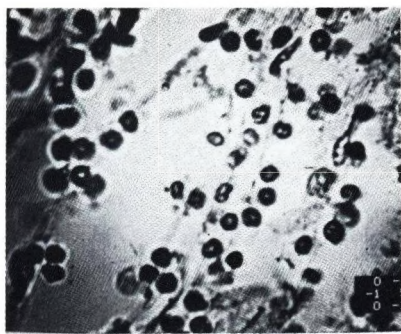


d)

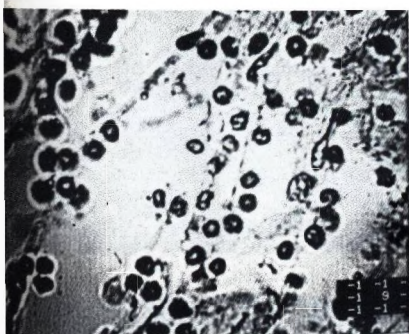
10. kép. Szűrés frekvenciatartományban:
a) eredeti kép (vö. 7/a képpel), b) Fourier-spektrum,
c) élkiemelés felülatéresztő exponenciális szűrővel, d) élkiemelés a c) képre



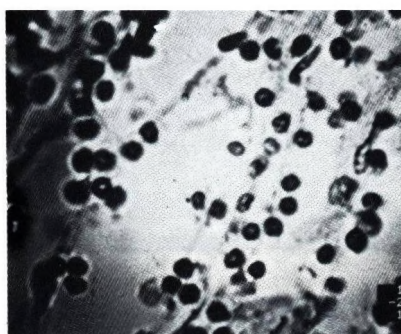
a)



b)



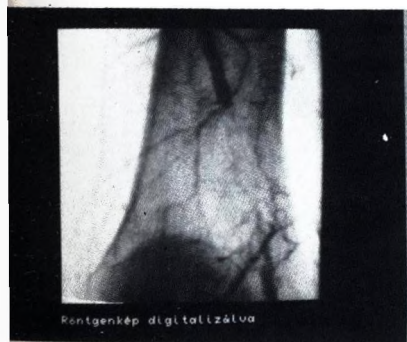
c)



d)

11. kép. Élkiemelés konvolúciós szűrőkkel:

a) eredeti kép, b) L1-szűrő hatása, c) L2-szűrő hatása, d) L3-szűrő hatása



a)



b)

12. kép. Pszédoplasztikus kép előállítása:

a) eredeti röntgenfelvétel, b) R-szűrő hatása



a)



b)

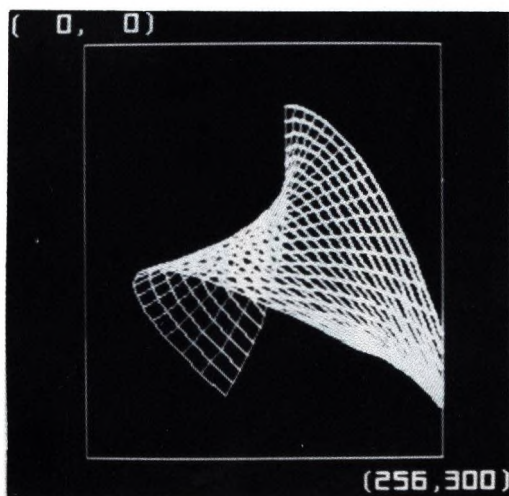


c)

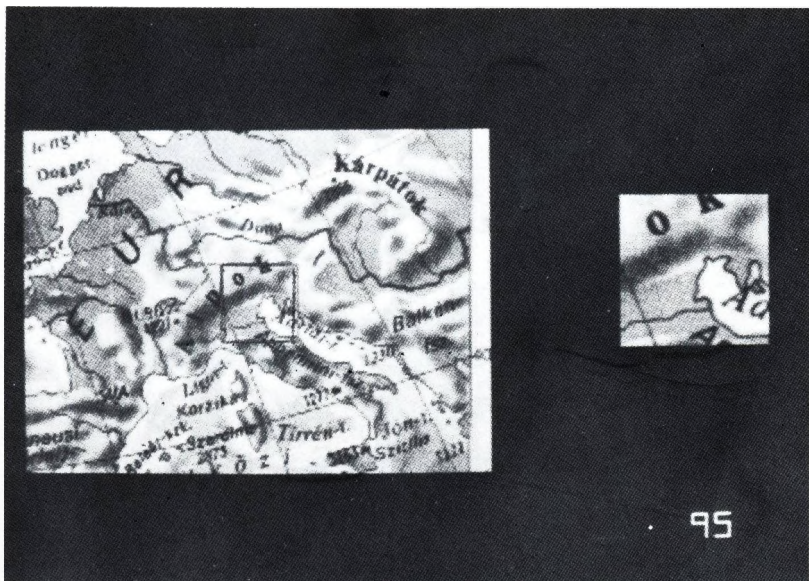


d)

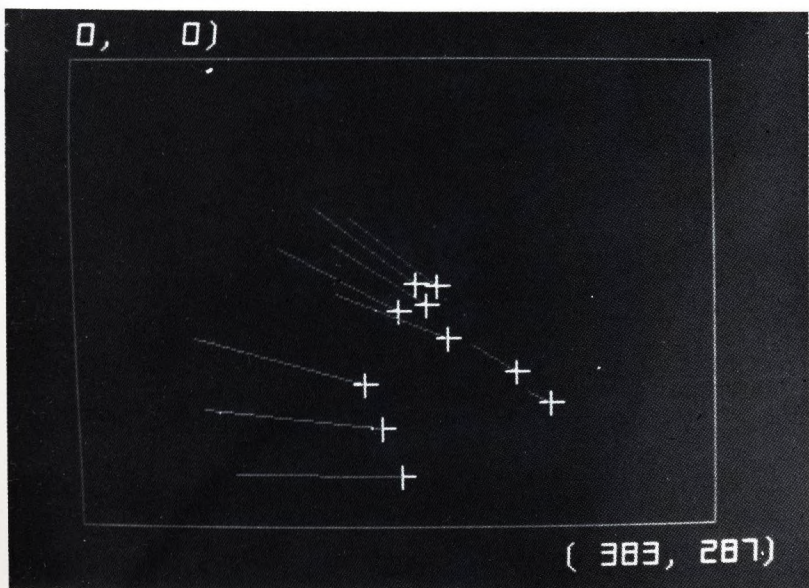
13. kép. Mintakép (LANDSAT-felvételek hisztogramkiegyenlítés után):
a) MSS4-, b) MSS5-, c) MSS6-, d) MSS7-sáv



14. kép. Geometriai korrekciót vezérlő adatok hálózata



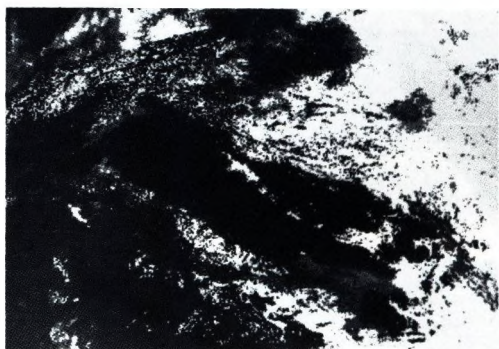
15. kép. Azonosítási pontok definiálása



16. kép. Azonosítási pontok ellenőrzése



a)

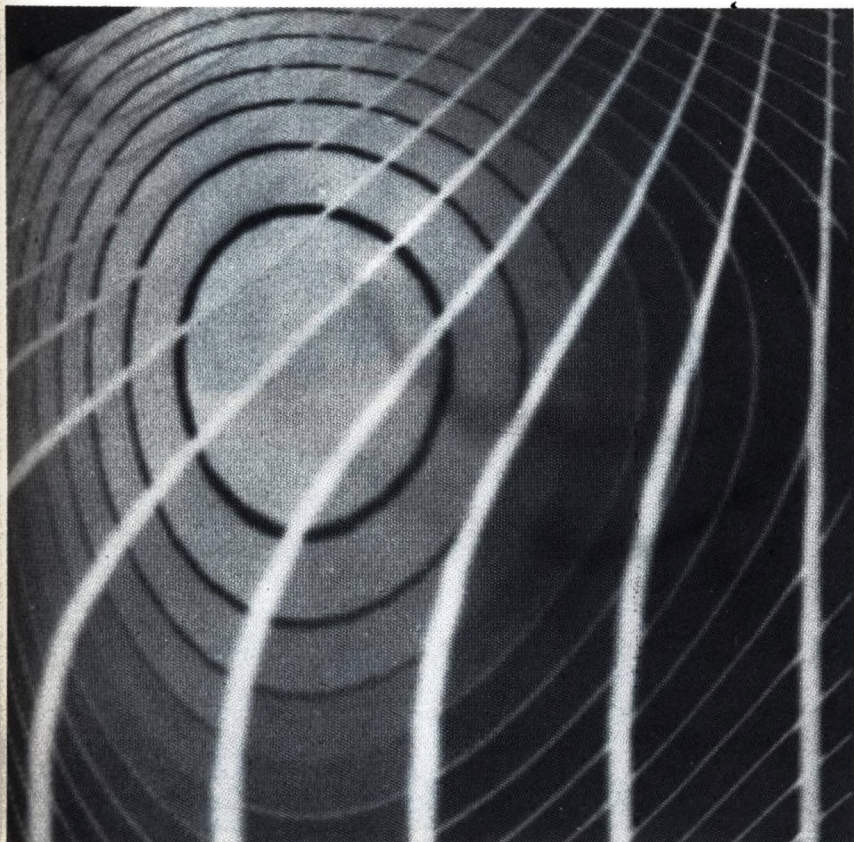


b)

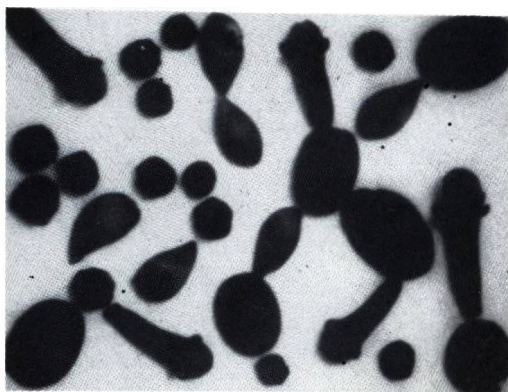


c)

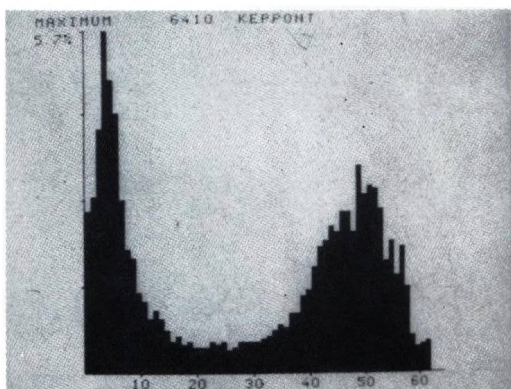
17. kép. Geometriai korrekció:
 a) bemenő kép (korrigálandó), b) referenciakép,
 c) kimenő kép (korrigált)



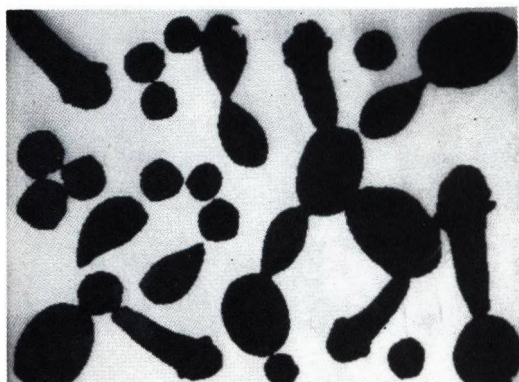
18. kép. Erősen nemlineáris, nem invertálható geometriai korrekció (a bemenőkép szabályos négyzetrács volt)



a)

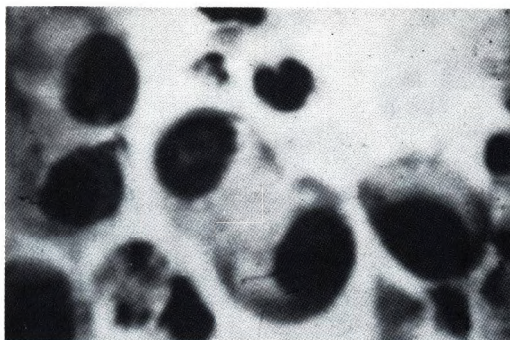


b)

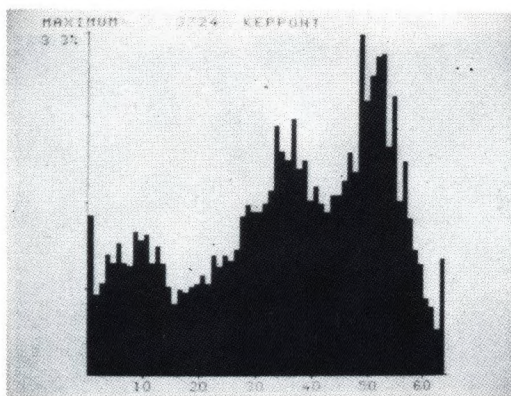


c)

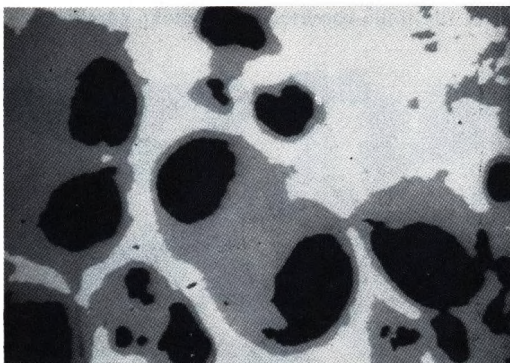
19. kép. Két szintre vágás hisztogramparticionálással:
a) tesztkép, b) a tesztkép hisztogramja, c) kétszintű kép



a)



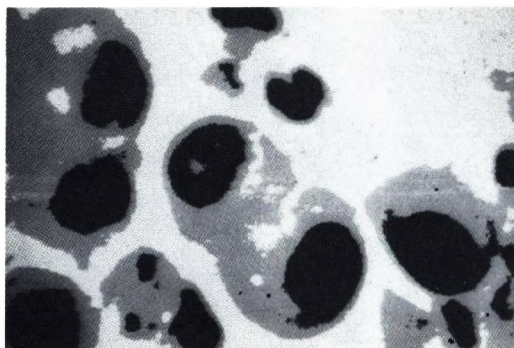
b)



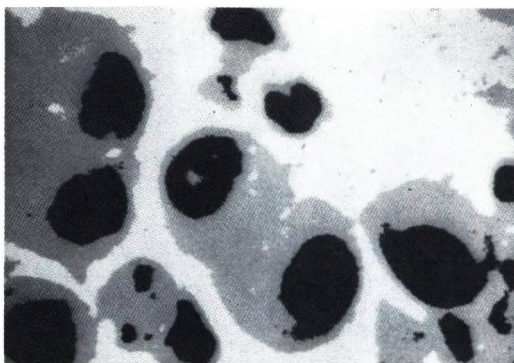
c)

20. kép. Három szintre vágás:

a) eredeti mikroszkópi sejt felvétel, b) az a) kép/hisztogramja, c) háromszintű kép

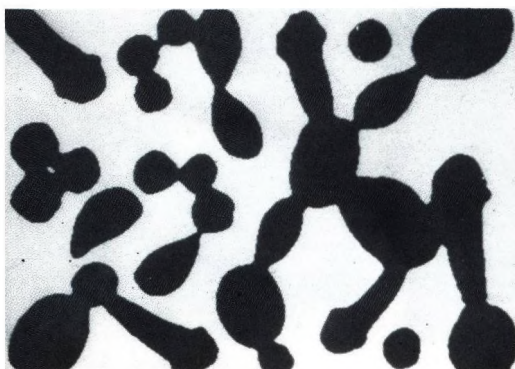


a)

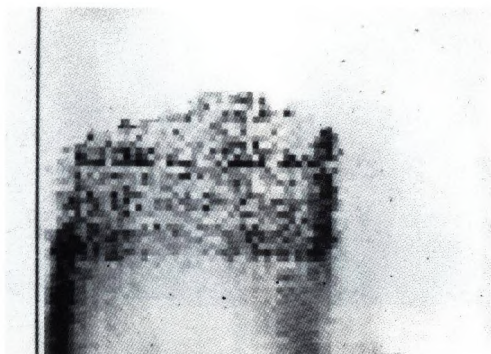


b)

21. kép. A 20/a kép három szintre vágása a görbületfüggvény alapján:
a) 10-edfokú közelítés, b) 25-edfokú közelítés



22. kép. Többségi szűrés hatása a 19/c képre

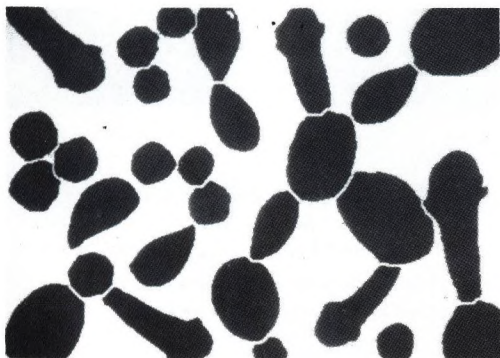


a)



b)

23. kép. A tesztkép (19/a kép) két szintre vágása:
a) gradiensthistogram, b) kétszintű kép



24. kép. A tesztkép (23/b kép) szegmentálása a szintvonalkép (lásd VI. kép) alapján



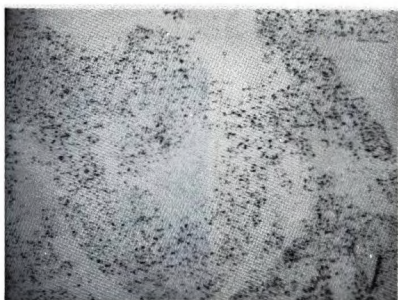
a)



b)



c)



d)

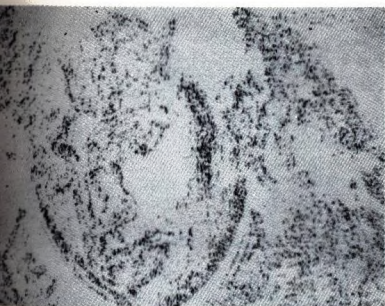


e)



f)

25. kép. Élkitűző gradiensszűrők hatása a 2/a képre:
a) Roberts-, b) Söbel-, c) Prewitt-, d) Laplace-,
e) módosított Laplace-, f) „búr kalap” szűrő



a)



b)



c)



d)

26. kép. Iránykereső szűrők hatása a 2/a képre:
a) Robinson-, b) Kirsch-, c) Frei-, d) főirányszűrő

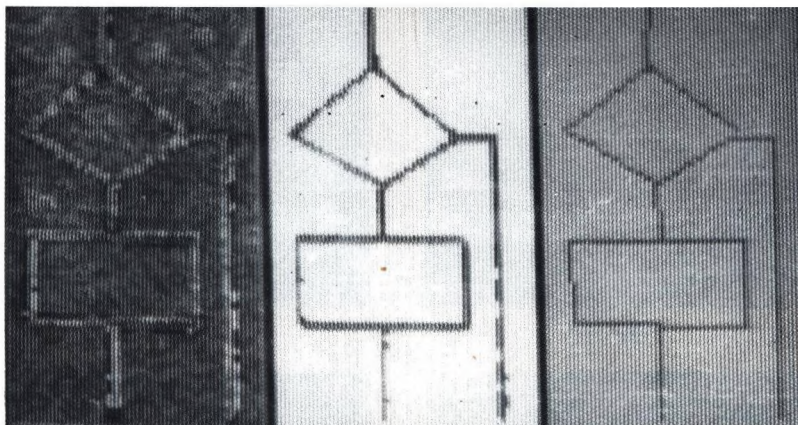


a)



b)

27. kép. Élkorrekció vonalvékonyítással a 26/c képre:
a) 1. iteráció eredménye, b) végeredmény



a)

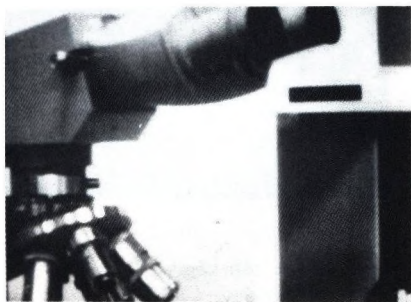
b)

c)

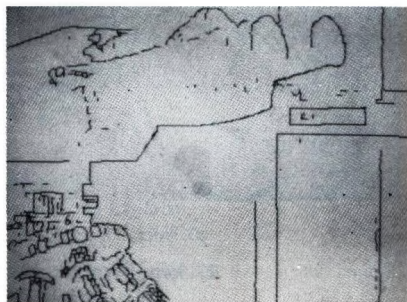
28. kép. Példa egyenesvonalú vonalvékonyításra:
a) eredeti kép, b) kétszintű kép, c) végeredmény



a)



b)

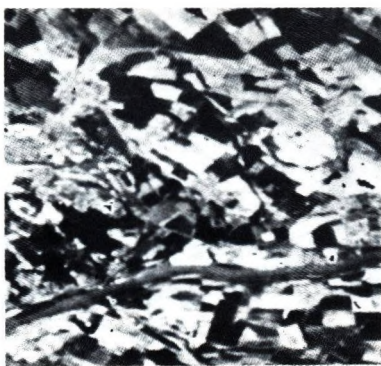


c)

29. kép. Példák heurisztikus élkorrekcióra:
a) a 26/d kép korrigált képe, b) mintakép, c) korrigált élkép

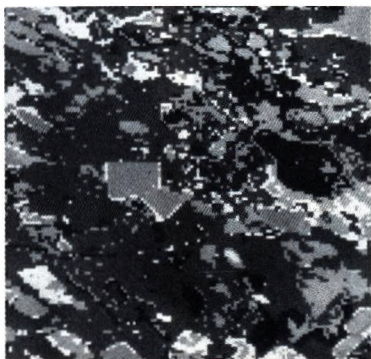


a)

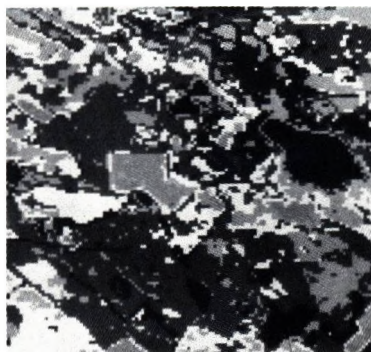


b)

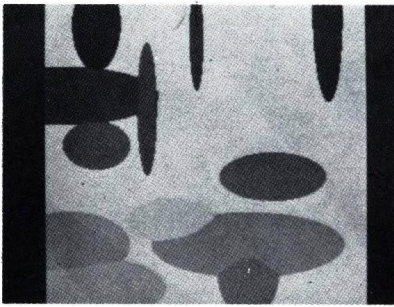
30. kép. Karhunen–Loève-transzformáció a mintaképre (13. kép):
a) KL1 sáv, b) KL2 sáv



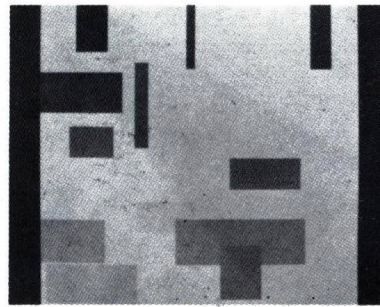
31. kép. A mintakép (13. kép) osztályozása Bayes-módszerrel



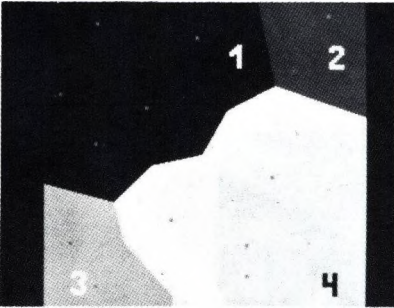
33. kép. A mintakép (13. kép) osztályozása dobozmódszerrel



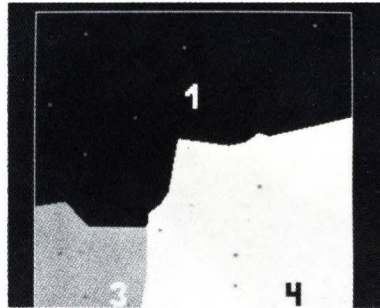
a)



b)

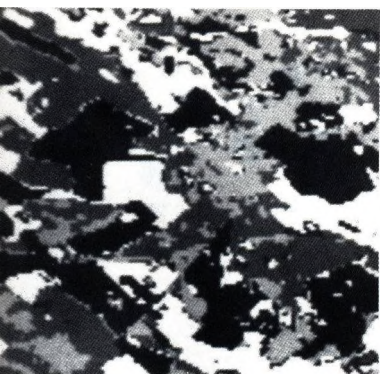


c)

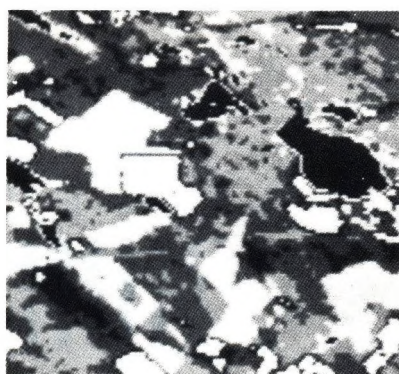


d)

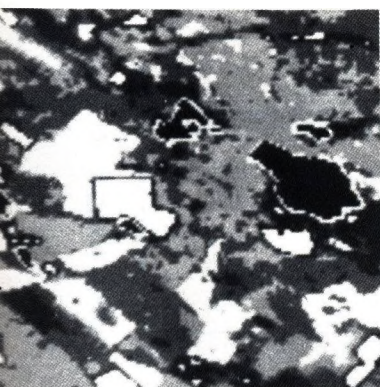
32. kép. Kétdimenziós döntési régiók:
a) Bayes-, b) doboz-, c) LKS-, d) (3,2)-LKS-módszer esetén



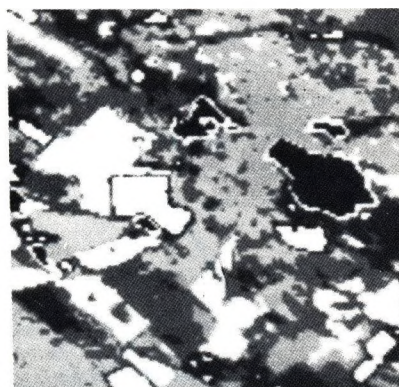
a)



b)



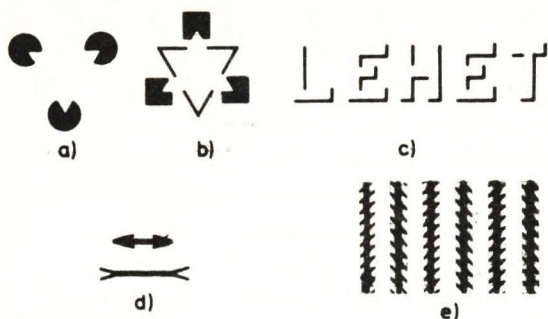
c)



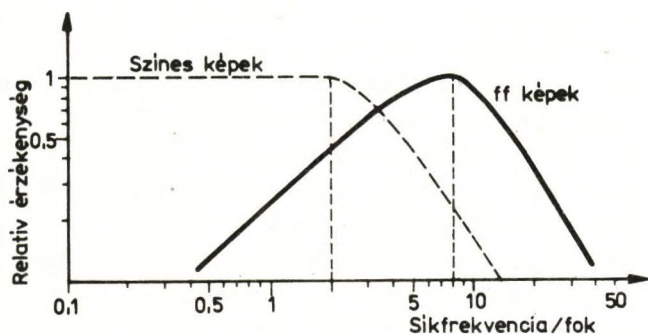
d)

34. kép. ISODATA klaszterezés a mintakép (13. kép) egy részletére:
a) az első, b) a negyedik, c) a tizenegyedik, d) a huszadik ciklus eredménye





1-6. ábra. Optikai csalódások a szűrkelátásban: a)-c) kontrasztugrá-
sok hatása, d)-e) adaptációs csalódások



1-7. ábra. Síkfrekvencia-érzékenységi görbék

érdekes módon függ még az észlelés a kontrasztosságtól is: növekvő kontraszttal javul, vagyis a görbe kiszélesedik [1-44].

Mindezeket erősen befolyásolja a szem *alkalmazkodása* (adaptáció) a környezethez. Az 1-6/d ábrán pl. különböző hosszúnak látjuk a két vonalszakaszt, az 1-6/e ábrán pedig összefutónak a párhuzamos vastag vonalakat, mivel szemünk a vékony vonalsereg által kijelölt irányokhoz alkalmazko-

1.1.3.3 Szinlátás

A szinlátás alapja - mint láttuk - a csapok eltérő spektrális érzékenysége. Bár az ingerületek kialakulásának pontos mechanizmusát nem ismerjük, azt tudjuk, hogy a szem lényegében *integrátorként* működik. Ez azt jelenti, hogy ha egyidejűleg *különböző* hullámhosszu és energiájú (= *spektrális energiaeloszlásu*) fénysugarak érik, akkor a csapok ingerválaszát (a hullámhosszt nm-ben mérve) az

$$i_j = \int_{380}^{780} I(\lambda) e_j(\lambda) d\lambda \quad (j \in P, D, T) \quad (1-1)$$

összefüggés adja, ahol $I(\lambda)$ a beeső fény spektrális energiaeloszlását, e_P, e_D , illetve e_T pedig rendre a P-, D-, illetve T-típusú csapok érzékenységi függvényét jelenti; ez utóbbiakat az 1-4. ábrán adott görbékből lehet leszámoltatni. (Ha a fényforrás csak véges számú hullámhosszon sugároz, az (1-1) integrálok súlyozott összegre egyszerűsödnek.) Ebből következik, hogy ha két fényforrás spektrális energiaeloszlása ugyan különböző, de az (1-1) integrálok azonos értéket adnak, akkor a fényüket *nem tudjuk megkülönböztetni*. Más szóval szemünk nem tudja *külön érzékelni* az egyidejűleg beérkező, különböző hullámhosszuságú és energiájú fényimpulzusokat, csak ezek *eredőjét*. (Ellentétben a fülünkkel, amely differenciális működésű; kis gyakorlattal egy zenekari akkordban minden hangszer hangját külön is észlelni tudjuk.)

Az agyunkban kialakuló *szinérzéketnek* 3 jellemző sajátossága van:

1. A *színezet* (hue), amit a hétköznapi életben pontatlanul csak színek mondunk, a fény hullámhosszától függ.

A szemünk mintegy 200 féle színezetet tud megkülönböztetni, azonban különböző monokromatikus

fények *keverékét* (különböző spektrális energiaelozslású fényforrások fényét) is azonos színezetűnek érzékelheti, ez az 1-4. ábra és az (1-1) képlet alapján nyilvánvaló. (*Monokromatikusnak* nevezünk egy fényforrást, ha egyetlen hullámhosszon (a gyakorlatban 1-5 nm sávszélességben) sugároz.)

Additív színkeverésről beszélünk, ha a keverékszint monokromatikus fényforrások fényének egymásra vetítésével állítják elő, így működik pl. a színes képcső.

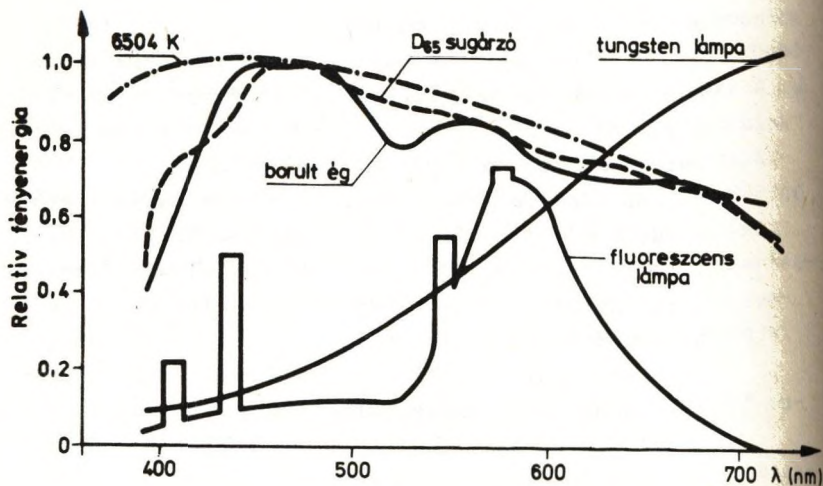
Szubtraktív színkeverés úgy jön létre, hogy egy széles sávban sugárzó fényforrás fényéből színszűrőkkel "*kivonunk*" egyes hullámhosszakot (vagy tartományokat), ekkor a keverékszín a megmaradók eredője. Ez az alapelve pl. a színes nyomtatásnak is: a színszűrők a különböző festékanyagok (lásd I. kép; a színes képek a mellékletben találhatóak, a 48. oldaltól kezdődően).

Színkeveréssel olyan színeket is létre lehet hozni, amelyek nem fordulnak elő a spektrumban. Ilyen pl. a *bibor*, amelyik a vörös és a kék színezet keveréke; de ilyen a *fehér* szín is.

Az *ideális fehér* fényforrás spektrális energiaeloszlása *egyenletes*, vagyis minden hullámhosszon *egyenlő energiával* sugároz. Ezt a jelenlegi eszközeinkkel nem tudjuk megvalósítani; legjobban bizonyos hőmérsékletű *abszolút fekete testtel* lehet megközelíteni. (Ismeretes, hogy az abszolút fekete test spektrális energiaeloszlását - s ezzel együtt a színezetét is - egyértelműen meghatározza az abszolút hőmérséklete, amit éppen ezért *színhőmérsékletnek* is nevezünk.) Az (1-1) képletből következik, hogy a fehér színezetet 3 (pl. vörös, zöld

és kék színezetű) monokromatikus fényforrás fényéből is elő lehet állítani.

Az 1-8. ábrán megadjuk néhány fehér színezetű fényforrás spektrális energiaeloszlási görbéjét. (A különböző spektrális energiaeloszlású, fehér színezetű fényforrások "színezete" egymástól természetesen eltérő.) A CIE D65 jelű görbe a CIE (Commission Internationale de l'Eclairage = Nemzetközi Világítástechnikai Bizottság) által elfogadott "standard fehér" színezetű fényforrást jellemzi, amely az ultraibolya sugárzást is tartalmazó közepes nappali sugárzásnak felel meg [1-23]. Ennek korrelált színhőmérséklete 6504 K, ami azt jelenti, hogy a sugárzás *színezete* megegyezik a 6504 K hőmérsékletű abszolút fekete testéével. (De a spektrális energiaeloszlása nem!)

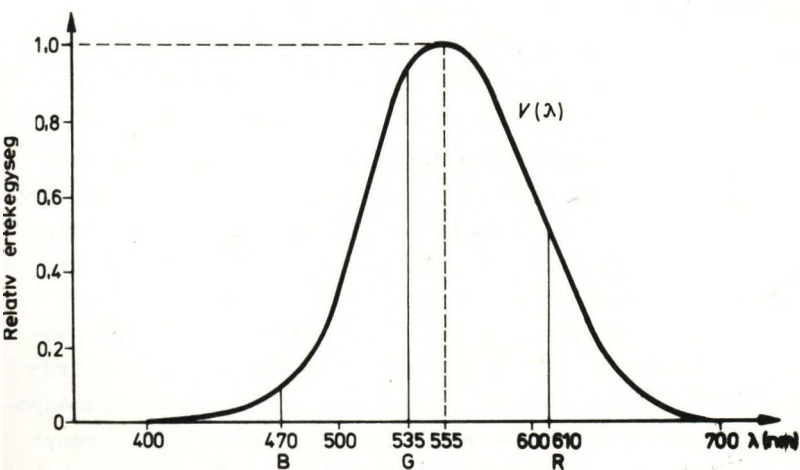


1-8. ábra. Fehér fényforrások spektrális energiaeloszlása

2. A *telítettség* (saturation) attól függ, hogy mekkora a fény fehér összetevője. (Pontosabban: az észlelt fényenergiának hányad része oszlik el úgy a spektrumban, hogy fehér színezetűnek látjuk.) A spektrumszínek *monokromatikusak*, azaz fehér összetevőt nem tartalmaznak; ezért 100% telítettségűek. A *rózsaszín* pl. néhány % telítettségű vörös színezet.

A szemünk egy adott színezetben kb. 20 telítettségi fokozatot tud megkülönböztetni.

3. A *világosság* (intensity) - mint már említettük - az egységnyi térszögben a szemünkbe érkező fényenergia mennyiségétől függ. A szem eredő *spektrális érzékenysége* változik a hullámhosszal: azonos energiájú, különböző hullámhosszu fénysugarakat különböző világosságúnak látunk, az 1-9. ábra szerint. (Az ábrán a CIE által nemzetközileg elfogadott, nagy számu egyedi mérésen alapuló, $V(\lambda)$ láthatósági függvényt adtuk meg.) Ha egy fényforrás



1-9. ábra. A fotopos látás láthatósági függvénye

spektrális energiaeloszlását $I(\lambda)$ -val jelöljük, akkor fénysűrűségét

$$F = k \int_{380}^{780} I(\lambda) V(\lambda) d\lambda \quad (1-2)$$

értékűnek látjuk, a szem integrális működése miatt. (A k konstans értéke a mértékegységek megválasztásának függvénye.) A világosságban mintegy 500 fokozatot tudunk megkülönböztetni.

Példaként megemlítjük, hogy a barna szinezet spektrális energiaeloszlása azonos a sárgáéval, csak kisebb a világossága.

Két fényforrás fényét akkor látjuk azonosnak, ha mindhárom jellemzőjük megégyezik. A színérzéklet keletkezését egyelőre még nem sikerült pontosan megmagyarázni.

A *Young-Helmholtz-elmélet* szerint a színérzéklet közvetlen összefüggésben van a csapokban kiváltott ingerületekkel. A *Hering-elmélet* szerint viszont a retinában 3 féle *bipoláris* neuron van, amelyek a "fehérséget" (vagy világosságot) és a vörös-zöld, illetve a kék-sárga szinezetet észlelik, ellentétes (inverz) "üzemmódban" [1-24].

Hubel és Wiesel felfogásában a két elmélet dialektikus ötvözetét ismerhetjük fel [1-22]. Eszerint a neuronok "körlap-körgyűrű" típusu érzékelési területében a körlaphoz az egyik színre ingeradó, a körgyűrűhöz egy másikra gátláskeltő csapok kapcsolódnak. Így alakulnak ki a "vörös-zöld" (vörösre aktiv, zöldre gátolt), a "zöld-vörös", illetve a "kék-sárga" (a sárga gátlást a körgyűrűhöz kapcsolódó vörös és zöld gátláskeltő csapok eredményezik) színérzékítő neuronok. Ez megmagyarázza azt a tényt is, hogy egy fényforrást sohasem látunk *egyidejűleg* vörösnak és zöldnek vagy kéknek és sárgának.

Bár igen kis színezetkülönbségeket is tudunk észlelni, döntésünk csak kvalitatív: nem, alig vagy jól észrevehető, nagy stb. Vagyis színérzékelésünk nem abszolút, a fenti 3 jellemző egyikére nézve sincs "kalibrálva". Objektív színmérési célokra ezért a *színegyeztető* (colour matching) technikát dolgozták ki. Ennek az a lényege, hogy a megfigyelő szemébe egyidejűleg két, egymással jól érintkező (rendszerint félkör alakú) fényfoltot vetítenek. Az egyiket egy meghatározott fénysűrűségű monokromatikus fényforrás fényével generálják; a másikat viszont úgy hozzák létre, hogy egymásra vetítik 3 pontosan meghatározott hullámhosszu (R = red (vörös), G = green (zöld) és B = blue (kék)), ugyancsak monokromatikus fényforrás fényét. Az utóbbiak fénysűrűségét addig változtatják, amíg a megfigyelő számára a két fényfolt egybe nem olvad. A kiegyenlítő fényforrások fénysűrűségét *színinger-összetevőknek* (tristimulus value) nevezik, ezek szolgálnak a vizsgált monokromatikus szín jellemzésére. (Részletesebben lásd pl. [1-23].) Sok megfigyelő adatainak átlagolásával a teljes spektrumra (5 nm-enként) meghatározták a *színegyeztető függvényt*, amelyet a CIE szabványosított [1-51].

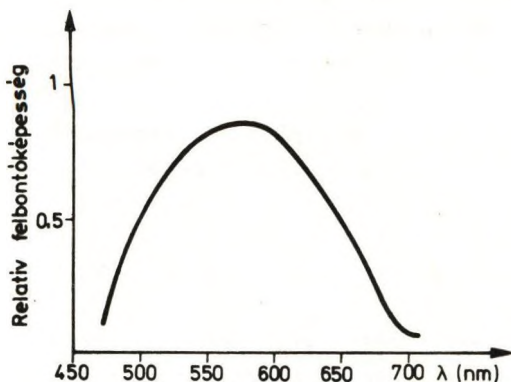
A mérések az alábbi meglepő, de a látáselmélet szempontjából *alapvető axiómák* felállításához vezettek:

- I. Bármely színingert elő lehet állítani *legfeljebb 3* monokromatikus fény keverékeként úgy, hogy a keveréket semmilyen módon nem tudjuk megkülönböztetni a kiegyenlített monokromatikus fénytől.
- II. A színegyezés széles világosságtartományban *megmarad*, bárha a színezet és/vagy a telítettség meg is változik; feltéve, hogy a fény a 2^0 -os látómezőben éri a szemet. A szem alkalmazkodása miatt

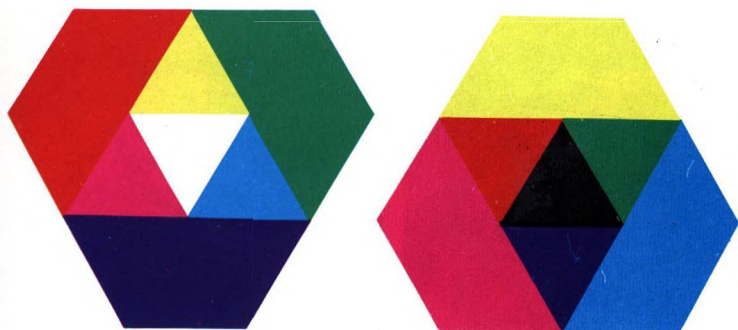
- a fényforrás és a környezet fénysűrűségétől füg-
gően - megváltozik ugyan egyik vagy másik csap-
típus spektrális érzékenysége (vagyis az (1-1)
szerinti primer ingerület nagysága, ez a *kromati-
kus adaptáció*), de a görbe *alakja nem* változik:
csak eltolódik.

III. A színezés *additív* (ha az A szín megegyezik a
B színkeveréssel és a C szín a D-vel, akkor A + C
megegyezik a B + D-vel) és *transzitiv* (ha A megegye-
zik B-vel és B a C-vel, akkor A is megegyezik C-vel).
Ezen az eredményeken alapulnak a lineáris látási model-
lek.

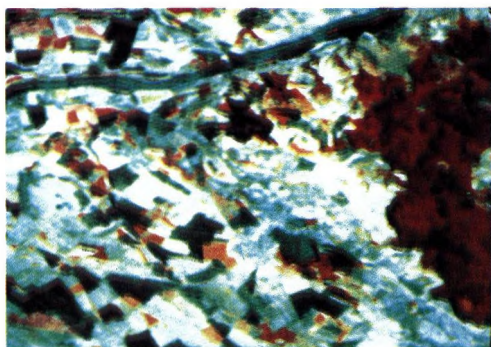
Megjegyezzük még, hogy a szem *felbontóképessége* (vagy-
is az észrevehető legkisebb fényfoltok legkisebb távolsá-
ga) is függ a hullámhossztól az 1-10. ábra szerint: a lát-
ható tartomány szélei felé erősen romlik, ami azt elenti,
hogy a finomabb részleteket kevésbé érzékeljük. A *inek*



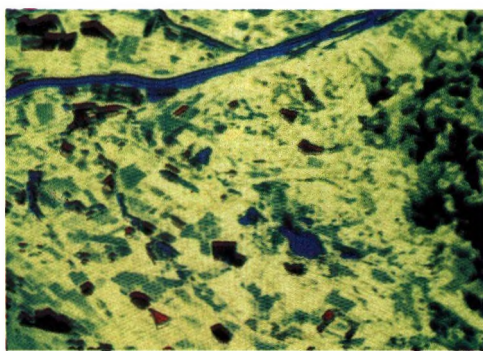
1-10. ábra. A szem spektrális felbontóképessége



I. kép. Additív és szubtraktív színkeverés

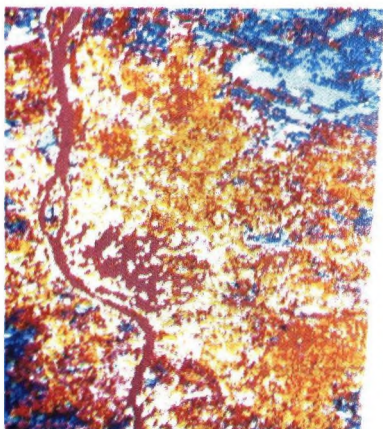


a)

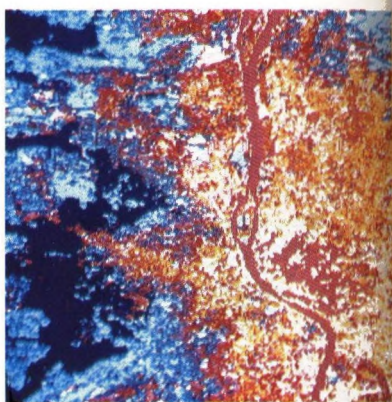


b)

II. kép. A 13. sz. mintaképek alapján készített kompozit
a), ill. álszínes b) kép



a)



b)

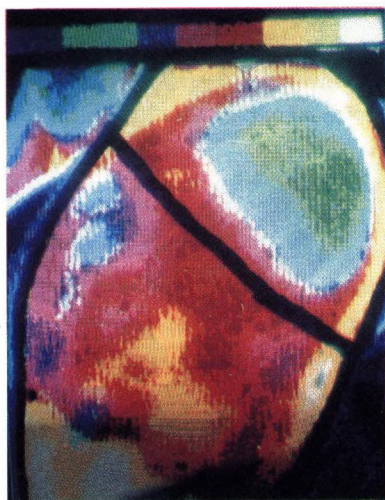
III. kép. Űrfelvétel lineáris geometriai korrekciója:
a) eredeti, b) korrigált kép



IV. kép. Felület ál-háromdimenziós megjelenítése



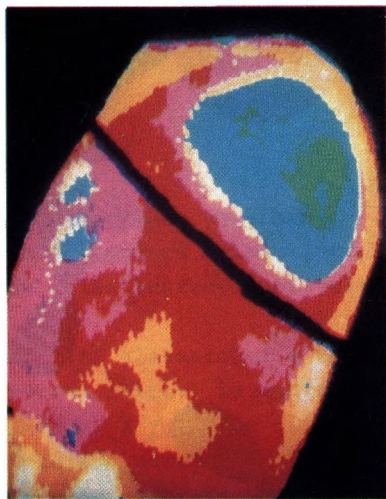
a)



b)



c)



d)

V. kép. Valódiszínés kép létrehozása:

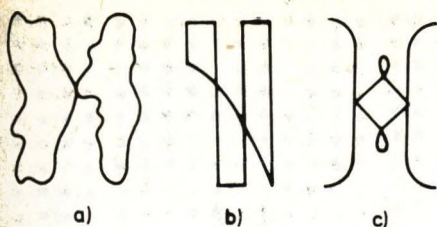
- a) eredeti hőkamera-felvétel,
- b) 4096 színre digitalizált kép, c) átszínezés 10 színre,
- d) a kiértékelendő értékes terület

alapszik, hogy adott irányu vonaldarabokat képesek felismerni. Az itt található egyszerű neuronok térbeli elrendezése lehetővé teszi, hogy ezekből összefüggő vonalak és élek képe alakuljon ki, amiből aztán az összetett neuronokban létrejönnek a geometriai helyzetűl már független, absztrakt képzetek. A magasabb szintű absztrakciós lépések mechanizmusát egyelőre még közelítőleg sem sikerült tisztázni.

Agyunk *alakfelismerő képessége* rendkívüli: a tárgyakat - méretüktől, állásuktól, helyzetüktől, környezetüktől, fénysűrűségüktől, színüktől stb. függetlenül - szinte hihetetlen biztonsággal ismerjük fel. A látás automatizálása területén ez a legnagyobb probléma. (Megjegyezzük, hogy bár a felismerés sebessége nem független a zavaró tényezőktől - pl. függőleges torzulások esetén kisebb - a jelenlegi elektronikus "látó" eszközeinkénél 2-3 nagyságrenddel nagyobb.) Ma az az elfogadott vélemény, hogy "képi adatbázis"-unk *prototípusokat* tárol, felépítése *relációs* jellegű. A konkrét felismerés egy sor transzformáció elvégzése után *mintaillesztéssel* történik. A folyamatban valószínűleg van *visszacsatolás* is: a fel nem ismert részletekről "erősebb odanézéssel" próbálunk meg több információt szerezni.

A felismerés folyamatában *fokozatosan* haladunk a mind bonyolultabb alakzatok felé úgy, hogy az alapvető jellemzők (szin, alak, mélység, mozgás) tekintetében hasonló alakzatokat *összekapcsoljuk*. Az eddig felismert főbb csoportosítási szabályok [1-50]:

- *közelség és hasonlóság*: minél közelebb vannak egymáshoz és minél hasonlóbbak az alakzatok, annál inkább vagyunk hajlamosak egy csoportba vonni őket;
- *zárttság*: az egy konturral körülzárt alakzatokat könnyebben egyesítjük (pl. az 1-11/a ábrán nem ön-



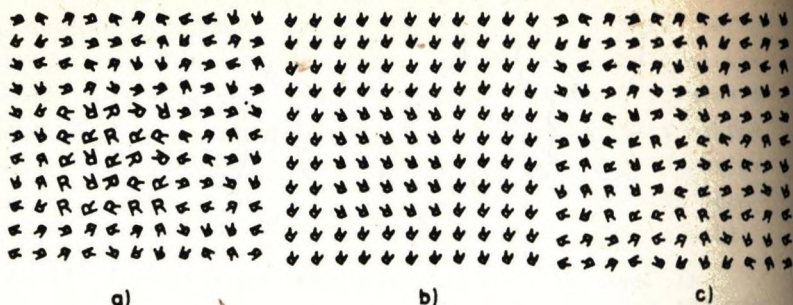
1-11. ábra. Példák az alakfelismerési szabályokra

magát metsző görbét, hanem két zárt területet látunk);

- *sima folytonosság*: olyan csoportokat képzünk, amelyeknek a lehető legsimább a konturvonaluk (pl. az 1-11/b ábrán nem 3 zárt területet, hanem egy négyszöghullámot és egy "sima" görbét látunk);
- *szimmetria és ismerőség*: **lehetőleg szimmetrikus** csoportokat képzünk és lehetőleg *már ismert* elemekből.

A "szabálysértő" alakzatokat nem vagy csak nehezen ismerjük fel. (Pl. az 1-11/c ábráról el tudjuk képzelni, hogy egy W és egy M betüből áll, de *nem látjuk* annak.)

A *mikroszerkezet*, vagyis a *textura*, illetve a *texturaelemek tudatos felismerésének* mechanizmusa hasonló; merőben más azonban a *nemtudatos felismerési folyamat*, amelyben a *textura* különböző rendű *statisztikái* (együttes valószínűségi eloszlásai) játszanak szerepet. (Egy *textura* *n*-edrendű statisztikája arról tájékoztat, hogy ha véletlenszerűen rádobunk minden lehetséges alaku és méretű *n*-oldalú sokszöget, akkor milyen valószínűséggel esik minden csucs egy meghatározott világosságú vagy színű képpontra.) Az 1-12/a, b), illetve c) ábrán [1-25] alapján mutatott *texturák* rendre különböző első-, másod-, illetve harmadrendű statisztikával bíró részekből állnak. (Az eltérő statisztikájú részlet az 5-9. sor 3-7. oszlopában van). Kísérletileg bebizonyították [1-25], hogy az olyan *texturákat*, amelyeknek első- és másodrendű statisztikájuk me-

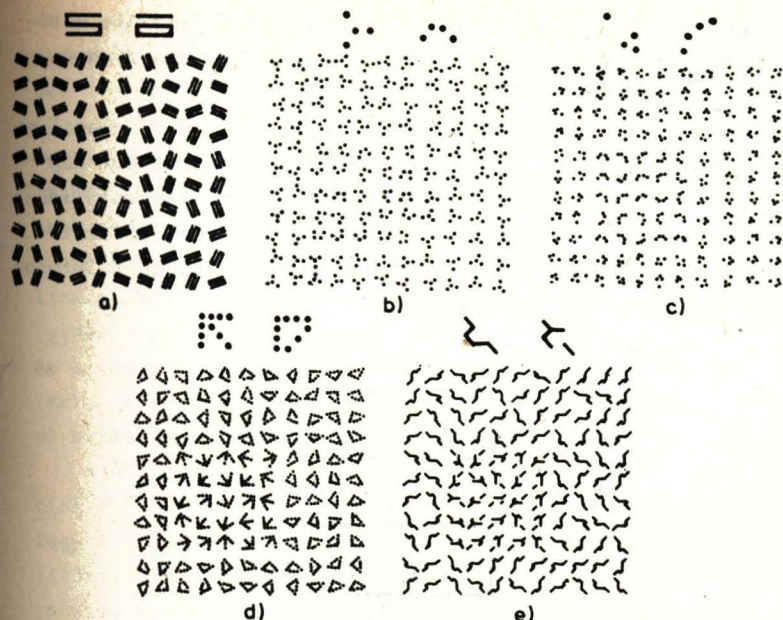


1-12. ábra. Első- (a), második- (b), illetve harmadrendű (c) statisztikában eltérő texturarárszletek

egyeznek és csak a harmad- vagy magasabbrendű statisztikájuk különböző, nemtudatos megfigyelés esetén *általában* azonosnak látjuk. (Lásd pl. az 1-12/c ábrát: a túkrözött R betűket csak tudatos odafigyeléssel vesszük észre.)

A további kísérletek során azonban kiderült, hogy azonos első és másodrendű, de különböző harmad- és/vagy magasabbrendű statisztikával rendelkező texturák *is megkülönböztethetők* nemtudatosan, ha bizonyos jellemző tulajdonságaik, amelyeket *textonnak* neveznek, különbözők. Ez idő szerint 3 textontípust sikerült felfedezni, ezek: a *szín*; adott irányu és szélességű, valamint hossz/szélesség arányu *vonaldarabok*; illetve ezek *végpontjainak száma*. Az 1-13. ábrán ugyancsak [1-25] alapján mutatott texturák másodrendű statisztikája *azonos* rendű. Az a) és b) ábrán levő eltérő mintákat nem tudjuk megkülönböztetni, mert minden textonjuk megegyezik. A c) ábrán a b) ábrához nagyon hasonló texturaelemek hossz/szélesség aránya már túlnőtt egy küszöbön; a d) és az e) ábrán pedig a vonalvégződések száma különbözik a kétféle texturárszletben. A különböző textonok hatására a texturák jól megkülönböztethetővé váltak nemtudatos megfigyeléssel is.

Természetesen más és sokkal bonyolultabb, sőt hie-



1-13. ábra. Azonos rendű statisztikájú texturaminták eltérő textonszerkezettel

rarchikusar többszintű texturák is léteznek; az alapvető felismerési szabályok azonban ezekre is érvényesek.

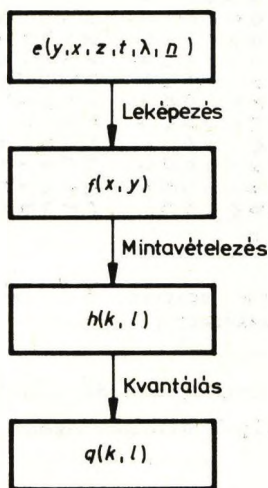
1.2 DIGITÁLIS KÉP LÉTREHOZÁSA ÉS MEGJELENITÉSE

A továbbiakban képen a vizuális információt hordozó jelet (elektromágneses hullámot) értjük. (A falon függő műalkotás, amelyet a hétköznapi szóhasználatban képnek neveznek, ebben az értelemben nem kép, hanem képforrás.)

A természetes kép analóg, térben (x, y, z) és időben (t) folytonos, különböző színezeteket (λ) tartalmazó jel,

amely még függ a megfigyelés irányától (n) is; ezt matematikailag valamilyen $e(x, y, z, t, \lambda, n)$ folytonos függvény-nyel lehet leírni.

Feldolgozás céljára a képet *fel kell venni* valamilyen leképező eszközzel (pl. kamerával stb.). A felvett kép kétdimenziós, és mivel a felvétel egy meghatározott időpillanatban, frekvenciasáv(ok)ban és irányból történt, most már leírható valamilyen $f(x, y)$ kétváltozós folytonos függvény-nyel (lásd 1-14. ábra).



1-14. ábra. Digitális kép létrehozása

A képsík a továbbiakban mindig az (x, y) sík. A tv-technikában szokásos megjelenítési mód miatt az x tengelyt balról jobbra, az y tengelyt viszont felülről lefelé irányítjuk, s az origót úgy választjuk meg, hogy a feldolgozandó kép az első síknegyedbe essen. (A gyakorlatban a kép téglalap alakú; ennek bal felső sarokpontja

az origóra, két megfelelő oldala pedig a koordinátatengelyekre illeszkedik.) Fontos megjegyezni, hogy egy jeletről egyidejűleg gyakran több frekvenciasávban is készül felvétel (pl. ürfelvételek esetében). Ezeket együttesen kell feldolgozni, ami ugyan információnyereséget jelent, de számos problémát is felvet; ezekre a későbbiekben visszatérünk.

Számítógépes feldolgozás céljából a képjelet digitalizálni kell. Ez azt jelenti, hogy a képet *képpontokra* (pixel = *picture's element*) *bontjuk* (=mintavételezés), és az egyes képpontokhoz tartozó pillanatnyi amplitúdó(k)at, amely(ek) a megfelelő képpont világosságát, illetve színét határozzá(k) meg, *kvantáljuk*.

A *digitális kép* - bár megőrzi a vizuális információt - "képiségét" elveszti. Ahhoz, hogy ismét látható legyen, vissza kell alakítani analóg jellé (= *visszadlitás*). A két folyamatnak felváltva ismételtetőknek kell lennie, ezért csak olyan eljárásokat szabad használni, amelyek ezt lehetővé teszik.

A képfüggvényt általában determinisztikusnak tekintjük, de egyes eljárásokban valószínűségi változóként (lásd F2.2 pont) kezeljük. Az utóbbi főleg a zajhatások és a mérési pontatlanságok figyelembevételkor hasznos.

1.2.1 MINTAVÉTELEZÉS

A digitalizálás első lépése az analóg kép felbontása képpontokra. Ez úgy történik, hogy a képsík meghatározott pontjaiban mintákat veszünk a képjelből, amelyek a folytonos képnek valamilyen mérhető tulajdonságát - ff képek esetében a *fénysűrűségét*, színes képek esetében a *színét* - reprezentálják. Más szóval a mintavételezés tulajdonképpen *mérés*, s mint minden mérés, valamilyen bizonytalansággal,

mérési hibával jár. Fontos, hogy a mérési hibák nagyságát meg lehessen határozni, mert ezáltal lehet értékelni a mintavételezés "jóságát".

Legyen a leképező rendszer által produkált folytonos, véges kiterjedésű, ideális (hibátlan) kép fénysűrűségét leíró függvény $f(x,y)$. (Emléztetünk rá, hogy a szín 3 adattal jellemezhető, ezért a színes képek leírásához 3 egymástól független függvényre van szükség, pl. a 3 színösszetevőre; ezeket külön-külön f képként kezelhetjük. Az eljárást csak egy függvényre írjuk le, de nehézség nélkül kiterjeszthető háromra is.)

Az általánosság megszorítása nélkül feltételezhetjük, hogy a képfüggvény az (x,y) síknak egy téglalap alakú tartományában különbözik 0-tól. (Célszerű megállapodni abban, hogy az információt nem hordozó "fekete" (azaz 0 fénysűrűségű) pontokban $f(x,y) = 0$. Tetszőleges alakú képtartomány helyett tekintsük a legkisebb befoglaló téglalapot, és ennek az eredeti képtartományon kívüli részében egészítsük ki a képfüggvényt "fekete" függvényértékekkel.)

Jelöljük a mintavételi pontok x irányu távolságát Δx -szel, az y irányut Δy -nal, és legyen a képméret $(K-1)\Delta x \times (L-1)\Delta y$. Válasszuk az origót a téglalap bal felső sarokpontjában, de a könnyebb írásmód kedvéért tükrözzük a képtartományt először az y , majd az egészet az x tengelyre. Így az origó a $2(K-1)\Delta x \times 2(L-1)\Delta y$ méretű képtartomány középpontjába kerül; ezen kívül pedig egészítsük ki a képfüggvényt "fekete" értékekkel a teljes (x,y) síkban. (A valóságban a mintavételezés után természetesen csak az első síknegyedbe eső téglalap alakú képtartományhoz tartozó függvényértékekkel fogunk dolgozni, ez azonban nem érinti az alábbiakat.)

Az ideális mintavételt a

$$t(x, y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \delta(x - m\Delta x, y - n\Delta y) \quad (1-3)$$

mintavételi függvénnyel lehet leírni, amely az (F4-5)-tel definiált Dirac-delta kiterjesztése az összes mintavételi pontra. A mintavételezés folyamatát első közelítésben a képfüggvény és a mintavételi függvény konvolúciója (lásd F4.1.2 alpont) fejezi ki:

$$\begin{aligned} h(k, l) &= f(x, y) \otimes t(x, y) = \\ &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \lim_{\substack{x \rightarrow m\Delta x \\ y \rightarrow n\Delta y}} \iint \delta(x - m\Delta x, y - n\Delta y) f(x, y) \, dx dy = \\ &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} f(m\Delta x, n\Delta y). \end{aligned} \quad (1-4)$$

A képlet középső sorában felhasználtuk a Dirac-delta (F4-5)-beli definícióját, az utolsó sorban pedig az (F4-6)-ban levezetett "toló" tulajdonságát. Látható, hogy az eredményül kapott *mintavételezett kép* nem más, mint az eredeti képfüggvényértékek ("pontforrások") halmaza a ($k = m\Delta x$, $l = n\Delta y$) koordinátájú képpontokban (=mintavételi pontokban).

Az (1-4) összefüggés egzakt levezetéséhez szükséges, hogy a képfüggvény rendelkezzen bizonyos - itt nem részletezendő - tulajdonságokkal (lásd pl. [1-17]). Ezek közül számunkra a két legfontosabb az, hogy *stacionáriusnak* tekinthető (vagyis értéke nem vagy csak elhanyagolható mértékben változik a mintavétel folyamán), és hogy "gyengén" *sávkorlátos*.

Az utóbbi tulajdonság a gyakorlatilag megvalósíthatatlan végtelen összegzés problémáját teszi áthidalhatóvá. Arról van szó, hogy a képfüggvényt - mivel általában

ugysem lehet megadni analitikus formában, s így matematikai kezelése nehézkes – célszerű felbontani valamilyen ortogonális bázisfüggvényrendszer szerint (lásd F4.3.1 alpont), s a kapott függvénysorral számolni. Matematikai szempontból a legkényelmesebb ilyen felbontást a *Fourier-transzformációval* (lásd F4.3.2 alpont) nyert *frekvenciaspektrum* adja, amely azonban általában szintén végtelen kiterjedésű. Tekintsünk viszont egy sávkorlátos függvényt, amelyre

$$F(u, v) = 0, \text{ ha } u \geq u_{\tilde{h}} \text{ és/vagy } v \geq v_{\tilde{h}}, \quad (1-5)$$

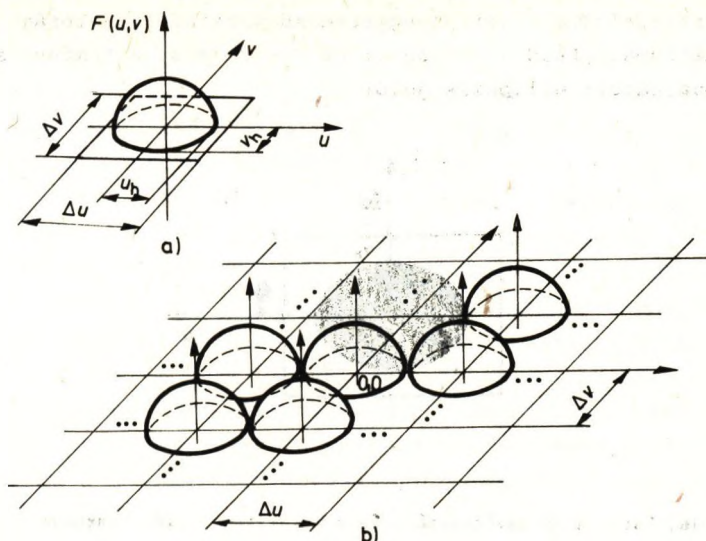
ahol az $u_{\tilde{h}}$ és a $v_{\tilde{h}}$ az (u, v) síkfrekvencia-tartományban az x , illetve y irányu fénysűrűség-változásoknak megfelelő határfrekvencia. (Itt és a továbbiakban *nagybetűvel* jelöljük egy függvény Fourier-transzformáltját.) Be lehet bizonyítani (lásd pl. [1-17]), hogy ez esetben egyrészt az eredeti képfüggvény frekvenciaspektruma

$$F(u, v) = \frac{1}{4\pi^2} \int_{-u_{\tilde{h}}}^{u_{\tilde{h}}} \int_{-v_{\tilde{h}}}^{v_{\tilde{h}}} f(x, y) \exp[-i2\pi(ux + vy)] dx dy \quad (1-6)$$

alakba írható (lásd 1-15/a ábra); másrészt a mintavételezett képé ennek *végtelen szuperpozíciója* a síkfrekvencia-tartomány $(m\Delta u, n\Delta v)$ koordinátájú rácspontjaiban ($m, n = 0, \pm 1, \pm 2, \dots$; lásd 1-15/b ábra):

$$H(u, v) = \frac{\Delta u \Delta v}{4\pi^2} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} F(u - m\Delta u, v - n\Delta v). \quad (1-7)$$

Az (1-7) képletben



1-15. ábra. Sávkorlátos analóg (a) és mintavételezett (b) képfüggvény fiktív síkfrekvencia-spektruma

$$\Delta u = \frac{2\pi}{\Delta x}, \text{ illetve } \Delta v = \frac{2\pi}{\Delta y} \quad (1-8)$$

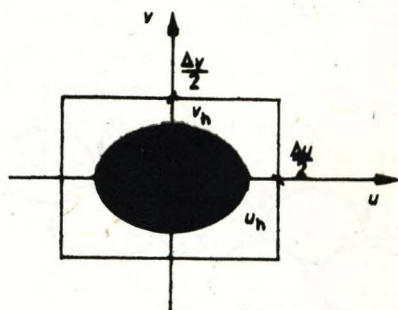
a mintavételezési frekvencia x , illetve y irányban.

Ha a képfüggvény "gyengén" sávkorlátos, ez azt jelenti, hogy az (1-5) feltétel egzaktul *nem* teljesül ugyan, de a gyakorlatban *elfogadható*, mivel $F(u, v)$ "elég gyorsan" (a frekvencia négyzetével fordított arányban) eltűnik a határfrekvenciáknál nagyobb síkfrekvenciákon.

Az 1-15/b ábráról az is látható, hogy a mintavételezett kép frekvenciaspektruma akkor és csak akkor *átfedésmentes*, ha teljesül az

$$u_h < \frac{\Delta u}{2} = \frac{\pi}{\Delta x}, \text{ illetve } v_h < \frac{\Delta v}{2} = \frac{\pi}{\Delta y} \quad (1-9)$$

Nyquist-feltétel, ami az egyértelmű visszaállithatóság kritériuma. (Lásd 1-16. ábra: az $F(u, v) \neq 0$ tartományt a bevonalkázott ellipszis jelöli.)



1-16. ábra. A mintavételezési és a határfrekvenciák viszonya

A mintavételezés gyakorlati megvalósítása szempontjából a Nyquist-feltétel azt jelenti, hogy a mintavételező (optikai- vagy elektron) sugárnyaláb átmérője, illetve sortávolsága nem lehet nagyobb a képen előforduló legfinomabb képrészlet periodicitásának a felénél. Mivel azonban az (1-5) sávkorlátossági kritérium nem teljesül pontosan, a mintavételezés még ebben az esetben is információvesztéssel jár; emiatt a képet már nem lehet hibátlanul visszaállítani.

Anélkül, hogy további részletekbe mennénk, megemlítjük, hogy a zajok síkfrekvencia-spektruma általában nem sávkorlátos, vagy legalábbis szélesebb a képénél. Mivel ily módon nem teljesül a Nyquist-feltétel, a zajos kép általában *alulszelektált* lesz (ami azt jelenti, hogy az ismétlődő spektrumminták átfedik egymást), és ezért nem lehet egyértelműen visszaállítani (lásd még az 1.2.3 alpon- tot is).

1.2.2 KVANTÁLÁS

A digitalizálási folyamat második lépésében az A/D átalakító összehasonlítja az egyes képpontokhoz kapott mintavételi függvényértékeket (jelamplitudókat) a lehetséges kimeneti szintjeivel, s eredményül a legközelebbi kimenő szintek kódját adja (lásd 1-14. ábra). Ezt a folyamatot nevezzük *kvantálásnak*, miáltal a mintavételezett képből létrejön a *digitális kép*. Ebben minden képponthez egy a fénysűrűségétől függő kódszám tartozik, amit *világosságkódnak* fogunk nevezni, és q -val jelölünk. (Színes képek esetében minden képponthez egy *szinkódvektor* tartozik, amelynek 3 összetevője a 3 képfüggvény egyidejű digitalizálásával áll elő.) A világosságkód nemnegatív egész szám; értékkészlete

$$0 \leq q \leq J = 2^b - 1,$$

ahol b az A/D átalakító bitszáma. (Jelenleg nálunk 6 és 8 bites átalakítók vannak forgalomban, de 10 és 12 bites is léteznek. Így az értékkészlet: $0 \leq q \leq 63$, illetve 255; ezt a $q \in \mathbb{E} 63$, illetve $q \in \mathbb{E} 255$ szimbólummal jelöljük.)

Tegyük fel, hogy a mintavételezéskor kapott függvényértékek a

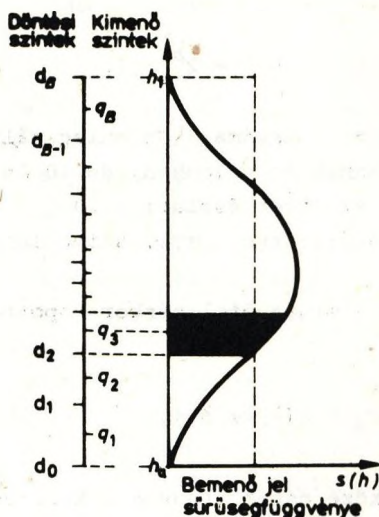
$$h_a \leq h(k, l) \leq h_f$$

alsó és felső határ közé esnek. Legyen a kvantumszintek száma $B = 2^b$, és jelöljük ezeket q_1, \dots, q_B -vel, a döntési szinteket pedig d_j -vel ($d_0 = h_a$ és $d_B = h_f$). Ekkor a kvantálást a

$$q(k, l) \leftarrow \{q_j | d_{j-1} \leq h(k, l) < d_j; j = 1, 2, \dots, B\} \quad (1-10)$$

kifejezéssel írhatjuk le (lásd 1-17. ábra). Láthatjuk, hogy a kvantálás is *információvesztéssel* jár. Az alábbiakban az egyszerűség kedvéért elhagyjuk a (k, l) argumentumot. Tekintsük most a képfüggvényt valószínűségi változónak, amelynek sűrűségfüggvényét jelölje $s(h)$. A kvantálás "jóságát" a négyzetes hibával fejezhetjük ki (lásd pl. [A-14], 101. old.)

$$\epsilon = \int_{h_a}^{h_f} (h - q)^2 s(h) dh = \sum_{j=1}^B \int_{d_{j-1}}^{d_j} (h - q_j)^2 s(h) dh. \quad (1-11)$$



1-17. ábra. Kvantálási szintek kialakítása

A kvantálást akkor tekintjük optimálisnak, ha ez a négyzetes hiba minimális, adott B kvantumszint esetén. A minimum megkeresése céljából számítsuk ki (1-11) differenciálhányadosát a d_j -k, illetve a q_j -k szerint:

$$\frac{\partial \epsilon}{\partial d_j} = (d_j - q_j)^2 s(d_j) - (d_j - q_{j+1})^2 s(d_j) = 0$$

$$(j = 1, 2, \dots, J); \quad (1-12)$$

illetve

$$\frac{\partial \epsilon}{\partial q_j} = -2 \int_{d_{j-1}}^{d_j} (h - q_j) s(h) dh = 0$$

$$(j = 1, 2, \dots, B). \quad (1-13)$$

Ha $s(d_j) \neq 0$, (1-12)-ből és (1-13)-ból következik, hogy

$$d_j = \frac{q_j + q_{j+1}}{2}, \quad (j = 1, 2, \dots, J); \quad (1-14)$$

illetve

$$q_j = \frac{\int_{d_{j-1}}^{d_j} hs(h) dh}{\int_{d_{j-1}}^{d_j} s(h) dh} \quad (j = 1, 2, \dots, B). \quad (1-15)$$

Vagyis az optimális kvantáló döntési szintjei a lehetséges kvantumszintek felezőpontjaiban vannak, míg a kvantumszintek a megfelelő két döntési szinttel és a sűrűségfüggvénnyel határolt görbe vonalú trapéz súlyvonalán helyezkednek el (lásd 1-17. ábra). Könnyű belátni, hogy ha a képfüggvényértékek eloszlása egyenletes (vagyis a sűrűségfüggvény állandó, lásd az 1-17. ábrán a szaggatott vonalat), akkor (1-15) a

$$q_j = \frac{d_{j-1} + d_j}{2}$$

kifejezésre egyszerűsödik, vagyis ebben az esetben az egyenlő közü kvantálás az optimális. Különbö a q_j és a d_j értékeket az (1-14) és az (1-15) egyenletrendszerből az alábbi iterációval lehet meghatározni $s(h)$ ismeretében:

- válasszuk meg q_1 -et tetszés szerint és határozzuk meg (1-15)-ből d_1 értékét;
- q_1 és d_1 felhasználásával (1-14)-ből kiszámíthatjuk q_2 -t;
- d_1 és q_2 ismeretében (1-15)-ből kiszámítható d_2 ; és így tovább.

Ha az utolsó lépésben kapott q_B -re és $d_B = h_f$ -re teljesül az (1-15) egyenlőség, akkor az eljárás véget ért, különben előlről meg kell ismételni q_1 megfelelően módosított értékével.

Az ilyen típusu eloszlásfüggő (tapered) kvantálót úgy valósítják meg, hogy az A/D átalakító előtt a képfüggvényt nemlineáris transzformációnak vetik alá, aminek hatására a transzformált függvény sűrűségfüggvénye állandó lesz. Ezután egyenlő közü kvantálást végeznek, majd a kapott kódokból inverz transzformációval állítják elő a tényleges világosságkódokat [1-27]. Ha pl. a képfüggvényhez exponenciális eloszlású valószínűségi változó rendelhető (lásd F2.2.1 alpont), akkor a

$$h^* = 1 - \exp(-\lambda h)$$

transzformált képfüggvényt kell egyenlő közüen kvantálni. Ezután a világosságkódokat a

$$q = \frac{1}{\lambda} \ln(1 - q^*)$$

inverz transzformációval határozhatjuk meg (A-13, 146. old.)

Mivel $B = 2^b$, a $K \times L$ képpontból álló digitális kép tárolásához $K \times L \times b$ bitre van szükség. Kísérletileg bebizonyították [1-21], hogy ezt a kapacitást akkor lehetne a legjobban kihasználni, ha a mintavételezési intervallum *nem volna állandó*. Olyan képeknél, ahol lassan változik a fénysűrűség (kicsi a síkfrekvencia-tartomány), növelni kellene a kvantálás finomságát és csökkenteni a mintavételi pontok számát, míg a részletgazdag (nagy síkfrekvencia-tartományu) képek esetében a felbontást kellene finomítani, s emellett megengedhető a kvantumszintek számának csökkentése. Az elvet a 2. és a 3. képen szemléltetjük, egy kis (arckép), illetve egy nagy (mikroszkópi sejt felvétel) síkfrekvencia-spektrumu képen. Az a) képeken a kiindulási digitális képek $384 \times 288 \times 6$, a b) képeken $192 \times 144 \times 6$, a c) képeken a $384 \times 288 \times 2$ bitre vannak digitalizálva. Az arcképen a kisebb síkbeli felbontás, a sejt felvételen a durvább kvantálás hatása a kevésbé zavaró. Megjegyezzük, hogy a gyakorlatban ilyen digitalizálók nem léteznek.

1.2.3 VISSZAÁLLÍTÁS (REKONSTRUKCIÓ)

A *viSSzaállítás* vagy *rekonstrukció* (reconstruction) során a digitális képből *interpolációval* állítjuk elő az analóg képjelet, digitális-analóg (D/A) átalakítókkal.

A (1-5) képlet értelmében sávkorlátos függvényekre érvényes az

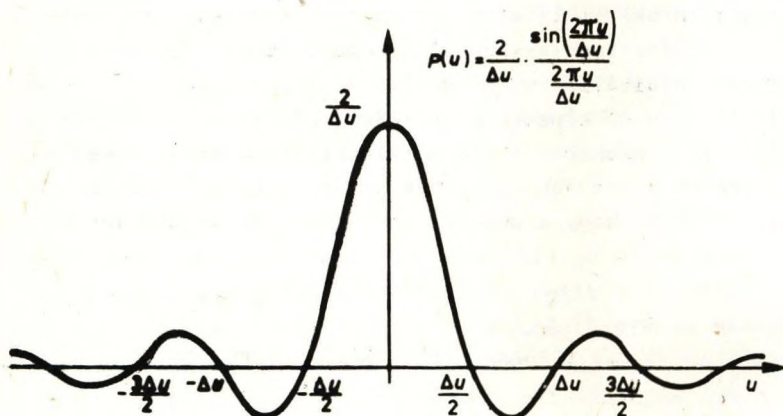
$$f(x, y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} f(m\Delta x, n\Delta y) \frac{\sin[u_h(x - m\Delta x)]}{u_h(x - m\Delta x)} \frac{\sin[v_h(y - n\Delta y)]}{v_h(y - n\Delta y)}$$

(1-16)

Shannon-Kotyelnyikov-tétel [1-40]. Ez lehetővé teszi, hogy ha ismerjük a (pontos) mintavételezett képet a képsík $(m\Delta x, n\Delta y)$ pontjaiban, akkor ebből *tökéletes pontossággal* visszaállítsuk az eredeti képfüggvényt. Az (1-16)-ban szereplő

$$\text{sinc}(x, y) = \frac{\sin x}{x} \frac{\sin y}{y} \quad (1-17)$$

ideális interpolátorfüggvényt egydimenziósan az 1-18. ábrán adtuk meg.



1-18. ábra. Sinc típusu interpolátorfüggvény

Sajnos - mint láttuk - a gyakorlatban egyrészt a képfüggvény csak "gyengén" sávkorlátos, másrészt a mintavételezett kép is csak véges számú függvényértékből áll össze, amelyeket ezen kívül még a kvantálási hiba is terhel. Induljunk ezért ki abból az $\hat{f}(x, y)$ folytonos képfügg-

vényből, amelyik az eredeti $f(x,y)$ -től csak annyiban tér el, hogy a mintavételezési pontokban a kvantált függvény-értékek szerepelnek. Mivel ezek általában nem egyeznek meg az eredetivel (hanem alatta vagy felette vannak), a "képfelület" - folytonos lévén - általában a közbülső pontokban is el fog térni az eredetitől. Erről az (1-11) alapján annyit állithatunk, hogy

$$\iint_{-\infty}^{+\infty} [f(x,y) - \hat{f}(x,y)]^2 dx dy = \text{minimum}$$

az adott digitalizálási technika esetében. (Ami persze nem jelenti azt, hogy más technikák nem eredményezhetnek jobb közelítést; lásd a megjegyzést az előző alpont végén.)

Az (1-4) alapján a $q(k, l)$ digitális képet úgy foghatjuk fel, mintha az $\hat{f}(x,y)$ -ből állt volna elő mintavételezéssel:

$$q(k, l) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \hat{f}(m\Delta x, n\Delta y); \quad (1-18)$$

sikfrekvencia-spektrumát pedig az (1-7) alapján a

$$Q(u, v) = \frac{\Delta u \Delta v}{4\pi^2} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \hat{F}(u - m\Delta u, v - n\Delta v) \quad (1-19)$$

összefüggésből kapjuk, ahol \hat{F} az \hat{f} Fourier-transzformáltja. Az (1-19) jobb oldaláról látszik, hogy az $\hat{f}(x,y)$ sikfrekvencia-spektrumának végtelen szuperpozíciója. Tegyük fel, hogy az (1-5) és az (1-9) feltétel $\hat{F}(u,v)$ -re is igaz, ekkor a visszaállítás egyértelműen elvégezhető.

A visszaállított $f^{(-1)}(x,y)$ képfüggvényt $q(k,l)$ -nek a

$p(x, y)$ interpolátorfüggvénnyel képzett konvolúciója eredményezi:

$$f^{(-1)}(x, y) = p(x, y) \otimes q(k, l). \quad (1-20)$$

Helyettesítsük be (1-18)-at, és végezzük el a konvolúciós integrálást, akkor

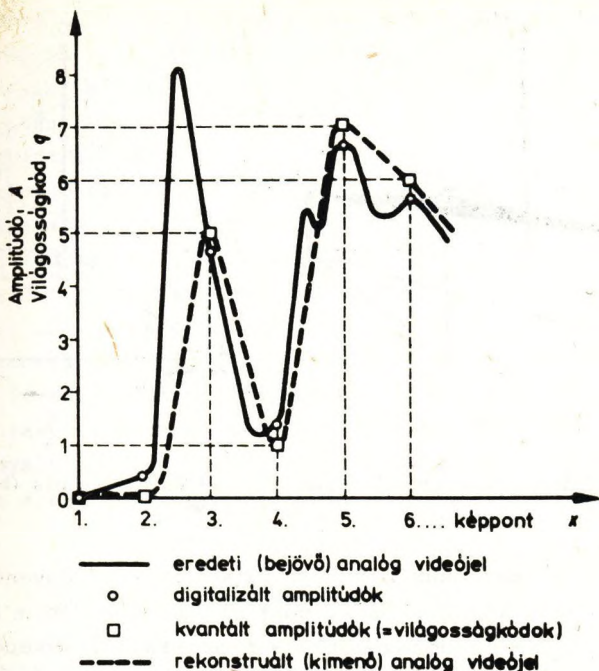
$$f^{(-1)}(x, y) = \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \hat{f}(m\Delta x, n\Delta y) p(x - m\Delta x, y - n\Delta y) \quad (1-21)$$

lesz. Ennek síkfrekvencia-spektruma (1-19) és (1-20) alapján - kihasználva, hogy a konvolúcióképzés a transzformált függvények szorzatával helyettesíthető:

$$F^{(-1)}(u, v) = \frac{\Delta u \Delta v}{4\pi^2} P(u, v) \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \hat{F}(u - m\Delta u, v - n\Delta v). \quad (1-22)$$

Az (1-21)-ből látható, hogy a képfüggvény visszaállításához egy adott pontban szükség van egy (elvből végtelen) környezetében levő mintavételi értékekre is. (Ha a visszaállítást sorfolytonosan végezzük, szemléletesen szólva ez azt jelenti, hogy nem csak az aktuális képpont "multját", hanem a "jövőjét" is ismernünk kell.) A valószínűségi interpolátor (D/A átalakító) természetesen csak véges környezetre támaszkodik; a "jövőőről" pedig *késleltetése* révén szerez információt. (Ez azt jelenti, hogy csak olyan pontban tudja visszaállítani a képfüggvényt, amelynek már "elég nagy" környezetében megtörtént a mintavételezés.)

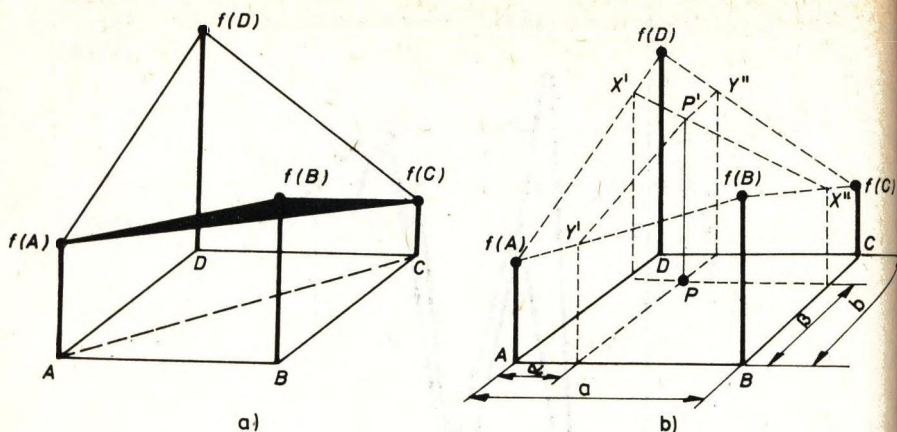
Az eredményt egydimenziós esetre az 1-19. ábrán szemléltetjük. Az ábrán jól látható a mintavételezési és a kvantálási hibák hatása.



1-19. ábra. Analóg képjel visszaállítása (rekonstrukció)

A gyakorlatban lineáris interpolációt valósítanak meg. Az 1-20/a ábrán a szakaszonként lineáris interpoláció elvét mutatjuk. A rácspontok által kijelölt téglalap felett a képfüggvényt két lépésben közelítjük: pl. az ABC háromszög felett az $f(A)$, $f(B)$, $f(C)$ pontra, az ACD háromszög felett pedig az $f(A)$, $f(C)$, $f(D)$ pontra illeszkedő síkkal.

Az 1-20/b ábrán a bilineáris interpoláció elvét szemléltetjük. Ekkor az $ABCD$ téglalap egy belső P pontjában a következő két módon határozhatjuk meg a képfüggvény értékét:



1-20. ábra. Szakaszonként lineáris (a) és bilineáris (b) interpoláció elve

- Megkeressük a P pont képét az $\overline{X'X''}$ egyenesen. (Az X' , illetve az X'' pont az $\overline{f(A)f(D)}$, illetve az $\overline{f(B)f(C)}$ egyenes dőféspontja a P pontra illeszkedő, a képsíkra merőleges és az x tengellyel párhuzamos síkon.)
- Megkeressük a P pont képét az $\overline{Y'Y''}$ egyenesen. (Az Y' , illetve az Y'' pont az $\overline{f(A)f(B)}$, illetve az $\overline{f(C)f(D)}$ egyenes dőféspontja a P pontra illeszkedő, a képsíkra merőleges és az y tengellyel párhuzamos síkon.)

Könnyű belátni, hogy mindkét esetben az ábrán jelölt P' ponthoz jutunk. Jelöljük az interpolációs téglalap \overline{AB} oldalát α -val, \overline{BC} oldalát b -vel, és legyen a P pont távolsága az A ponttól a megfelelő oldalakon mérve α , illetve β , akkor az interpolált képfüggvényértéket az

$$f(P) = \beta[\alpha f(C) + (a - \alpha)f(D)] + (b - \beta)[\alpha f(B) + (a - \alpha)f(A)] \quad (1-23)$$

képletből számíthatjuk ki.

Összefoglalásként a következőket állapíthatjuk meg:

- Sávkorlátos (lásd (1-5)) függvény ideális (végtelen síkfrekvencia-spektrumu) mintavételezés esetén egzaktul visszaállítható, ha teljesül az (1-9) Nyquist-feltétel.

- A képfüggvények általában "gyengén" sávkorlátosak, a mintavételezés nem ideális (véges), és a mintákat még kvantáljuk is; ráadásul az interpolátor sem ideális. Ezért a visszaállított képfüggvény *sohasem lesz azonos* az eredetivel; megfelelő interpolátor választásával legjobb esetben azt érhetjük el, hogy az (1-22) szerinti síkfrekvencia-spektruma - a mintavételezési és a kvantálási hibák hatását is figyelembe véve - megegyezék az eredetivel. Ekkor a két képet azonosnak tekintjük.

1.3 A KÉPFELDOLGOZÁS SAJÁTOS ESZKÖZEI

A képfeldolgozási problémák során az elméleti megfontolásokat konkrét megvalósítási tapasztalatokkal kívánjuk alátámasztani, így elkerülhetetlenül utalni fogunk az alkalmazott rendszerekre. Célszerűnek tartjuk ezért áttekinteni a képfeldolgozásban használatos speciális hardver eszközöket és megemlíteni a szoftver eszközökkel kapcsolatos egyes problémákat.

Az utóbbi évtizedben a képfeldolgozó eszközök területén bekövetkezett - és jelenleg is folytatódó - rohamos fejlődés eredményeként újabb és újabb rendszerek jelennek meg. Nem kívánunk naprakész áttekintést adni ezekről, egyrészt terjedelmi okokból, másrészt a későbbiek szempontjából nem is tartjuk szükségesnek. (Az érdeklődők ebben a témakörben számos összefoglalást találhatnak az irodalomban; pl. [1-11], [1-33], [1-36]). Inkább csak az

egyres berendezések működési elvét ismertetjük, de ezt is csak röviden, mivel ilyen jellegű áttekintés magyar nyelven is megtalálható [A-1]. Kicsit részletesebben foglalkozunk azokkal a berendezésekkel, amelyeken a könyvben bemutatásra kerülő feldolgozások történtek, mert az eszközök műszaki paraméterei a végeredményre is hatással vannak. Ebből a szempontból a képfeldolgozó eszközök terén járatos olvasóknak is érdemes átlapozniuk ezt a pontot, amelynek azonban elsősorban az a célja, hogy az ebben a vonatkozásban kellő ismerettel nem rendelkezők könnyebben megértsék a későbbiekben található eszközfüggő hivatkozásokat. Az egyes eszközöket a jobb áttekintés kedvéért funkcióik szerint csoportosítottuk, bár a gyakorlatban egy-egy képfeldolgozó rendszeren belül ezek sokszor nehezen választhatók szét.

1.3.1 SZÁMITÓGÉPEK

A képfeldolgozás területén felhasznált számítógépek skálája a személyi számítógépektől a mammutméretű szuper-számítógépeken keresztül a speciális architektúrájú sokprocesszoros rendszerekig terjed.

A kiválasztást - a megoldandó feladaton túl - döntően befolyásolja a felhasználó műszaki-technikai környezete és anyagi ereje. A lehetőségek határain belül figyelembe kell venni a feldolgozandó képek méretét és mennyiségét. Általában előnyösebbek az interaktív beavatkozást lehetővé tevő, többfelhasználós rendszerek. Ez érthető, hiszen a feldolgozási folyamatban sokszor van szükség szubjektív, vizuális kiértékelésre, ilyenkor az egyfelhasználós rendszer nincs kihasználva.

A digitális kép - a felbontás finomságától függően - 10^5 - 10^6 képpontból és képpontonként 6-12 bithől áll, így

tárolásához 0,1-1,5 Mbyte tárkapacitásra van szükség. A feldolgozás bonyolultságától függően képpontonként 10-10³ műveletet kell elvégezni. Nyilvánvaló, hogy a felmerülő feladatokat - elviselhető futásidők mellett - csak igen nagy tárkapacitású és nagy sebességű gépekkel lehet megoldani.

A képfeldolgozási feladatok sem egyformák, egyes esetekben a számítási sebesség, máskor a tárkapacitás és igen gyakran mindkettő együtt jelenti a szűk keresztmetszetet [1-34].

A Neumann-elvű (soros) számítógépek igazában sohasem optimálisak a képfeldolgozási feladatokra, a hatékonyság a párhuzamos szervezésű gépekkel növelhető.

A megvalósított rendszerek három csoportba sorolhatók:

- futószalag szervezésű (pipe-line) gépek,
- párhuzamos működésű homogén rendszerek és
- speciális inhomogén rendszerek.

A futószalag szervezésű gépekben a képi adat az egymást követő részműveleteket végrehajtó, párhuzamosan működő processzorokból álló "futószalagon" halad végig. A feladatot lehetőleg közel azonos idő alatt elvégezhető részműveletekre bontják fel. A rendszer inicializálását és az első adat végigfutását követően minden újabb adat feldolgozása a leghosszabb részműveletnek megfelelő idő alatt történik meg. Az ilyen rendszerek előnye az egyszerű adatkapcsolat a feldolgozási lépések között, valamint hogy a műveleteket az eredeti adatokon lehet elvégezni. A saját mikroprogrammal, önállóan működő futószalag-feldolgozómuvet a gyakorlatban rendszerint hozzákapcsolják valamilyen hagyományos elvű számítógéphez, vagy egybeépítik valamilyen képfeldolgozó egységgel.

A párhuzamos homogén rendszerek SIMD (single instruction stream - multiple data stream) strukturát valósítá-

nak meg. Általában egy *vezérgép* (master) dönti el, hogy milyen utasítást kell az összes *aldrendelt* (slave) feldolgozó egységnek végrehajtania. A feldolgozó egységek általában saját tárral is rendelkeznek és fejlettebb rendszerekben egymás között is tudnak adatot cserélni. Az ilyen típusu, eddig megvalósított rendszerek (CLIP-IV, DAP) processzorai rendszerint bitműveleteket végeznek, ezért használatuk a szokásostól erősen eltérő gondolkodást igényel. Ebben a kategóriában a 128×128 processzorból álló MPP rendszer jelenleg a legnagyobb.

A *speciális inhomogén rendszerek* igen változatos, egymástól is eltérő architektúrákat valósítanak meg. Közös jellemzőjük, hogy az egyes speciális funkciókat végrehajtó egységek egy-egy képfeldolgozási algoritmus *közvetlen* hardver implementációi, amelyek egymással és a gyors képtárolóval nagy sebességű *sin* révén vannak kapcsolatban. Tekintettel arra, hogy ezekben a rendszerekben a speciális funkciókat végrehajtó egységek saját programvezérléssel és tárolóval rendelkeznek, lényegében MIMD (multiple instruction stream - multiple data stream) elvű rendszereknek tekinthetők. Ismertebb megvalósításuk a PICAP II és a ZMOB rendszer.

Bár a képfeldolgozás területén feltétlenül a párhuzamos architektúrájú rendszereké a jövő, jelenleg a hazai gyakorlatban még nem jelentek meg. Ezért a továbbiak megértését nem akadályozza, ha részletekbe nem bocsátkozunk. (Lásd pl. [A-1], [1-34], [1-37], [1-38].)

1.3.2 SPECIÁLIS HÁTTÉRTÁRAK

A digitális képek mérete általában nem teszi lehetővé, hogy a teljes képet a főtárban tároljuk. A valósídejű rendszerektől eltekintve - ahol a kiértékelés után már

nincs szükség az eredeti képre - rendszerint igény van több kép *egyidejű* felolgozására és a képek *archiválásra* is. Ezért megfelelő háttértár-kapacitásról kell gondoskodni. A feldolgozás alatt álló képet rendszerint mágneslemezeken tárolják, archiválás céljaira viszont mágnesszalagot használnak. Bár a hagyományos mágneses tárolók kapacitása fokozatosan nő (a méret egyidejű csökkenése mellett), mégis elenyésző például a különböző műholdas rendszerek által rögzített hatalmas tömegű adat mellett.

A probléma megoldása szempontjából nagy reményekkel kecsegtet a *képlemez*, amelynek az elmúlt években megjelent néhány típusa máris lenyűgöző paraméterekkel rendelkezik. A jelenlegiek tárolókapacitása 2-10 Gbyte között van, az átviteli sebesség 5-50 Mbit/s között változik [1-39]. Az információt lézersugárral *beégetett lyukak* tárolják, a leolvasás is lézersugárral történik. A képlemez egyik hátránya jelenleg az, hogy nem törölhető. Ezt a nagy tárolási kapacitás némiképp kompenzálja, bizonyos esetekben pedig (pl. archiváláskor) ez a tulajdonság adatvédelmi szempontból előnyként is elkönnyvelhető.

Itt érdemes megemlíteni egy szerényebb, de viszonylag olcsó és mégis nagy tárolási kapacitású eszközt, a *képmagnót*. Az SZKI egyik kísérleti rendszerében sikeresen megoldottuk a számítógéppel vezérelt címzést és a cím szerinti visszakeresést is.

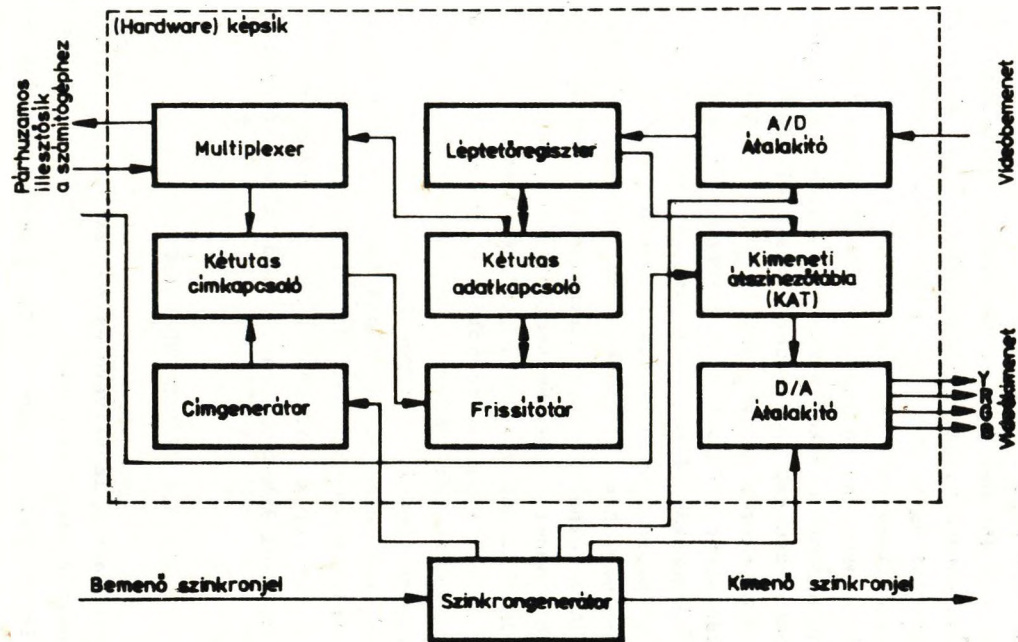
1.3.3 A KÉPMŰ

Azoknak a képfeldolgozó rendszereknek, amelyekben analóg képeket dolgoznak fel, illetve amelyekben az eredményeket vizuálisan is meg kell jeleníteni, alapvető kiegészítő egysége a *képmű*. Ez teremt kapcsolatot az analóg működésű, rendszerint tv-technikai képfelvevő és képmegjelenítő eszközök, valamint a digitális működésű számítógép között.

A képmű (hardver) *képsík(ok)*ből, valamint szinkronizáló és vezérlő áramkörökből áll. Egy képsík elvi felépítését az 1-21. ábrán adjuk meg. Elvi működése a következő [1-4]:

A bemenő videojel az *analóg-digitális (A/D) átalakító* kerül. Ez egyszerre végzi a mintavételezést a megfelelő időpillanatokban és a kvantálást. A kimenetén megjelenő világosság kódok a *frissítőtár* megfelelő című byte-jába íródnak, s így létrejön a digitális kép. (A tárcimek és a képpontok geometriai helyzete között kölcsönösen egyértelmű megfeleltetés van. A mintavételezési és a tárciklusidő közötti különbséget léptetőregiszter hidalja át.) A *frissítőtár* - a kétutas adatkapcsolón és a multiplexeren keresztül - közvetlen összeköttetésben van a számítógép főtárával, rendszerint *cikluslopásos* elven működő gyors csatolókon (DMA = direct memory access) révén. A digitális képet így be lehet olvasni a számítógépbe, illetve a feldolgozás eredményét vissza lehet írni a frissítőtárba. Az adatátvitel kezdő címét a számítógép adja meg; ez a *kétutas címkapcsolón* keresztül jut a frissítőtár címregiszterébe. Folyamatos átvitel esetén a soron következő képpont címét ez a regiszter szolgáltatja.

A képmű *digitális-analóg (D/A) átalakítója* automatikusan generálja a frissítőtár mindenkori tartalmának megfelelő, a *kimeneti átszinező táblával* (lásd 1.3.5 alpont) átszinezett videojelet; ennek felhasználásával a *képmonitron* megjeleníthető a frissítőtárban levő kép. A D/A átalakító az *Y világosságjelen* (fekete-fehér videojel) kívül a *színes (R,G,B) videojeleket* is előállíthatja, ha a világosság kódokhoz valamilyen szint rendelünk hozzá. Ekkor - ha a képmonitor erre alkalmas - a fekete-fehér képet *dlszínesen* (lásd 2.5.1 alpont) is láthatjuk. (A frissítőtár folyamatos kiolvasásához szükséges címeket a *címgene-*



1-21. ábra. Képmű egy képsíkjának elvi felépítése

rátor állítja elő.) A kétutas adatkapcsoló a címzési módnak megfelelően irányítja az adatokat a D/A, illetve a számítógép felé.

Az SZKI-ban kifejlesztett, első használható színes megjelenítő [1-45] még a vezérlő számítógép főtárát használta frissítőtárnak, ami természetesen erősen korlátozta a geometriai és a gradációs felbontást. A tapasztalatok felhasználásával 1978-79-ben kialakított CDP (Colour Display Processor) néven ismertté vált rendszer [1-13] lényegében a fenti elveket valósítja meg. A berendezést kisebb módosításokkal és korszerűsítésekkel a Híradástechnika Szövetkezet ma is gyártja ([1-1], [1-9]). A frissítőtár 288 sorból, soronként 384 képpontból álló képet tárol, képpontonként 6 biten. Ez azt jelenti, hogy 64 szürkéségi fokozat, vagy - álszínes kép esetén - 64 szín jeleníthető meg. Készült olyan képmű is, amely négy képsikből áll. A felhasználó egy mikrogéppel vezérelt videojel-kapcsoló segítségével határozhatja meg, hogy melyik képsik melyik kimenő videojele kerüljön a monitorra. (Ha pl. három képsik Y jelét kapcsoljuk a monitor R,G,B bemenetére, lehetőség van *valódiszínes és kompozit* képek előállítására is; lásd 2.5.1 alpont).

Az SZKI-ban jelenleg két képmű kifejlesztése folyik; az egyik frissítőtárában max. $512^2 \times 8$ bites, a másikéban max. $1024^2 \times 16$ bites képek tárolhatók. Megfelelő kiépítés esetén mindkettővel fel lehet dolgozni valódiszínes képeket is.

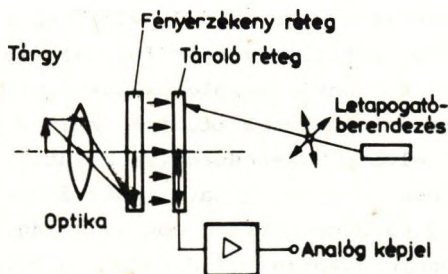
1.3.4 KÉPFELVEVŐ ESZKÖZÖK

A továbbiakban képfelvevő eszköznek csak az olyan berendezéseket tekintjük, amelyek a kimenetükön digitalizálható analóg képjelet, illetve közvetlenül digitális képet

adnak. (Ebben az értelemben a fényképezőgép nem képfelvévő eszköz.)

Az 1960-as években, amikor a képfeldolgozó rendszerek dinamikus fejlődésnek indultak, a televíziós képfelvételi, átviteli és vételi rendszer már széles körben elterjedt, így kézenfekvően adódott a képfeldolgozó rendszerek bemeneti eszközeként az (optoelektronikus) *tv-kamera*.

Az elektronikus képfelbontás elvét az 1-22. ábra alapján érthetjük meg. A képet - megfelelő optikával - egy fényérzékeny félvezető rétegre képezik le. A fotoelektromos hatás következtében ebben az egyes képelemek fénysűrűségével arányos elektromos jel keletkezik, amittől - töltésmegosztás révén - feltöltődik a fényérzékeny réteg mögötti tárolóréteg. Ezt a "töltésképet" egy megfelelően vezérelt elektronsugár soronként letapogatja, és a töltéskisülés hatására a tárolólemezhöz csatlakozó ellenálláson a fénysűrűséggel arányos feszültségesség jön létre. Ezt az analóg elektromos képjelet (*egyszerű videójelnek* nevezik. Ehhez legtöbbször hozzákeverik a kép- és a sor-szinkron jelet is, így áll elő az *összetett videójel*. (A szocialista országokban elfogadott *OIRT szabvány* sze-



1-22. ábra. Kamera működési elve

rinti SECAM rendszerben a képet 625 sorra bontják (ebből azonban csak 576 látható, a többi a képváltás idejére esik); a vízszintes/függőleges képarány 4:3, így egy sor 768 képpontból áll. A képet két félképből (páros, illetve páratlan sorszámú sorok) fésűsen egyesítik. A képváltási frekvencia 25 Hz, így egy kép felvételére 40 ms jut.)

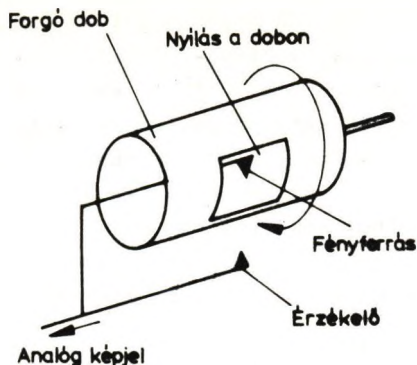
A fényérzékeny réteg fizikai tulajdonságainak változtatásával különböző kameratípusokat fejlesztettek ki (hazai vonatkozásban [1-30]). Legújabbban elterjedőben vannak a tértöltéscsatolású (CCD) kamerák, amelyek nagyfokú linearitásukkal tűnnek ki.

A *színeskamerában* az optika a képet piros (R), zöld (G) és kék (B) színösszetevőre bontva képezi le a megfelelő színsávra érzékeny félvezető rétegre, ennek megfelelően a tárolórétegen is egyszerre 3 töltéskép keletkezik. A "színes" videójelet 3 együttlutó, de egymástól függetlenül vezérelt elektronsugárral állítják elő.

A *hőkamerák* infravörös sugárzási tartományra érzékenyített optikával és félvezető réteggel rendelkeznek. (Általában a 2-5 μm , illetve a 10-12 μm hullámhossztartományban működnek.) A létező rendszerek egyik tv-szabvány-nal sem egyeznek meg, ezért a "hő"-videójel számítógépes feldolgozása külön csatolóegység kifejlesztését igényli.

A kamerák - viszonylagos olcsóságuk miatt - a mai napig a legelterjedtebb képbeviteli eszközök. További előnyük a nagy felvételi sebesség, valamint az, hogy ipari környezetben is alkalmazhatók. Hátrányuk a kis felbontóképeség és a geometriai pontatlanság.

A *forgódobos fotódigitalizáló* film vagy fénykép bevitelére alkalmas. A filmet a forgódob palástjában kivágott nyílásra erősítik, és a jól fókuszált, keskeny fénysugárnyalábot - a filmen való áthaladás után - fototranzisztor érzékeli (lásd 1-23. ábra). A dob forgása és az



1-23. ábra. Forgódobos fotódigitalizáló működési elve

együttmozgó megvilágító-érzékelő rendszernek a dob tengelyével párhuzamos elmozdulása a kép soronkénti letapogatását eredményezi. Az eszköz a felvett képet rögtön digitalizálja is, így kimenő jele közvetlenül tárolható a számítógépben. Korábban csak nyugati rendszerek léteztek (egyik legismertebb típusáról lásd [1-8]), de ujabban az NDK-ban is kifejlesztettek ilyen eszközöket.

A fotódigitalizáló ez idő szerint a legpontosabb képfelvevő eszköz: a képpont átmérője és a visszaállási pontosság 10-100 μm között van; a képet 512×512, ujabban 2048×2048 képpontra bontja, és általában 8 bitre kvantálja. Hátránya viszont, hogy a felvételi idő jelentősen megnő: 5-10 perc a képmérettől és a digitalizálási pontosságtól függően. Léteznek CCD-diódasoros síkdigitalizálók is, ezek sebessége lényegesen nagyobb. Megfelelő színszűrők alkalmazásával színes film digitalizálására alkalmas berendezéseket is kifejlesztettek.

Az ürfelvételek széles körű felhasználása miatt példaképpen vázlatosan ismertetjük a LANDSAT műhold *többcsatornás* felvevő rendszerének működési elvét [1-14]:

A Föld felszínéről visszavert fénysugár oszcilláló

mozgást végző tükörrre, majd onnan a fénybontó prizmára esik. A prizma a beeső sugarat különböző hullámhosszuságu komponenseire bontja, ezek közül a fényérzékelők négy hullámhossztartományba (két látható és két infravörös sávba) eső sugarakat érzékelnek. Minden sávban hat érzékelő van, azaz egyszerre hat sort pásztáznak végig. A földfelszín letapogatása a műhold haladásának és az oszcilláló tükör mozgásának megfelelően történik. Az érzékelők által keltett elektromos jelet már a műhold fedélzetén digitalizálják.

A *jeleneteket* (≈ 32 Mbyte/jelenet) rádión továbbítják a Földre. A vevőközpontokban nagy irásssűrűségű mágnesszalagra rögzítik az adást, majd ezekből különböző korrekciók után állítják elő a felhasználók részére a számítógéppel közvetlenül feldolgozható (CCT) mágnesszalagot. Megjegyezzük, hogy az újabb rendszerekben mind a sávok száma, mind a felbontás, mind pedig a kvantálási szintek száma nagyobb így az átviendő információ ennek többszöröse is lehet (lásd pl. [1-12], [1-32], [1-43], [1-49]).

Az űr- és légifelvételek készítésekor - az egyes felhasználási területek igényeinek megfelelően - számos más érzékelőt is alkalmaznak, beleértve a mikrohullámu (pl. radar) rendszereket is.

1.3.5 KÉPMEGJELENÍTŐ ÉS KÉPRÖGZITŐ ESZKÖZÖK

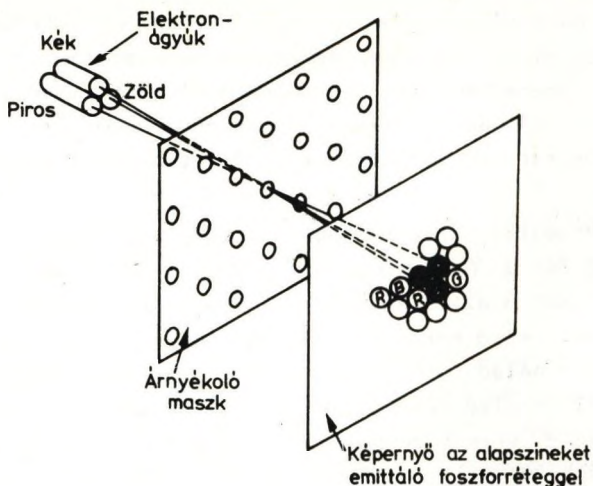
Kézenfekvőnek látszik, hogy a képek *megjelenítésére* is tv-technikai eszközöket alkalmazzunk. A nehézséget az okozza, hogy a monitor képernyőjén megjelenő képet állandóan (legalább másodpercenként 25-ször) újra kell rajzolni ("felfrissítés"), hogy villogásmentesen lássuk. Ez a feladat tulságosan leterhelné a számítógépet, hiszen pl. egy 512×512 képpontból álló kép és 25 Hz-es ismétlési frekven-

cia esetén 6,55 Mbyte adatot kell kivinni másodpercenként, megfelelő sebességű csatornán keresztül! A raszteres megjelenítőkből ezt a problémát a képmű, (lásd 1.3.3 alpont) oldja meg: a számítógépnek csak akkor kell a frissítőtárat feltöltenie, ha új képet akarunk megjeleníteni.

A *képmonitor* elektronsugara a képmű videojele alapján rajzolja fel a (két félképből álló) képet a képernyőre, a továbbiakban ezt *látványképnek* nevezzük. Az elektronsugár (a bal felső sarokból indulva) balról jobbra és felülről lefelé halad. Először a páratlan sorszámú sorokat rajzolja fel (= első félkép), majd - a 2. sor elejére visszaugorva - ezek közé berajzolja a páros sorszámú sorokat (= második félkép). Az elektronsugár erősségét - és ezáltal a monitor képernyőjének belső felületén egy adott pontban gerjesztett foszforszemcse fénysűrűségét - a frissítőtárban tárolt világosságkódoknak megfelelő videojel határozza meg.

Színes (raszteres) monitorok esetén három különböző (piros, zöld és kék) színben világító foszforpontokból álló réteget visznek a képernyő felületére, amelyeket három együttlfutó, de erősségét tekintve egymástól függetlenül vezérelt elektronsugár gerjeszt (lásd 1-24. ábra). Az egyes képpontok színezetét a megfelelő 3 foszforpont egyidejű, de különböző mértékű gerjesztésével állítják elő (*additív színkeverés*).

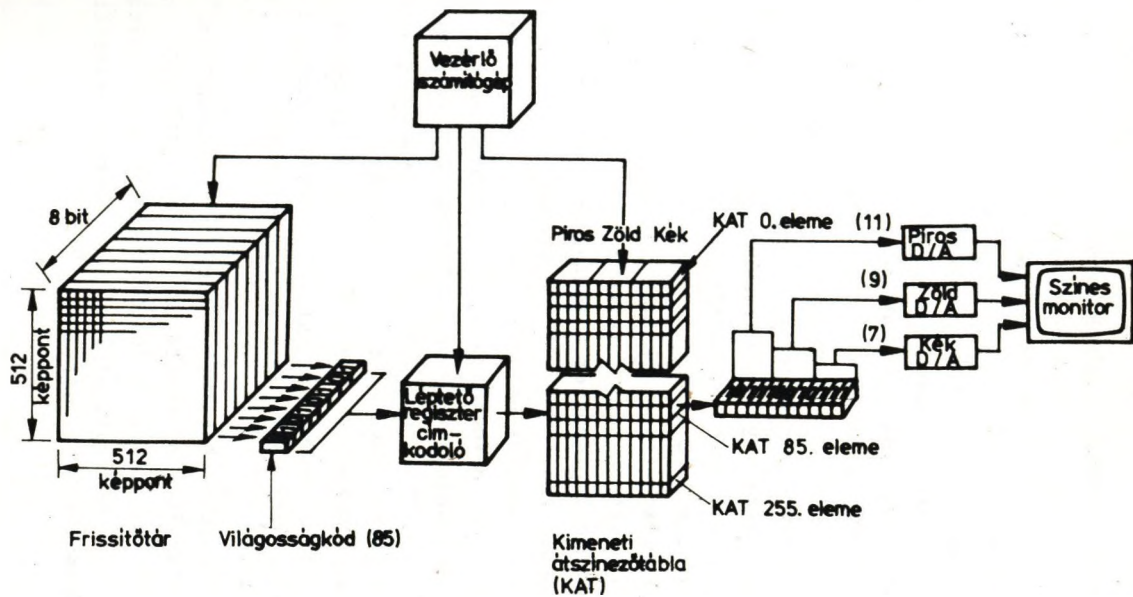
A mai rendszerekben általánosan elterjedt a *kimeneti dtsszínezőtábla* (a továbbiakban: KAT, angol elnevezés szerint LUT = look-up-table) használata. Ennek segítségével tetszés szerint meg lehet változtatni a látványképben a képpontok fénysűrűségét és/vagy színét anélkül, hogy a frissítőtár tartalma megváltoznék. A KAT elemeinek száma megegyezik a digitális képben megengedett világosság-, illetve szinkódok számával, értékészletük pedig a megje-



1-24. ábra. Színes raszteres monitor működési elve

lenithető szűrkeségi szintek, illetve színek kódjainak halmaza. Alkalmazása esetén - a frissítőtárban levő világosságkódok helyett - a KAT megfelelő című elemeiben található kódértékek fogják meghatározni a látványképben az egyes képpontok fénysűrűségét, illetve színét (természetesen a D/A átalakítás után; lásd 1-25. ábra). A KAT előnye, hogy programból feltölthető, módosítható és a kódcsere időveszteség nélkül történik.

A kép *rögzítésének* legegyszerűbb módja a képernyő *lefényképezése*. A polaroid technika elterjedése különösen meggyorsította ezt az eljárást. A fellépő torzító hatások (hordórtorzítás, fényvisszaverődés stb.) kiküszöbölésére speciális rendszereket fejlesztettek ki. Egy lehetséges megoldás, hogy különleges, sík felületű, monokromatikus képernyőn külön jelenítik meg az egyes színösszetevőket, és megfelelő szűrők közbeiktatásával fényképezik le a képet.



1-25. ábra. Kimeneti átszinezőtábla (KAT) működési elve

Az előző pontban ismertetett dobos fotodigitalizálókat általában úgy készítik, hogy megfelelő kiegészítésekkel alkalmasak legyenek a képek nagy pontosságú felrajzolására is. Ezekben a *filmiró* berendezésekben egy jól fókuszált 10-100 μm átmérőjű fénysugár fényerősségét modulálják a tárból kiolvasott világosságkódok szerint. A fénysugár a képsoroknak megfelelően pásztázza végig a forgódobra felerősített filmet, s így "rárajzolja" a képet, amely előhívás után válik láthatóvá.

A képrögzítők további fajtái a különböző szines *plotterek*, amelyek a képet képpontonként rajzolják ki. Az egyik impakt nyomtatási módszer a *görgős rajzolófejes eljárás*. A rajzpapírt egy forgó hengerre erősítik fel. A képpontokat az alapszínekkel (fehér papír esetén *magenta, sárga, türkizkék*, a jobb fedés miatt rendszerint még külön *fekete* is) festékezett görgők állítják elő. A dob egy körülfordulása alatt egy képsor felrajzolása történik meg úgy, hogy a vezérlő rendszer a megfelelő időpontban (elektromágnesek segítségével) a papírhoz nyomja a görgőket. Ezalatt a görgőket hordozó szán éppen egy sor szélességével mozdul el (a forgástengellyel párhuzamosan), így a következő körülfordulás során a következő képsor rajzolódik ki. A keverékszínnek az alapszínnek egymáshelyezésével jönnek létre (*szubtraktív* színkeverés).

Ilyen elven működik a hazai fejlesztésű, COROLLPRESS-4 elnevezésű szines, dobos plotter is, amelyet elterjedten alkalmazunk az SZKI képfeldolgozó rendszereiben ([1-11], [1-31], [1-48]).

A *tintasugaras* (tintalövő = ink-jet) eljárásnál az ábrát kis fuvókákból a megfelelő időpontban kilövellő festékszemesecskék hozzák létre. A festékcseppeket elektrosztatikus módszerekkel egyrészt porlasztják, másrészt eltérítik a kívánt tónusok elérése céljából.

1.3.6 INTERAKTIV ESZKÖZÖK

A képfeldolgozásban általában a vizuális ellenőrzés vagy kiértékelés befolyásolja a feldolgozás további menetét, ezért a kényelmes visszacsatolási lehetőséget nyújtó interaktív eszközök fontos szerepet töltenek be.

A *billentyűzet* (klaviatúra) a leggyakrabban alkalmazott interaktív eszköz, ezen keresztül lehet parancsokat és paramétereket megadni a programok vezérlésére. Gyakran találunk kiegészítő billentyűzetet, illetve - ezek hiányában - egyes billentyűkhöz programból rendelnek megkülönböztetett értelmet.

A legtöbb klaviatúrán található külön funkcionális billentyűzet, ezt használhatjuk a *grafikus fénypont* (cursor) mozgatására. (A grafikus fénypont a képmonitoron megjelenő villogó jel (szálkereszt, nyíl stb.), amely pozicionálás és azonosítás céljából tetszőleges irányban mozgatható. Ettől megkülönböztetjük a program futtatásra szolgáló terminál képernyőjén látható *alfanumerikus fénypontot*.)

A klaviatúra hátránya, hogy a billentyűk kezelése elvonja az operátor figyelmét a képmonitorról. Ennek kiküszöbölésére speciális kapcsolókat és vezérlő eszközöket használnak. Ilyen a gyakran alkalmazott *botkormány* (joystick), a *pozicionáló golyó* (track ball) vagy ujabban az *egér* (mouse). Ezek segítségével a felhasználó úgy mozgathatja a fénypontot, hogy a szemét nem kell levennie a képernyőről. Mindegyikben két, egymásra merőlegesen elhelyezett érzékelő jelzi az eszköz x , illetve y irányu elmozdulását, ezt használják azután a fénypont vezérlésére.

A *digitalizáló táblák* kölcsönösen egyértelmű megfeleltetést létesítenek a képmonitor és a tábla felülete között, s így közvetlenül vihetünk be grafikus adatokat.

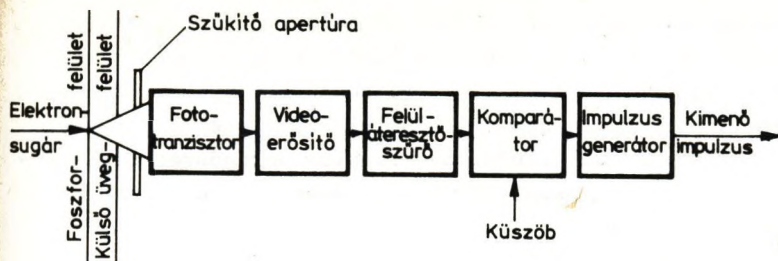
Működésük elvét tekintve két alaptípust különböztethetünk meg:

- A *vezetékrcsós rendszerűekben* a kezelő által vezetett szonda elektromágneses jelet bocsát ki, amely feszültséget indukál a szigetelő rétegek között elhelyezett vezetékhalóban. Ebből határozható meg a szonda x , illetve y koordinátája.
- Az *oldallos érzékelő rendszerekben* a tábla két szomszédos oldalán elhelyezett szalagmikrofonok a szonda által keltett ultrahangot érzékelik. Itt a szonda helyzetét az időzési viszonyokból határozzák meg.

A digitalizáló táblákat nemcsak pozíciómeghatározásra lehet használni: a tábla szélein fenntartott funkciómezőkhöz különböző funkciókat rendelnek, amelyeket rámutatással lehet aktiválni.

Megemlítjük még, hogy a digitalizáló táblák rajz vagy kézírás bevitelére is alkalmasak. A rajzelemek vagy a szöveg felismerésére természetesen megfelelő programokat kell kidolgozni.

A *fénytollal* (light pen) (gyakran *fényceruzának* is nevezik) a kezelő közvetlenül a képmonitoron keresztül teremthet kapcsolatot a rendszerrel. A kinézetre toll formájú eszköz - elnevezésével ellentétben - *nem ír* a képernyőre, hanem leolvassa egy képpont koordinátáit. Elvi felépítése az 1-26. ábrán látható: a belépő elektronsugarat a szűkítő fényrekesz egy fototranzisztorra fókuszálja; az erősítőt a fénytoll kapcsolójának benyomása aktiválja. (A kapcsoló általában úgy működik, hogy a fénytollat a képernyőhöz kell nyomni, de szoktak külön kapcsolót is beépíteni.) A képet felrajzoló elektronsugár elhaladásakor a fototranzisztor egy impulzust generál, amelyet - erősítés és szűrés után - a komparátor összehasonlít egy küszöbfeszültséggel. A küszöbértéket úgy kell megválasztani,



1-26. ábra. Fénytoll elvi felépítése

hogy a szórt világításból és a zajból eredő impulzusok hatástalanok maradjanak. Ha a feszültségváltozás meghaladja a küszöbértéket, az impulzusgenerátor előállítja a kimenő jelet. A képszinkronjel ismeretében az időkülönbségből egyértelműen meghatározható a fénytoll pozíciója. Erre a pozícióra a kezelőprogram generálhat egy pontot, így végül valóban rajzolhatunk a fénytollal. A fénytoll fő hátránya, hogy a sötét területen levő képpontokat nem tudja azonosítani; a problémát ez ideig nem sikerült kielégítően megoldani.

1.3.7 PROGRAMESZKÖZÖK

A képfeldolgozási feladatok speciális hardvereszközökkel történő megoldásához természetesen speciális programeszközökre is szükség van. A képfeldolgozó programokat általában magasan kvalifikált szakemberek fejlesztik ki, mivel ez a munka jelentős matematikai ismereteket feltételez, és jártasságot igényel a különböző határtudományágakban. Tovább nehezíti a helyzetet, hogy a futási idők csökkentése érdekében a programokat általában alacsony szintű (assembly) nyelven kell írni. A sokszor ellentmondó követelményeket csak különféle programozási forgásokkal

és kompromisszumok árán lehet teljesíteni. Ezért és a különböző speciális periférikus eszközök használata miatt a képfeldolgozó szoftver áttelepíthetősége (portabilitása) rossz, más gépeken nehezen adaptálható. Ez az oka annak, hogy - bár számos képfeldolgozó programrendszer létezik - szinte minden új eszközhöz többé-kevésbé új programokat fejlesztenek ki.

Egy széles körű felmérés [1-35] szerint a számítógéptől független képfeldolgozó programok és programkönyvtárak készítésekor legtöbb esetben a *FORTRAN* nyelvet részesítik előnyben. Általánosan elterjedt képfeldolgozás-orientált nyelv nem született még meg; ujabban a PASCAL és a C nyelv hódít tért. Egyes rendszerekben az adott rendszer architektúrájához igazodó specifikus nyelveket fejlesztettek ki.

A képfeldolgozó rendszerek kialakításánál a *felhasználói terület*, a felhasználói igények jelentős szerepet kapnak; nincsenek általános, mindenre használható rendszerek. Egész más megközelítést igényel pl. egy ipari környezetben működő, célorientált, kamera bemenettel rendelkező automatikus minőségellenőrzést végző rendszer kialakítása, mint egy ürfelvételek kiértékelése céljából létrehozandó képfeldolgozó központé. Ennek ellenére számos közös vagy hasonló módszerekkel megoldható feladat jelentkezik.

A képfeldolgozó programrendszerek fejlesztésekor különös gondot kell fordítani az *adatstruktúra* kialakítására és a be/kiviteli rutinok megtervezésére. Az alkalmazói programok kifejlesztése során célszerű közös csoportokba összevonni az azonos elven működőket. Ez csökkenti a programrendszer méretét és könnyebbé teszi a karbantartást. Számos példát találhatunk a különböző rendszerek felépítésére az irodalomban (pl. [1-7], [1-11], [1-18], [1-31], [1-35], [1-46], [1-48]).

2. KÉPJAVÍTÁSOK

A képfelvételi körülmények, a mintavételező és kvantáló rendszernek jelentősen befolyásolják a feldolgozandó digitális kép minőségét. A valós világ leképezése, esetenként a képnek a feldolgozás helyére történő továbbítása (pl. üreszközök által készített képek átvitele a vevőközpontba rádiócsatornán, vagy feldolgozó központok közötti adatcsere) során különböző torzító- és zajhatások érik a képet. Ezeknek a hatásoknak a kiküszöbölésére vagy legalábbis csökkentésére szolgáló módszereket és eljárásokat - a megvalósítandó cél szerint - két csoportba sorolhatjuk: az egyiket képhelyreállításnak, a másikat képjavításnak nevezzük.

A *képhelyreállítás* (restoration) során - mint az elnevezés is jelzi - arra törekszünk, hogy a rendelkezésre álló képből előállítsuk azt az ideális képet, amely megfigyelhető lett volna a különböző zavaró, torzító hatások nélkül.

Hatékony képhelyreállítási eljárás tervezéséhez alapvetően fontos a torzító hatásokat leíró matematikai modell kialakítása. Ezt követően meg kell határozni azt az inverz transzformációt, amely a modell által leírt hatásokat kiküszöböli, és ezt az inverz transzformációt kell alkalmazni a rendelkezésre álló képre. Nyilvánvaló, hogy csak *jó modell* esetén várható eredményes helyreállítás.

Bár a fenti, triviálisnak mondható elvek az egyszerű megoldás látszatát kelthetik, a valóságban a helyzet sok-

kal bonyolultabb, amit az is jól mutat, hogy a problémát teljes komplexitásban megoldó képhelyreállítási rendszer *nem létezik*. Figyelembe véve a potenciális zavaró tényezőket - például az optikai rendszer hibáit, az érzékelők nemlineáris voltát, az atmoszférikus hatásokat, a képelmozdulásból eredő homályosságot, a különböző geometriai torzulásokat stb. - érthetővé válik, hogy reménytelen teljes körű modell megalkotására vállalkozni. Ennek megfelelően egyes részproblémákra lehet csak modelleket kidolgozni, a torzulások sajátosságait kihasználva. A modell kidolgozása általában méréseken alapul, amelyeknek az a céljuk, hogy meghatározzák a vizsgált leképező rendszer hatását egy tetszőleges képmezőre. Néha lehetséges a determinisztikus leírás, máskor viszont a sztochasztikus megközelítés előnyösebb.

A képhelyreállítási eljárások alkalmazásának - a modellalkotás nehézségén túlmenően - másik problémája, hogy a felállított modell, illetve az ezen alapuló inverz transzformáció rendszerint olyan bonyolult összefüggéseket eredményez, amelyeket még a mai nagy teljesítményű számítógépekkel sem lehet elfogadható idő alatt megoldani.

A fentiekkel magyarázható, hogy a különböző általános célú képfeldolgozó rendszerekben (így a szerzők gyakorlatában is) ritkán kerül sor helyreállítási algoritmusok programjainak kidolgozására; ezért ezekkel *általánosságban nem* foglalkozunk. Kivételt képez a geometriai korrekciók kérdésköre, amelyet jelentősége és speciális módszerei miatt külön, a 3. fejezetben tárgyalunk. (A helyreállítás iránt érdeklődő olvasóknak ajánljuk az [A-13], [2-10] műveket.)

2.1 ELVI ÁTTEKINTÉS

A *képjavitási* (enhancement) eljárások célja, hogy a képet a további kiértékelés vagy feldolgozás szempontjából előnyösebb formába hozzuk, és nem törekszünk arra, hogy minél jobban megközelítsük az eredetit vagy az ideális leképezést. Éppen ellenkezőleg, esetenként egyes jellegzetességek kiemelésével, hangsúlyozásával kapjuk a számunkra kedvezőbb eredményt. (Itt említjük meg, hogy az angol terminológia jobban tükrözi a két megközelítési mód lényegét, ugyanis általában a "helyreállítást" (restoration) és a "fokozást" (enhancement) együttesen illeti "javítás" (improvement) megnevezéssel. A magyar szóhasználatban a képjavitási módszerek közé csak az angol terminológiában "image enhancement" névvel jelölt módszereket sorolják.)

A meghatározásból látható, hogy a képjavitási eljárások közötti választás vagy egy új algoritmus kialakítása a megoldandó problémától függ. Egy adott módszer az egyik alkalmazásban sikeres, egy másik feladatnál esetleg teljesen használhatatlan lehet. Amiatt, hogy a képjavitási eljárások célja sok esetben az emberi értékelő ("interpretátor") számára kedvezőbb látvány létrehozása, figyelembe kell venni az emberi látórendszerrel kapcsolatos pszichofizikai tényezőket, sőt szubjektív elemeket is. (Utalunk pl. az 1.2.2 alpontban említett kísérletre [1-21], amelyben a felbontás és a kvantálás finomsága közötti összefüggéseket vizsgálták; lásd 2. és 3. kép.)

Egyelőre nem sikerült meghatározni a képminőség objektív mértékeit, sőt még egységesen elfogadott normák sincsenek. Az egységes elmélet említett hiányán kívül valószínűleg ez is oka annak, hogy ez idő szerint a képjavitások általános elmélete *sem létezik*.

A képjavítási eljárások két kategóriába sorolhatók attól függően, hogy a *frekvenciatartományban* (a kép Fourier-transzformáltjával, lásd F4.3.2 alpont), illetve az *eredeti képsíkon* (a világosságkódokkal) dolgozunk.

A frekvenciatartományban végzett javítások a *konvolúcióelméleten* alapulnak (lásd F4.1.2 alpont). Jelölje $g(x,y)$ azt a képet, amelyet az eredeti $f(x,y)$ képből kapunk a $h(x,y)$ eltolásinvariáns operátorral végrehajtott konvolúció után azaz

$$g(x,y) = h(x,y) \otimes f(x,y). \quad (2-1)$$

Ismeretes, hogy a konvolúcióintegrál a frekvenciatartományban a transzformáltak szorzására vezethető vissza, vagyis (2-1)-nek

$$G(u,v) = H(u,v) \cdot F(u,v) \quad (2-2)$$

felel meg, ahol G, H , illetve F a g, h , illetve f függvény Fourier-transzformáltja. Számos képjavítási és helyreállítási eljárás ugyanis úgy fogalmazható meg, hogy a kép Fourier-transzformáltjának kiszámítása után keressük azt a $H(u,v)$ függvényt, amelynek alkalmazása esetén a (2-2) inverz transzformációjával nyert $g(x,y)$ eredményekép valamilyen szempontból előnyösebb, mint az eredeti $f(x,y)$ kép.

Egy másik lehetséges csoportosításban az *eljárások célját* veszik figyelembe. Mivel a legjellegzetesebb három hibatípus a kontrasztosság elszegényedése, zajok hozzákeveredése a képhez és az élek elmosódása, *kontrasztfokozó, zajelnyomó és élkiemelő* eljárásokat szokás megkülönböztetni. Tekintve azonban, hogy sok esetben ugyanaz a módszer más paraméterekkel más célra alkalmazható, valamint hogy a fejezet elején említett okok miatt az eljárások

zöme ad hoc jellegű, heurisztikus megfontoláson alapul, minden rendszerezés tartalmaz önkényes elemeket. Nem célnk teljes körű áttekintést adni (ez valószínűleg lehetetlen és értelmetlen vállalkozás volna), csak a leggyakrabban alkalmazott eljárásokat ismertetjük, kitérve a gyakorlati tapasztalatokra és néhány, az SZKI-ban kifejlesztett ujszerűnek számító megoldásra.

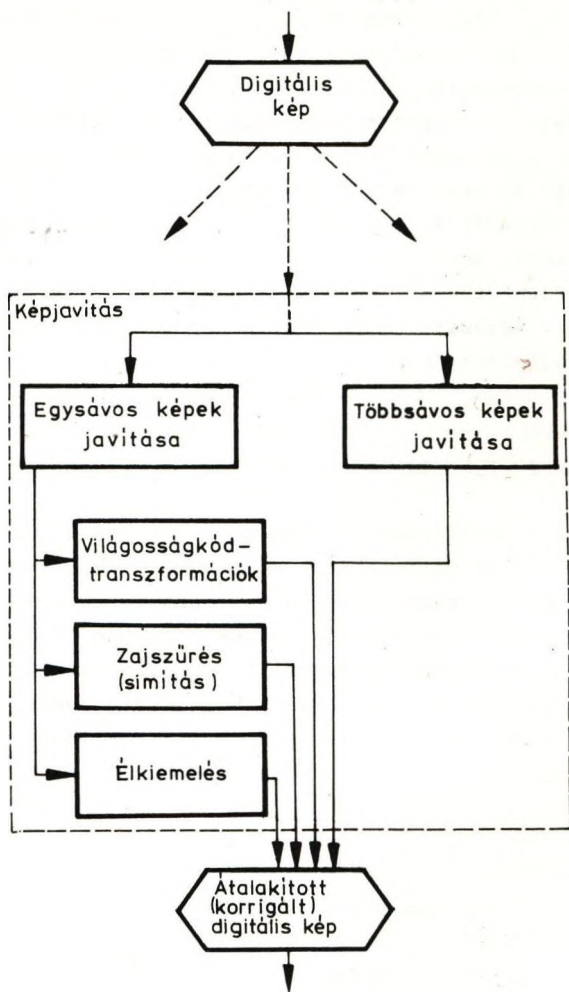
A továbbiakban a 2-1. ábrán látható csoportosítást vesszük alapul. Ennek megfelelően először a világosságkód-transzformációkkal, majd a zajelnyomással és a képesítéssel, végezetül a többsávós képek javításának kérdéseivel foglalkozunk.

2.2 VILÁGOSSÁGKÓD-TRANSZFORMÁCIÓK

Az egyik leggyakrabban előforduló képhiba a nem megfelelő fénysűrűségből, illetve a legképező rendszerben keletkezett fényvesztéséből adódó kontrasztszegénység. Ez szerencsés esetben az egész képen egyenletesen jelentkezik, de igen gyakran a képen belül is változik. A világosságkód-transzformációk célja a kontraszt növelése a világosságkódok *eloszlásának* megváltoztatásával.

2.2.1 A HISZTOGRAM

Elsősorban a képjavításokban - de egyéb képfeldolgozási eljárásokban is - fontos szerepet játszik a *hisztogram*, ami megadja a világosságkódok eloszlását az adott képen. Mivel a gyakorlatban a világosságkódok egyenlőközűek, csak ezzel az esettel foglalkozunk, bár a mondotak könnyen általánosíthatók. Ilyenkor a *hisztogramtáblázat* annyi elemű, amennyi a világosságkódok értékkészlete;



2-1. ábra. A képjavítás területei

jelöljük ezt $\{Q\} = \{0, 1, \dots, J\}$ -vel. Ekkor

$$\chi'(q) = N_q \text{ és } \sum_{q=0}^J \chi'(q) = N,$$

ahol N_q a q világosságkódu, N pedig az összes képpontok száma a digitális képben. A hisztogramot általában *lépcsős-függvénnyel* ábrázoljuk (lásd pl. 4. kép). A vízszintes tengelyen a lehetséges világosságkódokat mérjük fel, a $[q, q+1)$ intervallumhoz tartozó ordináta pedig a q világosságkódu képpontok *relatív gyakoriságával* arányos:

$$\chi(q) = c \frac{N_q}{N}. \quad (2-3/a)$$

A továbbiakban mindig feltesszük, hogy a hisztogram normalizált ($c = 1$), amiből következik, hogy ekkor

$$\sum_{q=0}^J \chi(q) = 1. \quad (2-3/b)$$

Vagyis a hisztogram a q diszkrét *valószínűségi változó* (statisztikai) *eloszlása* (lásd F2.3 pont); de semmit sem mond a világosságkódok *geometriai* eloszlásáról.

Ha a világosságkódot egy folytonos valószínűségi változó (ζ) közelítésekként kezeljük, akkor annak P valószínűsége, hogy értéke a q és $q+\Delta q$ közé esik, valamint a *sűrűségfüggvény* (lásd F2.2 pont) között fennáll az

$$s(\zeta) = \lim_{\Delta q \rightarrow 0} \frac{P(q \leq \zeta < q + \Delta q)}{\Delta q}$$

összefüggés. Ez éppen a (2-3/a) hányadossal közelíthető; vagyis a kvantálás finomításával

$$\chi(q) \rightarrow s(\zeta).$$

(2-4)

A kép általános jellemzőiről sok minden elmondható a hisztogram alapján; hasznos információkat nyújt nemcsak a kijavitáshoz, hanem például a különféle szegmentálási eljárásokhoz is. A 4. képen 6 bitre kvantált digitális képek hisztogramját adjuk meg ($q \in \mathbb{E}63$; lásd 1.2.2 pont). Ránézésre megállapíthatjuk, hogy a hisztogramok rendre egy sötét tónusu (a), egy világos összhatású (b), illetve egy közel kétszintű (c.; fehér papírra fekete betűkkel irt szöveg) képről készültek.

2.2.2 SKÁLÁZÁSOK

Sok esetben javíthatjuk a kép kontrasztját a világosságkódok ujraszkálázásával. A skálázást megfelelően választott átviteli függvényekkel valósíthatjuk meg, amelyek a bemenő kép és az eredménykép világosságkód-értékei közötti leképezést definiálnak.

A továbbiakban a következő jelöléseket használjuk:

$\{Q\}$ az eredeti bemenő értékkészlet;

$\{R\}$ az eredményül kapott kimenő értékkészlet, ($\{R\}$ nem szükségképpen egyenlő $\{Q\}$ -val);

q_m , ill. r_m a megfelelő halmaz legkisebb eleme;

q_M , ill. r_M a megfelelő halmaz legnagyobb eleme;

$q \in \{Q\}$, ill. $r \in \{R\}$ egy képpont eredeti, ill. transzformált világosságkódja (a (k, l) koordinátaargumentumot elhagyjuk, ahol nem zavaró).

Skálázásnak nevezzük a

$$T: [q_a, q_f] \rightarrow \{R\}$$

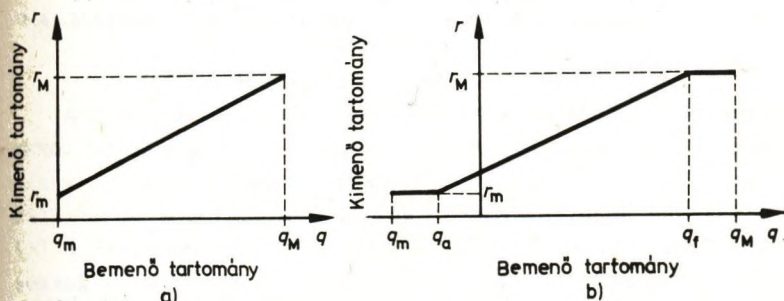
leképezését megvalósító *globális* transzformációkat, ahol $q_a (\geq q_m)$, illetve $q_f (\leq q_M)$ a bemenő világosságkód-intervallum alsó, illetve felső határa.

Legtöbbször *szakaszonként lineáris* transzformációt végzünk a következő összefüggés alapján:

$$r = \begin{cases} r_m' & \text{ha } q < q_a; \\ \text{int} \left[\frac{r_M - r_m}{q_f - q_a} (q - q_a) \right] + r_m' & \text{ha } q_a \leq q \leq q_f; \\ r_M' & \text{ha } q > q_f; \end{cases}$$

ahol *int* a "legközelebbi egész szám" függvény.

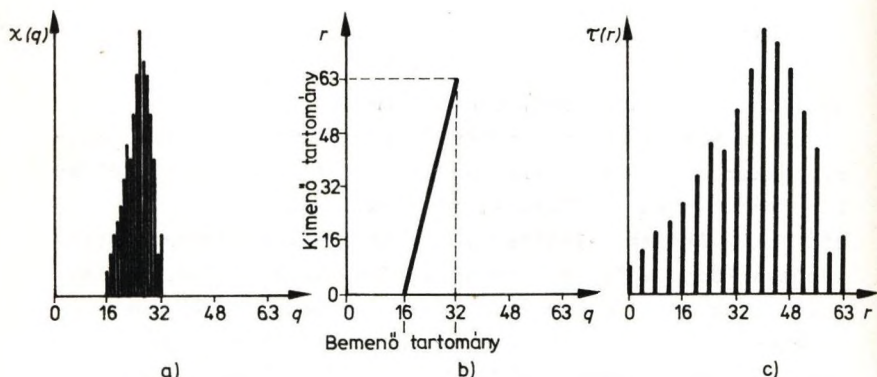
Ha a teljes bemenő intervallumot akarjuk transzformálni, akkor a $q_a = q_m$, illetve $q_f = q_M$ (2-2/a ábra), különben *kódzsugorításról* beszélünk (2-2/b ábra). Az utóbit akkor érdemes alkalmazni, ha az értékes képpontok világosságkódja a $[q_a, q_f]$ intervallumba esik. (Gyakran előfordul, hogy egy előző feldolgozási lépésben számított világosságkódok $\{q\}$ halmaza nagyobb, mint a megjeleníthető világosságkódoké; a 2-2/b ábrán pl. negatív bemenő



2-2. ábra. Lineáris transzformációk: a) a teljes bemenő intervallum transzformálása, b) kódzsugorítás

értékek is szerepelnek.) A kódzsugorítás akkor is célszerű, ha a szélsőséges világosságkódu képpontok száma kevés, és megtartásukkal feleslegesen szegényítendő a többség kontrasztját.

Az eljárást *kontrasztkiemelésnek* nevezzük, ha $\{Q\} \subset \{R\}$, ugyanis ilyenkor ténylegesen megnő a különbség a legsötétebb és a legvilágosabb képpontok világosságkódja között, az eredménykép hisztogramjában pedig "lyukak" keletkeznek (2-3. ábra).



2-3. ábra. Kontrasztkiemelés: a) eredeti hisztogram, b) lineáris transzformáció, c) a transzformált kép hisztogramja

Tetszés szerinti t (kvadratikuss, logaritmikus stb.) transzformációt valósíthatunk meg az

$$r = \text{int} \left[\frac{r_M - r_m}{q_{tM} - q_{tm}} [t(q) - t(q_m)] \right] + r_m$$

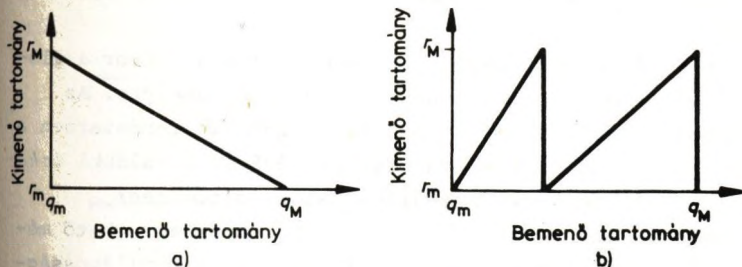
képlettel. Itt q_{tm} , illetve q_{tM} a bemenő kódokra alkalmazott t transzformáció során nyert legkisebb, illetve legnagyobb érték; monoton t függvény esetén természetesen

$$q_{tm} = t(q_m) \text{ és } q_{tM} = t(q_M).$$

Az eddig tárgyalt monoton növekvő átviteli függvények helyett monoton csökkenő vagy nemmonoton függvények is használatosak. *Inverz megjelenítést* érhetünk el a 2-4/a ábrán látható

$$r = \text{int} \left[\frac{r_M - r_m}{q_M - q_m} (q_M - q) \right] + r_m$$

invertáló transzformációval, ha pl. a sötétebb képrészek hordozzák a hasznos információt.



2-4. ábra. Speciális skálázó függvények: a) inverz transzformáció, b) "fűrészfog" transzformáció

A 2-4/b ábrán látható "fűrészfog" átviteli függvényt akkor alkalmazzuk, ha megjelenítő eszközünkön a megjeleníthető világosságok száma kevesebb, mint a képen találhatóké. Ezzel ugyan $\{Q\}$ több eleméhez rendeljük az $\{R\}$ ugyanazon elemét, de ez nem zavaró, ha az azonos kódú részek geometriailag elkülönülnek. (Pl. a Balaton és környéke megjelenítésekor lehetnek azonos színű foltok mind a vizet, mind a partot ábrázoló képrészekben; a tó jel-

legzetes alakja lehetővé teszi, hogy ezeket különbözőképpen értelmezzük.)

A skálázásokat *átszínezésnek* is nevezzük, mivel színes megjelenítés esetén megváltoztatják a képpontok színét. Legegyszerűbben *Q-táblákkal* valósíthatók meg. A *Q*-táblának annyi eleme van, mint a $\{Q\}$ halmaznak, az elemek értékészlete pedig az $\{R\}$ halmaz. A skálázást egyszerűen az

$$r = T(q) \quad (2-5)$$

kódcseré jelenti, ahol $T(q)$ a *Q*-tábla *q*-adik eleme. Különösen kellemes a helyzet, ha a kódcseré hardver uton valósul meg, ilyenkor az átszínezés csak a megjelenített látványképre vonatkozik, a tárolt eredeti kép változatlanul marad.

Ha a képek megvilágítása nem egyenletes, akkor a globális eljárások nem vezetnek kielégítő eredményre. Az SZKI-ban kifejlesztett [1-48] képfeldolgozó rendszerben a [2-7]-beli algoritmus továbbfejlesztésével az alábbi *lokális kontrasztkiemelési* eljárást valósítottuk meg:

Jelölje egy képpont rögzített, de paraméterezhető méretű környezetében a bemenő, illetve a kimenő világosságkódok helyi minimumát és maximumát q_{hm} és q_{hM} , illetve r_{hm} és r_{hM} ; a lehetséges kimenő világosságkódok minimumát és maximumát r_m és r_M ; a bemenő világosságkódot q , a kimenő értéket r . Feltéve, hogy a vizsgált pont környezetében a bemenő $[q_{hm} - q_{hM}]$ világosságkód-intervallum, sőt még ennek p -szerese sem nagyobb a lehetséges $r_M - r_m$ tartománynál, p -szeres nyújtást végzünk, vagyis

$$r_{hM} - r_{hm} = p(q_{hM} - q_{hm}).$$

A kimeneti intervallumot azonban úgy csúsztatjuk el, hogy a végpontok távolságának aránya ne változzék, vagyis

és

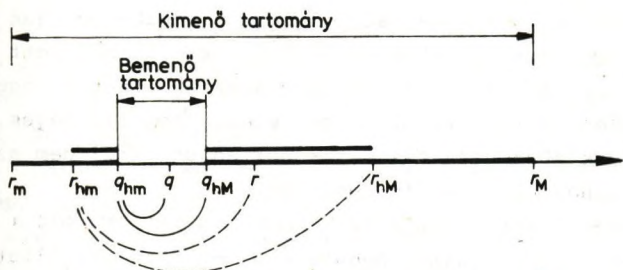
$$\frac{r_M - q_{hM}}{q_{hm} - r_m} = \frac{r_{hM} - q_{hm}}{q_{hm} - r_{hm}}$$

$$\frac{q - q_{hm}}{q_{hM} - q_{hm}} = \frac{r - r_{hm}}{r_{hM} - r_{hm}}$$

A képletekben szereplő értékek elhelyezkedését és az arányokban szereplő mennyiségeket a 2-5. ábra mutatja.

A transzformáció képlete:

$$r = pq - \frac{(p - 1)(r_M q_{hm} - q_{hM} r_m)}{(r_M - r_m) - (q_{hM} - q_{hm})}$$



2-5. ábra. Lokális kontrasztkiemelés

Mivel a kiemelés mértéke a környezettől függ, az eljárás alkalmazásával a háttér egyenletesebbé válik. Hátránya, hogy érzékeny a pontszerű hibákra, ezt a bemenő intervallum szélső értékeinek elhagyásával lehet kompenzálni.

2.2.3 KÉPVÁGÁSOK

A képpontok küszöbök szerinti átszinezését (átskálázását) *képvágásnak*, pontosabban $n-1$ *küszöb* megadása esetén n *szintre vágásnak* nevezzük. (A küszöbök meghatározásával kapcsolatban lásd a 4. fejezetet.) Legyenek a megadott küszöbök $\kappa_1, \kappa_2, \dots, \kappa_{n-1}$, és vegyük figyelembe a mindenkor érvényes $\kappa_0 = q_m$, illetve $\kappa_n = q_M + 1$ alapértelmezést. Ekkor az n -szintre vágást az

$$r(k, l) \leftrightarrow \{r_i | \kappa_{i-1} \leq q(k, l) < \kappa_i; i = 1, 2, \dots, n\} \quad (2-6)$$

hozzárendelési utasítással írhatjuk le; ahol r_i az i -edik küszöbhez előre megadott *sajátszinkód*.

Töltsünk fel egy Q -táblát úgy, hogy az $(i-1)$ -edik és az i -edik küszöb közé eső sorszámú elemeibe beírjuk az r_i világosság-, illetve szinkódot. Ekkor a megadott n szintre vágást a (2-5) képlet alapján lehet a legegyszerűbben végrehajtani. Természetesen nem szükséges, hogy mindegyik küszöbhez más szinkódot adjunk meg: nem szomszédosakhoz azonosak is megadhatók.

Igen gyakori eset a *két szintre vágás*, amikor a képpontokat - egy küszöb megadásával - *értékes*, illetve *háttérpontoknak* minősítjük. A kétszintes képet *bindris képnek* is szokás nevezni, mivel képpontjai egy biten ábrázolhatók.

A vágások speciális esete a *sávkivágás*, amihez két küszöbre (κ_1 és κ_2), valamint két szinkódra (r_1 : *háttérpontok*, r_2 : *értékes pontok*) van szükség. Az eljárás során a

$$\kappa_1 \leq q(k, l) < \kappa_2$$

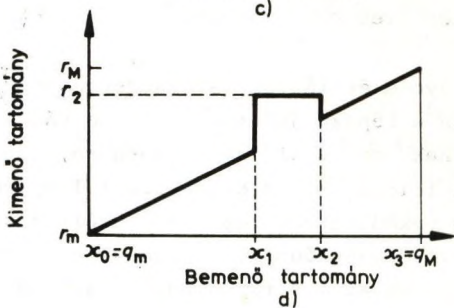
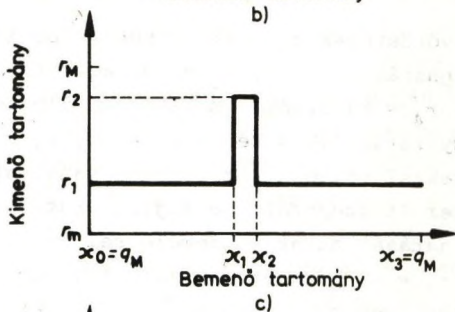
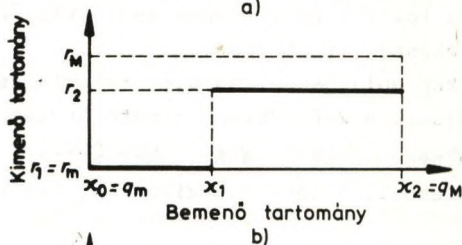
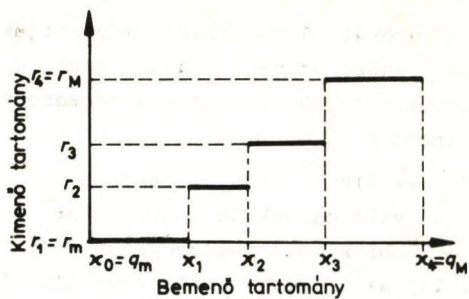
világosságkódu képpontokat változatlanul megtartjuk, a többi pedig háttérszinűre színezzük át. A *sávkivágás kiemeléssel* azt jelenti, hogy egyidejűleg a megtartott képpontokat át is színezzük r_2 színűre.

Sávkiemelésről beszélünk abban az esetben, ha a κ_1 és κ_2 küszöbök közé eső világosságkódu képpontokat r_2 színűre színezzük, de a többi képpont világosságkódját nem változtatjuk meg. Így az eredeti képből egy feltűnő színnel kiemelhetjük a további feldolgozás vagy vizsgálat szempontjából értékesnek ítélt részt.

A vágásokkal kapcsolatban elmondottak szemléltetésére tekintsünk egy példát. A 2-6. ábrán látható átviteli függvények sorrendben négy szintre vágást, két szintre vágást, sávkivágást kiemeléssel, illetve sávkiemelést valósítanak meg.

Az 5/a képen vörösvérsejtek mikroszkópi képe látható. A négy szintre vágásnál $\kappa_1 = 21$, $\kappa_2 = 33$, $\kappa_3 = 49$, $r_1 = 0$, $r_2 = 20$, $r_3 = 40$, $r_4 = 63$ értékek mellett az 5/b képhez jutunk. Tovább egyszerűsödik a kép a $\kappa_1 = 21$, $r_1 = 0$, $r_2 = 63$ paraméterekkel végzett két szintre vágással (lásd 5/c kép): így sikerült szeparálni a sejtmagokat. Az 5/d kép a sávkivágás hatását mutatja kiemeléssel, $\kappa_1 = 18$, $\kappa_2 = 27$, $r_1 = 5$, $r_2 = 50$ paraméterek mellett; míg az 5/e képen a sávkiemelés eredménye látható $\kappa_1 = 18$, $\kappa_2 = 27$, $r_2 = 50$ esetben.

A vágásokkal gyors és látványos eredményeket lehet elérni; ha azonban a kép zajjal torzított, a vágások általában nem vezetnek arra a kívánt eredményre, hogy kiválasszuk a további feldolgozás szempontjából lényeges képrészeket. A képre rakódott zaj ugyanis megváltoztatja egyes képpontok világosságkódját, ezért egy megadott küszöb a zajmentes képen azonos tartományba tartozó képpontokat a zaj miatt különböző osztályokba, illetve eltérő osztályokba tartozókat azonosba sorolhatja. Ilyen esetek-



2-6. ábra. Képvágások: a) négy szintre vágás, b) két szintre vágás, c) sávkivágás kiemeléssel, d) sávkkiemelés

ben a vágás előtt célszerű *zajsűrés*et végezni. (Zajsűrő eljárásokat a 2.3 pontban fogunk ismertetni.)

2.2.4 HISZTOGRAMTRANSZFORMÁCIÓK

Az eddigiektől némileg eltérnek azok az ugyancsak globális módszerek, amelyekben úgy transzformáljuk a világosságkódokat, hogy a kép *hisztogramja előre meghatározott alak* legyen. Erre a következő esetekben lehet szükség:

- *Össze akarunk hasonlítani* két képet, amelyek különböző megvilágítási viszonyok mellett készültek, s így a valóságban azonos fényességű pontokhoz a két képben különböző világosságkódok tartoznak. Ilyenkor az egyik hisztogramot a másikkal azonos alakúra kell hozni.
- A kép olyan tulajdonságait akarjuk elemezni, amelyek a világosságkódok *eloszlásától függenek*. Ekkor a megvilágítás változásai befolyásolják az eredményt, ezért célszerű a hisztogramot előzőleg valamilyen "normalizált" alakra hozni.
- Utólag meg akarjuk valósítani az *eloszlásfüggő* (tapered) kvantálást a kvantálási hiba csökkentése érdekében, rendszerint adott számú kvantumszint megtartásával. E célból süríteni kell a kvantumszinteket azokban a világosságkód-tartományokban, ahova a képpontok nagyobb része tartozik, míg a többi tartományokban a szinttávolságok nőnek.

Az eljárást *közvetlen hisztogramspecifikáció*nak is nevezzük, mivel hatására a kimenőkép hisztogramja előre meghatározott függvény (közelítése) lesz.

Jelölje a bemenő hisztogram ordinátáit $\chi(q)$, a kimenőit $\tau(r)$. A $(2-3/a)$ -ból következik, hogy ezek arányosak a q -adik, illetve az r -edik kvantumszinthez tartozó képpontok számával, és

$$\sum_{q=0}^J \chi(q) = \sum_{r=0}^I \tau(r) = 1, \quad (2-7)$$

ahol a bemenő, illetve kimenő világosságkódok maximális értékét jelöltük J -vel, illetve I -vel. A továbbiakban csak monoton transzformációkkal foglalkozunk, ezért a halmozott részhisztogramoknak minden lépésben meg kell egyezniük:

$$\sum_{q=0}^{m_n} \chi(q) = \sum_{r=0}^n \tau(r); \quad (m_n \leq J; n \leq I). \quad (2-8)$$

A (2-7) és (2-8) feltételek teljesülése mellett kell elérni, hogy a kimenő hisztogram valamilyen, számunkra előnyös, előre meghatározott alakot vegyen fel. A probléma explicit megoldása általában nehéz vagy nem lehetséges, de numerikus iterációval viszonylag könnyű.

Tekintsük a legegyszerűbb esetet, amikor azt kívánjuk elérni, hogy a kimenő hisztogram a vízszintes tengellyel párhuzamos egyenes közelítése legyen, azaz:

$$\tau(r) \approx \frac{1}{I}. \quad (2-9)$$

Látható, hogy ebben az esetben a kimenő hisztogramban az egyes világosságkód-intervallumokba közel egyenlő számú képpont esik, ezért ezt az eljárást *hisztogramkiegyenlítésnek* nevezzük. A (2-9) feltétel teljesülése esetén (2-8) jobb oldala n/I -re egyszerűsödik. A kiegyenlítést úgy végezzük, hogy $n = 1$ -től indulva rendre kiszámítjuk azokat az m_n értékeket, amelyekre

$$\sum_{q=0}^{m_n} \chi(q) \leq \frac{n}{I} < \sum_{q=0}^{m_n+1} \chi(q). \quad (2-10)$$

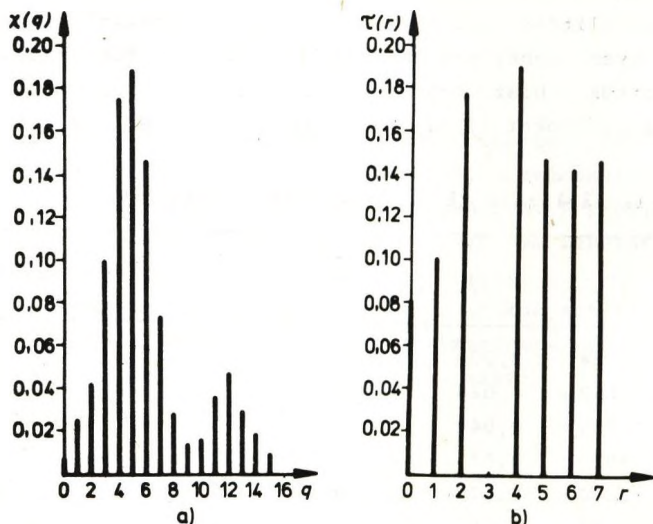
A (2-10) egyenlőtlenség alapján adódó n értékek kijelölik a kimenő hisztogramban a kvantumszinteket, az m_{r_c} index alapján pedig meghatározhatók az adott kvantumszinthez tartozó bemenő világosságkódok.

A fentiek illusztrálására lássunk egy példát. Tegyük fel, hogy egy 128×128 képpontból álló, $J = 16$ szintes képen az egyes világosságkódok a 2-1. táblázat második oszlopában megadott értékekkel fordulnak elő. Végezzünk hisztogramkiegyenlítést úgy, hogy a kimenő kvantumszintek száma $I = 8$ legyen! Ehhez kiszámítottuk, és a 2-1. táblázatban feltüntettük a hisztogramértékeket $[\chi(q)]$, a halmozott részhisztogramokat $[\Sigma\chi(q)]$, valamint az n/I értékeket.

2-1. táblázat: Példa hisztogramkiegyenítésre

q	N_q	$\chi(q)$	$\Sigma\chi(q)$	n/I	r	$\tau(r)$
0	124	0,008	0,008			
1	452	0,028	0,036			
2	725	0,044	0,080	0,125	0	0,080
3	1813	0,111	0,191	0,250	1	0,111
4	2930	0,179	0,370	0,375	2	0,179
5	3123	0,190	0,560	0,625	4	0,190
6	2442	0,149	0,709	0,750	5	0,149
7	1275	0,078	0,787			
8	514	0,031	0,818			
9	223	0,014	0,832			
10	315	0,019	0,851	0,875	6	0,142
11	623	0,038	0,889			
12	835	0,051	0,940			
13	519	0,032	0,972			
14	334	0,020	0,992			
15	137	0,008	1,000	1,000	7	0,149

Ezzel lényegében rendelkezésre állnak a szükséges adatok. A (2-10) egyenlőtlenység felhasználásával meghatároztuk az új kvantumszintek (r) alapján az összevonási határokat (a táblázatban ezeket a megfelelő sorba irtuk). Az összevonásokat elvégezve megkapjuk a táblázat utolsó oszlopában feltüntetett új hisztogramértékeket [$\tau(r)$]. A régi és az új hisztogramot a 2-7. ábra is mutatja.



2-7. ábra. Hisztogramkiegyenlítés

A 6. képen egy LANDSAT űrfelvételén végzett hisztogramkiegyenlítés hatását mutatjuk be. A 6/a eredeti kép meglehetősen kontrasztszegény; a képrészlet hisztogramjáról (6/c kép) viszont leolvasható, hogy a világosságkódok túlnyomó többsége a 0 és 38 érték közé esik. (A vízszintes tengely a [0-255] világosságkód-tartományt jelzi; a maximális hisztogramértéket feleltetjük meg a függőleges tengely hosszának, és ehhez viszonyítva ará-

nyosan ábrázoljuk a többi.) A 6/c képen (az igen kis előfordulási gyakoriság miatt) nem látható, de a hisztogram táblázatos kiíratásából megállapítható volt, hogy lényegesen szélesebb sávban jelentkeznek a világosságkód-értékek.

A hisztogramkiegyenlítés során a kimeneti kvantumszintek számát 64-ben határoztuk meg. A kiegyenlítés után kapott 6/b képen a kontraszt lényegesen javult, az eredeti képen nem látható részletek is előtűntek. A 6/d képen a kiegyenlítés utáni hisztogramot adtuk meg.

Látható, hogy sem a 2-7. ábrán, sem a 6. képen nem valósult meg a hisztogram teljes kiegyenlítése. Az eljárásból következik, hogy akkor várható közel konstans eredményhisztogram, ha a bemenő kép kvantálási szintjeinek száma jóval több, mint amennyit a kimenő képen megengedünk.

Teljes kiegyenlítés általában csak akkor lehetséges, ha a bemenő képen azonos világosságkódu képpontok halmazát felbontjuk, és a kimenő képben különböző világosságkódu halmazokba soroljuk. Önkényesen végrehajtott felbontás természetesen értelmetlen; jó eredményt akkor várhatunk, ha valamilyen módon a lokális tulajdonságokat is fel tudjuk használni.

Bár az előbbieken ismertetett hisztogramkiegyenlítés általában lényegesen javítja a kép kontrasztját, az emberi érzékelő rendszer szempontjából mégsem ad optimális megoldást. Ismeretes ugyanis, hogy az emberi szem nem lineárisan, hanem közel exponenciálisan érzékeli a fénysűrűség-változásokat (lásd 1.1.3.2 alpont). Ezt figyelembe véve olyan transzformációt kell végrehajtani, amely *hiperbolikus* lefutású hisztogramot eredményez [2-8], mert szemünk ezt egyenletes eloszlásúnak fogja érzékelni.

2.3 ZAJELNYOMÁS

A képhibák másik nagy csoportját a zajok alkotják. A képet érő zavaró hatások egy része abban nyilvánul meg, hogy egyes képpontok eredeti világosságkódja - általában véletlenszerűen előforduló izolált pontokban - megváltozik. Ezeknek a pontoknak a világosságkódja lényegesen eltér a környezetétől. A zajokat az érdemi feldolgozás előtt el kell távolítani a képről, de legalábbis csökkenteni kell a hatásukat. Az erre irányuló eljárásokat nevezzük *zajelnyomásnak* vagy *simitásnak*. A továbbiakban néhány jellegzetes eljárást ismertetünk, természetesen itt sem törekedhetünk teljességre.

2.3.1 SIMITÁS A FREKVENCIATARTOMÁNYBAN

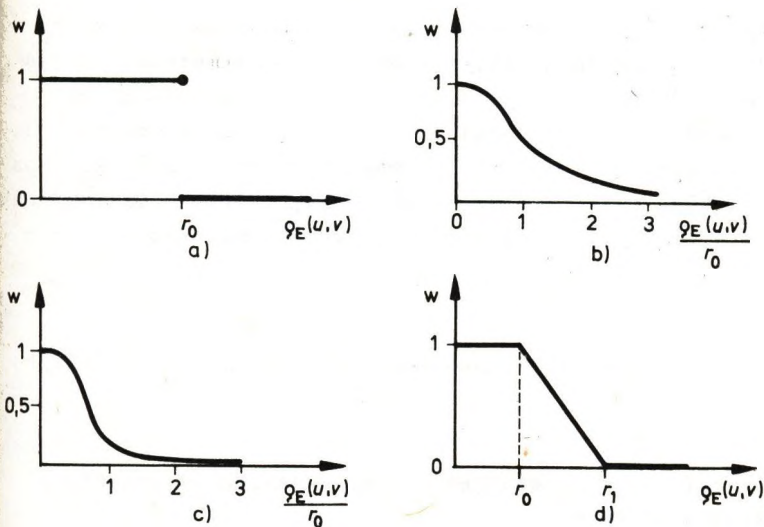
A zajok és a világosságkódokban mutatkozó éles átmenetek a kép Fourier-transzformáltjának magasfrekvenciás összetevőiben jelentkeznek. Ebből következik, hogy a zajok elnyomásának egyik lehetősége a frekvenciatartományban végzett szűrés. Ez azt jelenti, hogy a (2-2) egyenlőségben szereplő $H(u, v)$ függvényt úgy kell megválasztani, hogy a kép $F(u, v)$ Fourier-transzformáltjából kiszűrje a magasfrekvenciás összetevőket, de az alacsonyabb frekvenciájú komponenseket lehetőleg változatlanul engedje át. Az ilyen szűrőket *aluláteresztő szűrőknek* és az eljárást *aluláteresztő szűrésnek* (lowpass filtering) nevezzük. Az alkalmazott szűrőfüggvényt a Fourier-transzformált valós és képzetes összetevőjére is alkalmazni kell.

A legegyszerűbb, "ideális" aluláteresztő szűrő átviteli függvénye a következő:

$$H(u, v) = \begin{cases} 1, & \text{ha } \rho_E(u, v) \leq r_0; \\ 0, & \text{különben;} \end{cases} \quad (2-11)$$

ahol ρ_E az (F3-2)-ben megadott euklidészi távolságfüggvény. A (2-11) definícióból látható, hogy az ideális szűrő változatlanul átengedi az r_0 (>0) sugaru kör belsejébe eső alacsonyfrekvenciás összetevőket, a körön kívüli magasabb frekvenciájúakat pedig teljesen kiszűri. Azokat az (u_0, v_0) frekvenciapárokat, ahol a $H(u, v)$ függvénynek szakadása van (1-ből 0-ba ugrik) *vágási frekvenciáknak* nevezzük (a 2-8/a ábrán $r_0 = \sqrt{u_0^2 + v_0^2}$).

Mind az "ideális", mind a továbbiakban ismertetésre kerülő többi szűrő körszimmetrikus az origóra nézve, azaz a w tengelyre illeszkedő bármely síkban azonos metszetet kapunk. Ha pl. a $H(u, 0)$ függvényt ábrázoljuk az (u, w) síkban (lásd 2-8/a ábra) majd ezt a w tengely körül körbeforgatjuk, akkor megkapjuk a $H(u, v)$ függvény felületét. (Meggjegyezzük, hogy a körszimmetrikus szűrők al-



2-8. ábra. Aluláteresztő szűrők metszetei: a) ideális, b) Butterworth-, c) exponenciális, d) trapézszűrő

kalmazásakor feltételezzük, hogy a Fourier-transzformáció origója a frekvenciasík középpontjában van.)

Az "ideális" szűrő hatását tekintve bizonyos tekintetben korántsem ideális, minthogy a zajokkal együtt kiszűri a magasabb frekvenciatartományba eső *éleket* is, miáltal a kép igen jelentősen *elhomályosodik*. Ezen kívül fellép a "gyűrűsödési" hatás is, ami azt jelenti, hogy visszaállítás után a képen *periodikusan* jelentkező *folttok* tűnnek fel. A fenti problémák kiküszöbölésére különböző fokoza-
tos, "sima" átmenetű szűrőket szoktak alkalmazni.

Ilyen például az *aluláteresztő Butterworth-szűrő*, amely a következő összefüggéssel adható meg:

$$H(u, v) = \frac{1}{1 + (\rho_E/r_0)^{2n}}, \quad (2-12/a)$$

ahol ρ_E mellől elhagytuk az (u, v) argumentumot; n és r_0 pedig változtatható paraméterek. A keresztmetszet alakja a 2-8/b ábrán látható. Miután ebben az esetben nincs éles vágási határ, vágási frekvenciának azt a pontot szokták tekinteni, ahol $H(u, v)$ a maximális értékének bizonyos százalékára esik. (Látható, hogy $H(u, v)$ értéke maximumának 50%-ára esik, ha $\rho_E(u, v) = r_0$.) Gyakran szokták a szűrőt az alábbi módon megadni:

$$H(u, v) = \frac{1}{1 + 0,414(\rho_E/r_0)^{2n}}, \quad (2-12/b)$$

akkor $\rho_E(u, v) = r_0$ esetben $H(u, v)$ a maximális értékének $1/\sqrt{2}$ -szöröse lesz.

A 2-8/b ábrán látható elnyuló "farok" rész hatására kis mértékben érvényesülnek a magasfrekvenciás komponen-

sek is, így kevésbé homályosodik el a kép, és a sima átmenet miatt a gyűrűsödési hatás sem lép fel.

Előszeretettel alkalmazzák az *exponenciális aluláteresztő szűrőt* is, melynek keresztmetszete a 2-8/c ábrán látható, és a következő összefüggéssel adható meg:

$$H(u, v) = \exp[-(\rho_E/r_0)^n]. \quad (2-13/a)$$

A meredekebb lecsengés miatt kissé jobban homályosít, mint a Butterworth-szűrő, de korántsem annyira, mint az "ideális" szűrő, és itt sem lép fel a gyűrűsödési jelenség. Ha $\rho_E(u, v) = r_0$, akkor a $H(u, v) = 1/e$ értéket vesz fel. A $\ln(1/\sqrt{2}) = -0,347$ konstans alkalmazásával (2-13/a) a következő alakú lesz:

$$H(u, v) = \exp[-0,347[\rho_E/r_0]^n], \quad (2-13/b)$$

és ekkor $H(u, v)$ maximális értékének $1/\sqrt{2}$ -szöröse lesz a vágási frekvenciáknál ($\rho_E(u, v) = r_0$ esetén). Példaként a 7. képen az exponenciális szűrő hatását mutatjuk be. Egy ürfelvétel 64×64 -es részletének Fourier-transzformáltjára a (2-13/b) összefüggés szerinti szűrőt alkalmaztuk $r_0 = 5$ és $n = 1$ paraméterekkel. A 7/a az eredeti kép, a 7/b a Fourier-spektrum, míg a 7/c a szűrés utáni inverz Fourier-transzformációval visszaállított eredménykép. Az a) és a c) képet összehasonlítva jól látható a zajszűrő hatás, de a kép elhomályosodása is érzékelhető.

A 7/d képet a 7/c-ből kaptuk, frekvenciatartományban végzett kontrasztfokozás eredményeként. Ez utóbbi eljárásra a 2.4.1 alpontban térünk vissza. (Megjegyezzük, hogy mivel a feldolgozást csak egy kis képrészletre végeztük el, a 7. képet a képernyőn kétszeres nagyításban megjelentetett képről készítettük.)

Átmenetet képez az "ideális" és a "sima" szűrő között a következő összefüggéssel megadható trapéz alakú alulát-eresztő szűrő:

$$H(u, v) = \begin{cases} 1, & \text{ha } \rho_E < r_0; \\ \frac{\rho_E - r_1}{r_0 - r_1}, & \text{ha } r_0 \leq \rho_E \leq r_1; \\ 0, & \text{ha } \rho_E > r_1. \end{cases} \quad (2-14)$$

A szűrő keresztmetszetét a 2-8/d ábrán adtuk meg.

A frekvenciatartományban végzett simitási eljárások áttekintése után, amely elsősorban [2-10] alapján készült, röviden kitérünk a gyakorlati alkalmazásokkal kapcsolatos tapasztalatokra is. $N \times N$ méretű képre az (F4-33)-mal, illetve az (F4-34)-gyel definiált diszkrét Fourier-, illetve inverz Fourier-transzformáció az

$$F(u, v) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f(k, l) \exp[-i2\pi(\frac{ku + lv}{N})] \quad (u, v = 0, 1, \dots, N-1) \quad (2-15)$$

illetve az

$$f(k, l) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp[i2\pi(\frac{ku + lv}{N})] \quad (k, l = 0, 1, \dots, N-1) \quad (2-16)$$

alapösszefüggések felhasználásával N^4 nagyságrendű művelet elvégzését igényli. Ráadásul komplex lebegőpontos műveletekről van szó, ami tetemes futási időt jelent a hagyományos, soros elvű számítógépeken. Sajnálatos módon a

Fourier-transzformáció nemcsak számítás-, hanem főtár-igényes is, ugyanis egyetlen $F(u, v)$ érték kiszámításához a teljes $N \times N$ -es kép minden pontjára szükség van. Ez azt jelenti, hogy ha a teljes képet nem tudjuk a számítás idejére a főtárba tölteni, akkor - a sorozatos lemezműveletek miatt - nagyságrendekkel megnő a számítási idő. Léteznek különböző *gyors Fourier-transzformációs* (FFT) algoritmusok, amelyekkel a műveletszám és a lemezműveletek száma csökkenthető, de még ezek alkalmazásával sem lehet leszorítani a képtartományban végzett és a következő pontban ismertetésre kerülő számítási eljárások szintjére.

A [2-5] algoritmus felhasználásával TPA 1148 számítógépre elkészített programokban (lásd [2-16]), a műveletszám $N^2 \log_2 N^2$ -tel arányos. Ekkor egy 64×64 -es méretű képre kb. két perc időt vesz igénybe a Fourier-transzformáció végrehajtása, a rendszer leterheltségétől függően. A szűrés elhanyagolható idő alatt fut le, de az inverz transzformáció újabb két percét hozzávéve végül is mintegy 4 perc szükséges a szűrt kép előállításához. Mindez annak ellenére, hogy a futási idő csökkentése érdekében a program MACRO-11 nyelven készült, és a lemezműveletek csökkentése céljából az egy felhasználó részére általában engedélyezetten kívül további 32 kbyte-os memóriaterületet is lefoglal.

Az R-11 bázisu képfeldolgozó rendszerünkben [1-48] egy speciális *gyorsprocesszorral* [2-12] lényeges gyorsítást lehetett elérni. A kétdimenziós Fourier-transzformáció ugyanis visszavezethető két egydimenziósra, amelyeket a képmátrix soraira, illetve oszlopokra külön-külön kell végrehajtani; továbbá ezekben a nyolcas lepkealgoritmuson alapuló gyors transzformációt (FFT) valósítottuk meg [2-1]. A speciális processzor alkalmazásával és a lemezműveletek minimalizálására kidolgozott algoritmussal sikerült 50 s alatt előállítani egy 512×512 képpontból álló

kép Fourier-spektrumát. A frekvenciatartományban végzett szűrés és az inverz transzformáció futási idejét is beleszámítva a szűrt kép előállításához kb. 2 percre van szükség.

A frekvenciatartományban végzett szűrés előnye akkor jelentkezik, amikor egy vagy néhány diszkrét frekvencián fellépő zajt kell eltávolítani. Ekkor lehetőség van olyan szűrők alkalmazására, amelyek a frekvenciatartomány megadott pontjában, illetve kör alakú környezetében szűrnek, és a frekvenciatartomány többi részét átengedik. Ezek a szűrők nem az origó körül körszimmetrikusak, hanem a kérdéses (u_0, v_0) pontba történő eltolással keletkeznek.

A Fourier-transzformáció szimmetriája miatt ezeket a pontokat vagy tartományokat szimmetrikus párokként kell megadni, ideális szűrő esetében például az alábbi összefüggéssel:

$$H(u, v) = \begin{cases} 0, & \text{ha } \rho_E(u - u_0, v - v_0) \leq r_0, \text{ ill.} \\ & \rho_E(u + u_0, v + v_0) \leq r_0; \\ 1 & \text{különben.} \end{cases} \quad (2-17)$$

Nyilvánvaló, hogy ha $r_0 = \frac{1}{2}$, a "szűrési tartomány" egyetlen frekvenciára szűkül. Szükség lehet arra is, hogy egy frekvenciasávot szűrjünk ki. "Ideális" szűrő esetén ezt a következő átviteli függvény valósítja meg:

$$H(u, v) = \begin{cases} 1, & \text{ha } \rho_E(u, v) < r_0 - \frac{d}{2}; \\ 0, & \text{ha } r_0 - \frac{d}{2} \leq \rho_E(u, v) \leq r_0 + \frac{d}{2}; \\ 1, & \text{ha } \rho_E(u, v) > r_0 + \frac{d}{2}; \end{cases} \quad (2-18)$$

ahol d a sávszélesség.

2.3.2 SIMITÁS A KÉPTARTOMÁNYBAN

A (2-1) és a (2-2) képletből világos, hogy a frekvenciatartományban végzett, illetve a konvolúciós szűrés között erős összefüggés van, legtöbbször ekvivalensek. Az alternatívák közül gyakorlati szempontok szerint (futásidő, igényelt tárkapacitás minimalizálása) szokás választani. A Fourier-transzformáció megvalósításával kapcsolatban említett nehézségek miatt érthető, hogy képsimitásra az egyik legelterjedtebb módszer a *konvolúciós szűrők* alkalmazása (lásd F4.1.2 alpont). Az eljárást - az ott levezetett (F4-12) összefüggés szerint - az

$$r(k, l) = \sum_{s=-m}^{+m} \sum_{t=-n}^{+n} q(k+s, l+t) T(s, t)$$

képlettel valósítjuk meg; ami azt jelenti, hogy minden képpont világosságkódját helyettesítjük a szűrő által kijelölt környezetében levő világosságkódok valamilyen *súlyozott átlagával*. A szűrő *hatását* az *elemei* határozzák meg. A 2-9. ábrán adott szűrők zajtiszttító, simító hatásuk.

$$T_1 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad T_2 = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad T_3 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

2-9. ábra. Aluláteresztő konvolúciós szűrők

Ezek a szűrők a nagyfrekvenciás zajokat szűrik ki; elemeik pozitívak és egységre normalizáltak. (A szűrőelemek összegével osztást végzünk a konvolúciós összegzés után.) T_1 -et *egyszerű átlagoló szűrőnek* nevezzük. Többnyire

négyzetes alaku szűrőket alkalmazunk, és ritkán használunk 5x5-ösnél nagyobb méretűeket, mivel a szűrőméret növelésével egyre inkább elhomályosodik a kép, és a számítási idő is megnövekszik [2-11].

A homályosító hatás csökkentése érdekében számos eljárást dolgoztak ki. Legegyszerűbb módszer a *feltételes átlagoló* szűrő használata: eszerint csak akkor helyettesítjük egy képpont világosságkódját a környezetének átlagával, ha egy megadott küszöbnél nagyobb mértékben tér el tőle. Képlettel:

$$r(k, l) = \begin{cases} \frac{1}{8} \bar{q}, & \text{ha } |q(k, l) - \bar{q}| > \kappa; \\ q(k, l), & \text{ha } |q(k, l) - \bar{q}| \leq \kappa. \end{cases} \quad (2-19)$$

Itt \bar{q} a $q(k, l)$ képpont (8)-szomszédságát (lásd F3.1.1 alpont) alkotó képpontok világosságkódjából a

$$\bar{q} = \frac{1}{8} \sum_{s=-1}^{+1} \sum_{t=-1}^{+1} q(k+s, l+t) = q(k, l)$$

képlet szerint számított összeg, κ pedig tapasztalati küszöb.

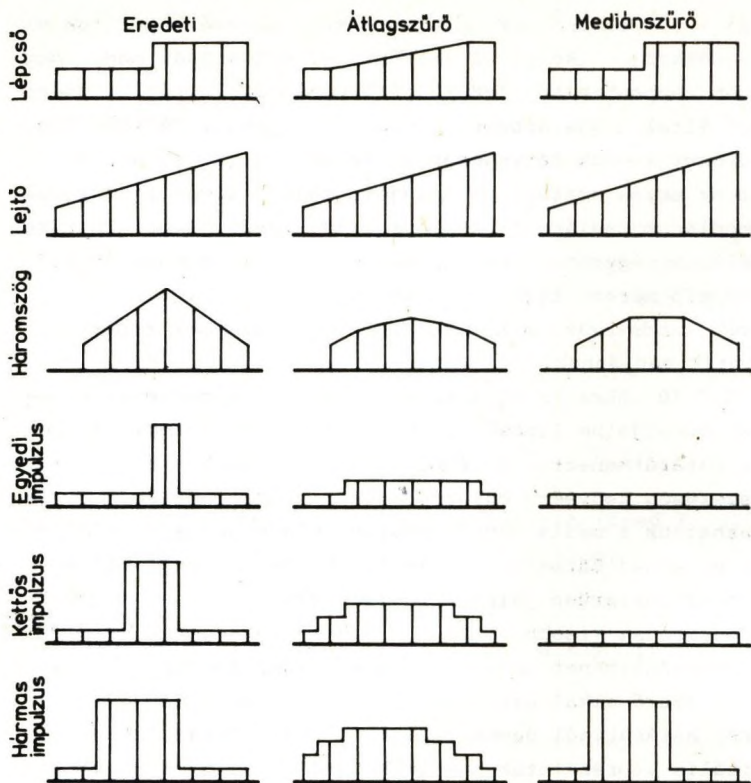
Egy másik módszer ([2-13] és [2-14]) a *lokális statisztikák* felhasználásával kívánja az eredményt javítani; ismét másik olyan algoritmust javasol [2-15], amelyik a *leghomogénabb szomszédságot* keresi meg, és ennek átlagával helyettesíti a vizsgált képpont világosságkódját. Ezek, és egyéb, kedvezőbb eredményt adó eljárások természetesen jelentős többletszámításokat igényelnek.

Kedvező tulajdonságai miatt szívesen alkalmazzák zajelnyomásra a *mediánszűrőt*, amely ugyan lokális jellegű, de nem konvolúciós típusú. (Egy digitális minta mediánján a rendezett minta középső elemét értjük.) A mediánszűrés

során a figyelembe vett mintát a szűrőbe eső képpontok világosságkódja alkotja. A szűrést úgy valósítják meg, hogy minden képpont világosságkódját kicserélik a ráillesztett szűrő által meghatározott minta mediánjával. Jó közelítő eredményt kapunk lényegesen rövidebb idő alatt, ha először az egyes oszlopok mediánját, majd a kiválasztott oszlopmediánok mediánját határozzuk meg [2-2]. A mediánszűrő általában négyzet alakú, de használnak pl. a képaránynak megfelelő méretű téglalapot is. Ha a szűrőelemek száma páros, rendszerint a két középső elem számtani közepét tekintik mediánnak.

A 2-10. ábra az egyszerű átlagoló és a mediánszűrő hatását mutatja be lépcső, lejtő, háromszög és impulzus jellegű határátmenetre, 5 képpontból álló ablak esetén (az egyszerűség kedvéért egydimenziós mintára). Az ábráról leolvashatjuk a mediánszűrő kedvező tulajdonságait: a lépcső- és a lejtőátmenet jellegét nem változtatja meg s az impulzusszerűen jelentkező zajt elnyomja, ha az impulzus periódusa kisebb, mint a szűrő szélességének fele. A háromszögátmenet csucsat viszont kissé letompítja. A 3x3-as szűrő alkalmazásával végzett átlagoló és mediánszűrés hatását jól demonstrálják a 2-11. ábrán látható digitális képrészletek, amelyek a világosságkódok változását mutatják egy y irányú lépcsőátmenet (felső rész) és folt (alsó rész) esetén. Az a) ábrán az eredeti kép, a b) ábrán az egyszerű átlagoló szűrő, a c) ábrán a mediánszűrő hatása látható.

A 8. képen porcszövetről készített mikroszkópi felvételen mutatunk be néhány szűrési eljárást. A 8/a képen a kiegyenlített eredeti felvétellel (9/b kép) kevert "zaj" főleg a bal felső részen és középen alul okoz pontszerű hibákat. A (2-19)-cel megadott feltételes átlagoló szűrő alkalmazása eltünteti a zajok egy részét (ahol a zaj kevésbé üt el a környezetétől), a többi helyen csak



2-10. ábra. Átlagoló- és mediánszűrő hatása egydimenziós mintára

tompítja a zajhatásokat (lásd 8/b kép). A 8/c, illetve a 8/d képen, 3×3 , illetve 5×5 méretű mediánszűrővel végzett szűrés eredményét mutatjuk. Látható, hogy nagyobb szűrőméret esetén (8/d kép) a több pontból álló zajok is eltűnnek; egyszeresmind azonban a kép is homályosabb lesz, mintha a 3×3 -as szűrőt használnánk.

Érdekességként megemlítjük az átlagoló szűrő egy sajátos alkalmazását a háttérmegvilágítás *egyenetlenségének*

	111	111	110	114	108		111	109	111	112	112	110	108	12	113	111	114	114	108			
121	119	121	120	122	118	117	116	116	118	118	116	109	108	119	117	117	117	114	112	10		
120	117	117	117	112	125	96	115	112	110	107	102	94	80	119	117	117	117	112	103	96		
111	99	94	87	76	62	52	91	80	85	77	72	65	59	101	99	94	87	76	62	52		
50	61	46	50	55	29	25	53	55	51	46	43	36	38	51	44	40	36	33	29	25		
25	22	22	22	22	22	21	30	30	28	27	25	24	2	25	25	25	25	25	25	22	2	
20	22	25	25	25	25	22	25	25	22	25	22	22	2	24	25	25	25	25	25	25	22	2
20	20	24	24	20	25	25	25	24	25	25	25	25	2	24	24	24	24	24	24	25	25	2
20	26	25	26	2	30	30	28	29	31	31	35	35	3	20	20	26	26	26	26	30	32	
30	39	40	42	50	50	65	40	40	40	50	55	56	60	30	34	40	42	50	56	65	7	
72	77	78	81	84	85	85	65	60	75	77	76	73	79	72	77	78	81	85	85	85	81	
82	89	102	111	106	95	84	75	85	95	99	98	92	87	77	80	89	102	96	86	87	82	
71	80	101	119	113	96	86	74	85	95	105	104	95	80	71	82	101	104	104	94	87	80	
67	70	86	104	101	91	86	70	76	89	99	100	94	89	70	71	86	101	97	94	88	7	
69	70	76	93	97	94	90	72	77	86	96	100	96	94	7	76	86	97	101	97	91		
	85	92	105	111	111	10	5	91	99	106	110	108	10		85	93	105	111	109	10		

2-11. ábra. Digitális képrészletek: a) eredeti kép, b) átlagszűrő hatása, c) mediánszűrő hatása

(inhomogenitásának) kiegyenlítésére [1-48]. A felhasznált összefüggés

$$r(k, l) = c q_M \frac{q(k, l)}{h(k, l)} \quad (2-20)$$

ahol c tapasztalati állandó, q_M a lehetséges legnagyobb világosságkód, $q(k, l)$ az eredeti, $r(k, l)$ a javítás utáni világosságkódokat jelöli. A $h(k, l)$ "háttér-referenciaképet" többszöri kicsinyítés és 5×5 -ös átlagoló szűrő alkalmazásával, majd a megfelelő számú nagyítás után kapjuk. Az eljárás eredményességét a 9. kép szemlélteti. (A 9/a-n az eredeti, a 9/b-n a javított kép látható; ez utóbbin a háttér csaknem teljesen homogén.)

Bizonyos esetekben nem a képsíkban, hanem időben végzett átlagolással javíthatjuk a kép minőségét. A többszörös képátlagolási eljárás akkor alkalmazható, ha a képvétel során a $q(k, l)$ képre additív zaj rakódik, azaz

$$q_i(k, l) = q(k, l) + z_i(k, l); \quad (2-21)$$

ahol i a képvétel sorszám, és feltételezzük, hogy minden (k, l) koordinátájú pontban a $z_i(k, l)$ zaj független és nulla átlagu. Ebben az esetben a zajos képek $q_i(k, l)$ sorozatát véve és képpontként átlagolva az

$$r(k, l) = \frac{1}{n} \sum_{i=1}^n q_i(k, l) \quad (2-22)$$

képet kapjuk. Az ily módon kapott $r(k, l)$ kép annál jobban megközelíti a zajmentes $q(k, l)$ képet, minél több képre végezzük el az összegzést, mivel feltevésünk szerint a zaj nulla átlagu minden képpontban.

Összehasonlítva a frekvenciatartományban végzett általános célu szűrések hatását a képtartományban végzettekével, általában megállapíthatjuk, hogy ez utóbbiak kisebb ráfordítással adnak azonos vagy esetleg jobb eredményt.

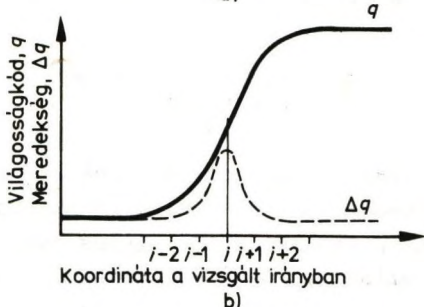
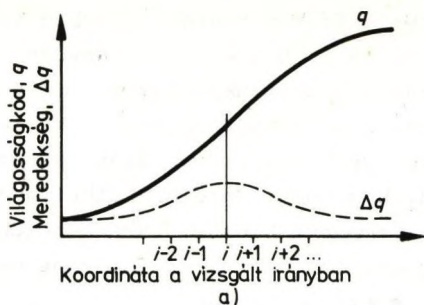
A simító szűrők gyakorlati használatánál azonban ügyelni kell arra, hogy nem megfelelő alkalmazásuk több kárt okozhat, mint amennyi hasznot várunk tőlük. Általában célszerű figyelembe venni a kép jellegét és az eredményt vizuálisan is ellenőrizni.

2.4 ÉLKIEMELÉS

A torzítások gyakran úgy jelentkeznek, hogy a határátmenetek kiszélesednek, az élek elmosódnak. Sok esetben akkor is alkalmazzuk az élkiemelési (élesítési) eljárásokat, ha ilyen torzítás fel sem lép, mivel pszichofizikai kísérletekkel bebizonyították, hogy szubjektíve előnyösebb érzetet kelt a tulhangsúlyozott élekkel rendelkező kép, mint a valóságghü ábrázolás.

Az eljárások célja az egyes képrészletek közötti átmeneti tartomány szűkítése, a határátmenetek hangsúlyozása, meredekebbé tétele, az *elmosódások korrígálása*. Az élkiemelés hatását szemlélteti a 2-12. ábra, ahol egy adott irányban ábrázoltuk az ideális átmeneti függvényt és szaggatott vonallal a megfelelő meredekséggörbét; élkiemelés előtt (a), illetve után (b). (Digitális esetben természetesen csak diszkrét pontokban értelmezettek a függvények, és az értékkészlet a kvantálási szintekhez igazodik.)

Az élkiemelést gyakran alkalmazzák a szegmentálásnál tárgyalt különböző élkitűzési eljárások előkészítő fázisaként is. (Lásd 4.4.1 alpont; ez utóbbiaknál a cél az élpontok geometriai helyzetének pontos meghatározása.)



2.12. ábra. Ideális átmeneti- és meredekségfüggvény:
 a) kiemelés előtt b) kiemelés után

2.4.1 ÉLKIEMELÉS A FREKVENCIATARTOMÁNYBAN

Mint korábban utaltunk rá, az élek (a zajokkal együtt) a kép Fourier-transzformáltjának magasabb frekvenciatartományában helyezkednek el. Így *felüláteresztő szűrők* alkalmazásával, *felüláteresztő szűréssel* (highpass filtering) tudjuk az élesítést elvégezni. Ezek a szűrők - elnevezésüknek megfelelően - kiszűrik az alacsonyfrekvenciás komponenseket és változatlanul átengedik a magas frekvenciatartományba esőket. A 2.3.1 alpontbeli aluláteresztő szűrők megfelelőit könnyen felírhatjuk.

A (2-11) alapján az "ideális" felüldtereszto szürő

$$H(u, v) = \begin{cases} 0, & \text{ha } \rho_E(u, v) \leq r_0; \\ 1, & \text{különben.} \end{cases} \quad (2-23)$$

A 2-8/a ábrán látható metszet megfelelőjét a 2-13/a ábrán tüntettük fel. A felüldtereszto Butterworth-, exponenciális, illetve trapéz alakú szürőt rendre a (2-24), (2-25), illetve (2-26) összefüggés adja meg (ρ_E mellől elhagytuk az (u, v) argumentumot); míg a megfelelő metszetek a 2-13/b, c), illetve d) ábrán láthatók.

$$H(u, v) = \frac{1}{1 + (r_0/\rho_E)^{2n}}; \quad (2-24/a)$$

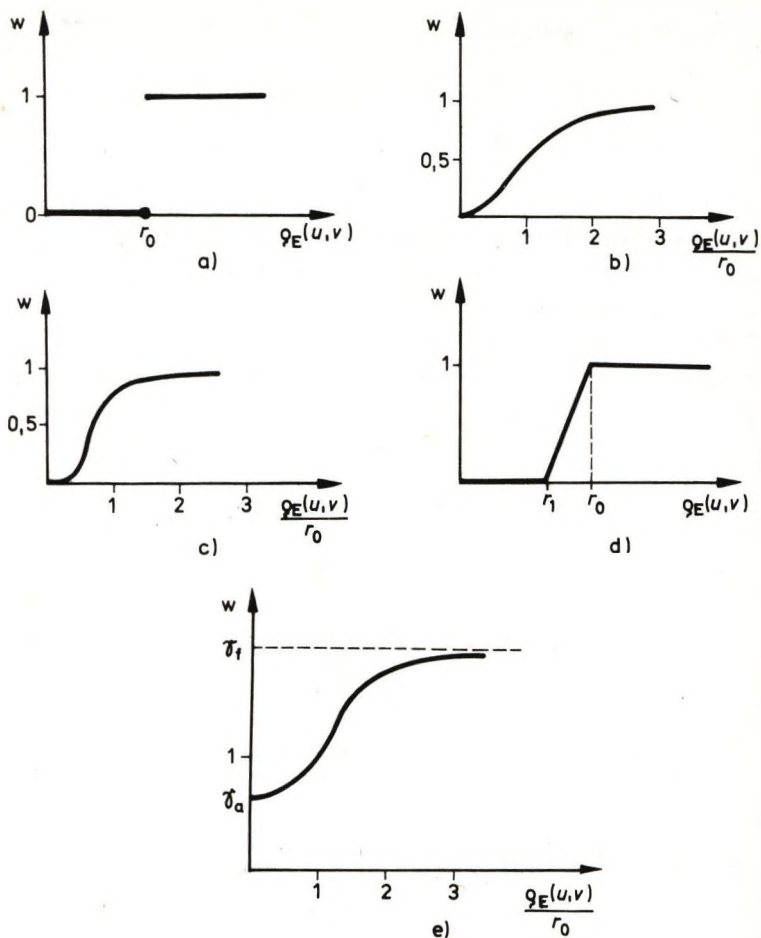
$$H(u, v) = \frac{1}{1 + 0,414(r_0/\rho_E)^{2n}}; \quad (2-24/b)$$

$$H(u, v) = \exp[-(r_0/\rho_E)^n]; \quad (2-25/a)$$

$$H(u, v) = \exp[-0,347(r_0/\rho_E)^n]; \quad (2-25/b)$$

$$H(u, v) = \begin{cases} 0, & \text{ha } \rho_E < r_1; \\ \frac{\rho_E - r_1}{r_0 - r_1}, & \text{ha } r_1 \leq \rho_E \leq r_0; \\ 1, & \text{ha } \rho_E > r_0. \end{cases} \quad (2-26)$$

A (2-18)-ban megadott sávszürőnek megfelelő sávkiemelő függvény:



2-13. ábra. Felüláteresztő szűrők metszetei: a) ideális, b) Butterworth-, c) exponenciális, d) trapéz, e) korrigált (kontrasztfokozó hatású) exponenciális szűrő

$$H(u, v) = \begin{cases} 0, & \text{ha } \rho_E < r_0 - \frac{d}{2}; \\ 1, & \text{ha } r_0 - \frac{d}{2} \leq \rho_E \leq r_0 + \frac{d}{2}; \\ 0, & \text{ha } \rho_E > r_0 + \frac{d}{2}. \end{cases} \quad (2-27)$$

Az aluláteresztő szűrőknél elmondottak értelemszerűen átvihetők a felüláteresztő szűrőkre is, ezeket nem ismételjük meg.

A (2-25/b) exponenciális felüláteresztő szűrő hatását szemlélteti a 10. kép: a 10/a az eredeti kép, a 10/b képen a Fourier-spektruma látható. Az élkiemelést a (2-25/b)-vel megadott szűrővel végeztük, $r_0 = 16$ és $n = 1$ paraméterek mellett; ennek eredményét mutatja a 10/c kép. Az élkiemelés utáni kontrasztfokozás (lásd alább) eredménye a 10/d kép. A felvételeket kétszeres nagyítású képről készítettük, a 7. képhez hasonlóan.

Amint arra már a 2.3.1 alpontban utaltunk, a zajelnyomáson és az élkiemelésen túlmenően a frekvenciatartományban kontrasztjavítás is végezhető. Az eljárás a következő modellen alapul:

A $q(k, l)$ kép kifejezhető az $m(k, l)$ megvilágítás és az egyes objektumok visszaverő képességtől függő $v(k, l)$ függvények szorzataként is:

$$q(k, l) = m(k, l) v(k, l). \quad (2-28)$$

A megvilágítástól függő tényező térben lassan változik, míg a visszaverő képesség hirtelen változhat, különösen a különböző objektumok határán. Durva közelítésben azt mondhatjuk, hogy a Fourier-transzformált alacsonyfrekvenciás összetevői a megvilágítással, a magasfrekvenciás összetevők pedig a visszaverő képességgel hozhatók összefüggésbe. A kép összhatását kedvezően befolyásolhatjuk olyan szűrő

alkalmazásával, amely csökkenti az alacsonyfrekvenciás összetevők hatását és erősíti a magasfrekvenciásakét.

A (2-28) összefüggés azonban nem alkalmazható közvetlenül, miután két függvény szorzatának Fourier-transzformáltja nem szeparálható, azaz nem egyenlő az egyes függvények Fourier-transzformáltjainak szorzatával. Ezért - a Fourier-transzformáció végrehajtása előtt - az eredeti képről a logaritmusára térünk át. Természetesen a szűrés és az inverz Fourier-transzformáció után exponenciális leképezést kell végrehajtani, hogy az eredményképet megkapjuk. A feldolgozás folyamatát a 2-14. ábra szemlélteti.



2-14. ábra. Kontrasztjavítás frekvenciatartományban

Céljainknak megfelelő szűrőt a "sima" átmenetű felüláteresztő szűrőkből nyújtással és a w tengellyel párhuzamos eltolással kaphatunk, például a (2-25/b) összefüggésből:

$$H(u, v) = (\gamma_f - \gamma_a) \exp[-0,347(r_0/\rho_E)^n] + \gamma_a. \quad (2-29)$$

A (2-29) szerinti szűrő keresztmetszetét a 2-13/e ábrán adtuk meg; az alsó, illetve a felső határra a $\gamma_a < 1$, illetve a $\gamma_f > 1$ feltételeknek kell teljesülniük. A $7/d$ és a $10/d$ képen bemutatott kontrasztfokozásnál a (2-29) szűrőt használtuk $r_0 = 10$, $n = 1$, $\gamma_a = 0,75$ és $\gamma_f = 10$ paraméterekkel.

2.4.2 ÉLKIEMELÉS A KÉPTARTOMÁNYBAN

Mint a 2.3.2. alpontban láttuk, a zajelnyomás egyik lehetséges módja a képpontok környezetében végzett átlagolás; ami egyuttal a képet homályosabbá teszi. Az átlagképzés az integrálás diszkrét megfelelője; kézenfekvő tehát, hogy a kép élesítését *differenciálással* próbáljuk elérni.

A képfeldolgozásban erre a célra leggyakrabban a *gradiens*, illetve a *Laplace-operátort* (lásd F4.2.1 alpont) használják. A differenciálhányadost a digitális technikában *differenciahányadosokkal* szokás közelíteni, ez azonban nem mindig vezet egyértelmű eredményre. A továbbiakban az F4.2.2. alpontban levezetett értelmezést használjuk.

A leggyakrabban alkalmazott *élkiemelő függvény* a

$$q(k, l) - \nabla^2 q(k, l) \quad (2-30)$$

kifejezés, ahol ∇^2 az (F4-19)-ben, illetve az (F4-20)-ban definiált Laplace-operátor, ami a másodrendű parciális, illetve totális differenciálnak felel meg. Kiszámítására - diszkrét esetben - az (F4-24)-ben, illetve (F4-25)-ben levezetett

$$\begin{aligned} \nabla^2 q(k, l) = & q(k + 1, l) + q(k, l + 1) + q(k - 1, l) + \\ & + q(k, l - 1) - 4q(k, l); \end{aligned} \quad (2-31)$$

illetve

$$\begin{aligned} \nabla_{\frac{t}{2}}^2 q(k, l) = & q(k + 1, l) + q(k + 1, l + 1) + q(k, l + 1) + \\ & + q(k - 1, l + 1) + q(k - 1, l) + q(k - 1, l - 1) + \\ & + q(k, l - 1) + q(k + 1, l - 1) - 8q(k, l) \end{aligned} \quad (2-32)$$

kifejezést fogjuk használni. Ezeket, valamint egy további kedvelt élkiemelő szűrőt a 2-15. ábrán látható $L_1 - L_3$ konvolúciós szűrőkkel lehet realizálni (vö. még F4-6/a-c ábra). Figyeljük meg, hogy ez esetben a szűrőelemek összege (a szűrő súlya) mindig 1.

$$L_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad L_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad L_3 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

2-15. ábra. Élkiemelő konvolúciós szűrők

A Laplace-típusú élkiemelő szűrők hatását a 11. képen mutatjuk be. Az a) az eredeti kép, a b), c), illetve a d) képen rendre a fenti L_1 , L_2 , illetve L_3 szűrő hatása látható.

Érdekes *pseudoplasztikus* hatást érhetünk el a különböző irányú differenciálhányadosokat közelítő szűrőkből a (2-30) mintájára leszámaztatott élkiemelő szűrőkkel. A 2-15. ábrán pl. az R jelű szűrőt az ÉK irányú "iránytű" (Robinson) szűrőből (lásd 4.4.1.2. alpont) nyertük. Ezt a 12/a képen megadott röntgenfelvételtre alkalmazva a 12/b képen látható eredményt kaptuk. Itt az erek plasztikusan, domborműszerűen kiemelkednek, ami elősegíti a kép értelmezését. (Felhívjuk a figyelmet arra, hogy a kiemelkedések semmilyen összefüggésben sincsenek a valódi térbeli viszonyokkal!)

Definiálhatók nagyobb méretű és más alakú szűrők is, a szűrőméret növekedésének a számítási igény szab határt.

Az élkiemelő konvolúciós szűrőket felüláteresztő szü-

rőknek is nevezzük, mivel ugyanaz a hatásuk, mint a frekvenciatartományban végzett felüláteresztő szűréseknek.

Megemlítjük még, hogy az él irányának kijelölésére alkalmas konvolúciós szűrőket is használnak élkiemelésre. Mint látni fogjuk, az él irányának kijelölése az élkitűzési eljárásoknál fontos szerepet kap, ezért ezekre a szűrőkre később még visszatérünk.

Élkiemelést lehet megvalósítani a *statisztikus differenciaoperátorral* is. Az eljárás során egy képpont korrigált világosságkódját úgy számítjuk ki, hogy elosztjuk a környezetében levő képpontok világosságkódjának eloszlása alapján meghatározott szórással. Az eredményképben a hirtelen változások helyén kis, a homogén tartományokban nagy számértékek jelennek meg. Hátrányos, hogy az eljárás a zajokat is kiemeli. Az esetleges tulcsordulások elkerülésére és egyéb hiányosságok kiküszöbölésére több módosított változatot dolgoztak ki.

A 2-12. ábrán szemléltetett élkiemelő hatás elérésére az SZKI-ban egy más elvű eljárást is megvalósítottunk [2-17], amely a mediánszűrő "felüláteresztő" változatának tekinthető. Eszerint minden képpontra (kivéve a szélsőket a kép szélein) $j = m \times n$ -es ablakot illesztünk, és az ebbe eső képpontok világosságkódját növekvő sorba rendezzük. Jelöljük ezeket q_1, q_2, \dots, q_j -vel, ekkor a (k, l) koordinátájú képpont új világosságkódját az

$$r(k, l) = \begin{cases} q_1, & \text{ha } |q_1 - q(k, l)| < |q_j - q(k, l)|; \\ q_j, & \text{ha } |q_1 - q(k, l)| > |q_j - q(k, l)|; \\ q(k, l), & \text{ha } |q_1 - q(k, l)| = |q_j - q(k, l)| \end{cases} \quad (2-33)$$

összefüggésből határozzuk meg. Az eljárással bizonyos mértékű *zajszűrést* is el lehet érni, ha a (2-33)-ban nem a sorba rendezett világosságkódok legszélső értékeit vesszük

figyelembe, hanem az ezektől jobbra, illetve balra eső q_{1+i} , illetve q_{j-i} -edikkel számolunk [$i < \text{int}(m \times n / 2)$]. Esetünkben i a felhasználó által megadható tapasztalati paraméter.

Végezetül megemlítjük, hogy érdekes módon aluláteresztő konvolúciós, azaz *homályosító szűrők* segítségével is lehet élkkiemelést elérni. Az eljárás azon alapul, hogy megfelelő súlyozó tényezőkkel figyelembe vett finomabb és durvább felbontású kép különbségét képezve az éllátmeneteknél tullövés keletkezik, ami az élek hangsúlyozását eredményezheti. A durvább felbontású képet például átlagoló szűrővel nyerhetjük. Az eljárást továbbfejlesztve [2-9]-ben - *különböző méretű homályosító szűrőkkel* kapott képek összegzésével - az emberi konturérzékeléshez közel álló élkkiemelő algoritmust fejlesztettek ki.

2.5 TÖBBSÁVOS KÉPEK JAVITÁSA

A képfeldolgozás alkalmazási területei között az úrfelvételek kiértékelése igen jelentős helyet foglal el. Ezen a területen - mint arra röviden már kitértünk az 1.3.4 alpontban - kezdettől fogva használnak többsávós (*multi-spektrális*) képeket.

A többsávós kép ugyanarról a területről (jelenetről) egyidejűleg több frekvenciasávban készített felvételtől áll. Ezt a

$$\underline{q}^T(k, l) = \{q_1(k, l), q_2(k, l), \dots, q_n(k, l)\} \quad (2-34)$$

szimbólummal jelöljük; $q(k, l)$ -et $n \times 1$ elemű oszlopvektorként kezeljük. Lényegében minden (k, l) koordinátájú ponthoz egy n dimenziós vektort rendelünk hozzá, a vektor i -edik

komponense $q_i(k, l)$ értékének felel meg. A $q_i(k, l)$ függvény-nyel leírt digitális képet i -edik színösszetevőnek vagy sávnak nevezzük.

Bár a "többsávos kép" elnevezést eredetileg az ürfelvételekre használták, mi a továbbiakban többsávosnak nevezünk minden képet, amelynek egynél több színösszetevője van, így pl. a szineskamerával vagy szinszűrőkkel felvett 3 színösszetevős képeket is.

Az előző pontokban ismertetett képjavítási eljárások természetesen alkalmazhatók a többsávos képek egyes színösszetevőire külön-külön is, de számos speciális eljárás is született, amelyekben kihasználják a több sáv által nyújtott lehetőségeket.

2.5.1 SZINKORREKCIÓK

Napjainkban a színek alkalmazása általánossá vált a képfeldolgozó rendszerekben, ami azzal magyarázható, hogy a szemünk mintegy 4000 színt tud megkülönböztetni, míg világosságban - így fekete-fehér képek esetében "szürkességben" is - csak kb. 500 fokozatot képes érzékelni (lásd 1.1.3.3 alpont).

Az egyik gyakran alkalmazott szinkorrekciónak eljárás a *hamis színezési* (false colour) technika. Az eljárás neve is utal arra, hogy alkalmazásával az eredményképen az egyes objektumok az eredetitől eltérő színekben jelennek meg. A hamisszínes kép előállításának célja lehet például a megfigyelő figyelmének felkeltése a szokásostól eltérő, meghökkentő színezéssel, vagy az emberi szem színérzékenységének jobban megfelelő kép előállítása. Ismeretes például, hogy az emberi szem a látható spektrum zöld tartományában a legérzékenyebb a fényességi változásokra, így például az eredetileg más színű objektumok zöldes árnyala-

tokban történő megjelenítéssel jobban érzékelhetővé válnak.

Általános megfogalmazásban a hamis színezés egy pont-pont leképezés, ahol a multispektrális kép sávjait vagy színes kép esetén az eredeti színösszetevőket a három alapszínnel megadott szintérbe képezzük le. Képlettel:

$$\begin{aligned} r_R(k, l) &= T_R\{q_1(k, l), q_2(k, l), \dots, q_n(k, l)\}, \\ r_G(k, l) &= T_G\{q_1(k, l), q_2(k, l), \dots, q_n(k, l)\}, \\ r_B(k, l) &= T_B\{q_1(k, l), q_2(k, l), \dots, q_n(k, l)\}; \end{aligned} \quad (2-35)$$

ahol T_R , T_G , T_B a piros, zöld, illetve kék színhozzárendelést leíró transzformáció; $r_R(k, l)$, $r_G(k, l)$, $r_B(k, l)$ pedig a piros, zöld, illetve kék folytonos színjelek előállításának alapjául szolgáló digitális színösszetevők.

Természetes színes képek lineáris leképezése esetén a transzformációt 3×3 elemű mátrix definiálja. A q_i színösszetevőknél a megfelelő betűindex alkalmazásával (2-35) a következő módon egyszerűsödik:

$$\begin{bmatrix} r_R \\ r_G \\ r_B \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} q_R \\ q_G \\ q_B \end{bmatrix} \quad (2-36)$$

Egy egyszerű zöld-piros, kék-zöld, piros-kék szincseré az alábbi mátrixszal valósítható meg:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Az ürfelvételek több frekvenciasávban készülnek, de miután sok szempontból (pl. hőmérséklet, vegetáció megfigyelése) a nem látható spektrumok előnyösebbek, ezért a látható tartományba csak egy vagy két sáv esik. A LANDSAT erőforráskutató műhold MSS rendszere [1-14] például a következő hullámhossztartományokban dolgozik:

MSS4	500-600 nm	zöld;
MSS5	600-700 nm	vörös;
MSS6	700-800 nm	infravörös 1;
MSS7	800-1100 nm	infravörös 2.

Miután a kék szín hiányzik, valódiszines képet nem lehet létrehozni, de a rendelkezésre álló sávokból a legkülönbözőbb hamisszines, vagy - a távérzékelési terminológia szerint - *kompozit* kép állítható elő. Legegyszerűbb esetben úgy készül a hamisszines kép, hogy 3 sávot kiválasztunk és megfeleltetjük a 3 alapszinjelnek, általában azonban a kiválasztott sávokra alkalmas transzformációt is végre szoktunk hajtani (lásd [2-4]).

Válasszuk ki például Komárom megye egy részletét lefedő LANDSAT felvétel MSS4, MSS5 és MSS7 összetevőjét, és végezzük el az alábbi transzformációt:

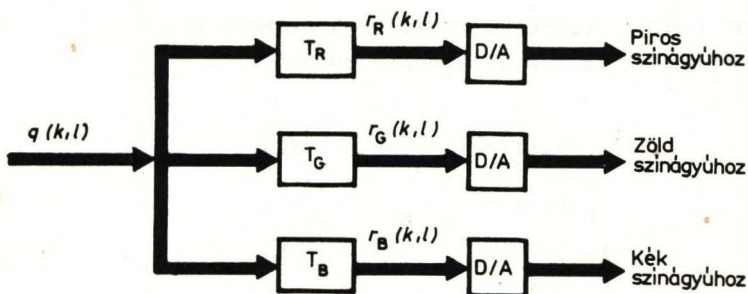
$$\begin{aligned}
 r_R(k, l) &= T_R \{q_{MSS7}(k, l)\}, \\
 r_G(k, l) &= T_G \{q_{MSS5}(k, l)\}, \\
 r_B(k, l) &= T_B \{q_{MSS4}(k, l)\},
 \end{aligned}
 \tag{2-37}$$

ahol T_R , T_G , T_B transzformáció olyan, a 2-7. ábrán bemutatott hisztogramkiegyenlítést jelöl, amely - az alkalmazott képmű (lásd 1.3.3 alpont) gradiációs felbontásának

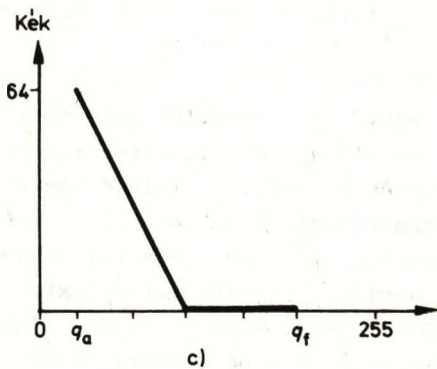
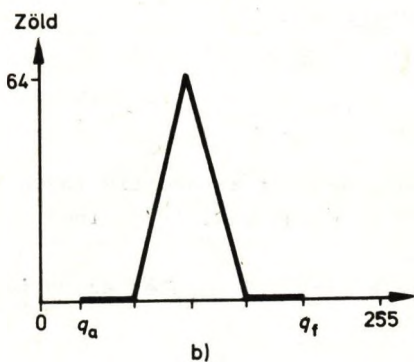
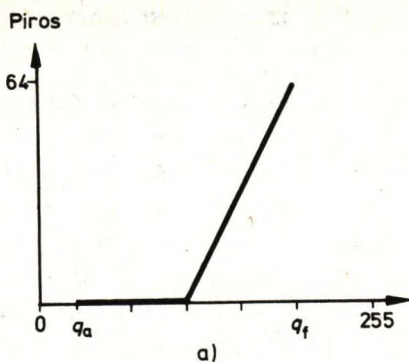
megfelelően - a lehetséges 64 világosságkódra képezi le az eredeti értékeket. A három különböző képsíkba elhelyezett r_R , r_G , r_B digitális képből előállított analóg videojeleket a színes monitor R, G, B bemenetére kapcsolva kapjuk az eredményképet.

A teljesség kedvéért mind a négy sávot bemutatjuk hisztogramkiegyenlítés után fekete-fehér képen (lásd 13. kép); a továbbiakban ezekre *mintaképként* fogunk hivatkozni. A színes kompozit kép a II/a képen látható. Jól megfigyelhető, hogy az élő vegetáció a képen - az infravörös összetevő hatására - élénk vörös színben jelentkezik.

Az *álszínes* (pseudo colour) kép előállítása alapvetően abban tér el a hamisszínesétől, hogy egyszávos, azaz monokromatikus képből indulunk ki. A három alapszint előállító transzformációkat ugyanarra a képre alkalmazva kapjuk a három színösszetevőt, amint a 2-16. ábra szemlélteti. A II/b képen egy álszínes képet mutatunk be, amely a 2-17. ábrán látható transzformációk alapján állítja elő az eredeti képből (MSS7) az egyes színösszetevőket.



2-16. ábra. Álszínes kép előállításának elve



2-17. ábra. Álszines képet előállító piros (a), zöld (b), illetve kék (c) transzformáció

2.5.2 KÉPSÁVOK KÖZÖTTI MŰVELETEK

Igy nevezzük azokat az eljárásokat, amelyekben két vagy több képsáv azonos geometriai koordinátájú képpontjaihoz tartozó világosságkódokból állítjuk elő az eredményképet. Természetesen nemcsak az azonos időpontban, különböző frekvenciatartományokban készített (*multi-spektrális*) képsávok, hanem a különböző időpontokban, azonos területről, azonos frekvenciatartományban készített (*multitemporális*) felvételek között is végezhetők ezek a műveletek.

Gyakran végeznek kivonást az

$$r_{mn}(k, l) = q_m(k, l) - q_n(k, l) \quad (2-38)$$

összefüggés szerint az additív zajok vagy a leképező rendszer helytől függő additív jellegű torzításainak kiszűrésére.

A másik egyszerű művelet az osztás az

$$r_{mn}(k, l) = \frac{q_m(k, l)}{q_n(k, l)}, \quad q_n(k, l) \neq 0 \quad (2-39)$$

képlet szerint, a megvilágítási hibákból eredő világosságkód-torzítások kiegyenlítésére. (A $q(k, l) = 0$ előfordulása esetén előzőleg megfelelő átskálázást kell végezni.)

A hányadosképzés azon a 2.4.1 pontban már említett modellen alapszik, hogy egy színösszetevő világosságkódjai a megvilágított felület objektív $v(k, l)$ visszaverő képességének (amely képpontonként változik) és a képpont $m(k, l)$ -lel jellemezhető megvilágításerősségének szorzataként állnak elő, ahol az $m(k, l)$ ugyan képpontonként változhat, de egy képpontra közel azonos minden frekvencia-

sávban. Ily módon ez eredményképből ki lehet szűrni az egyenetlen háttérmegevilágítás torzító hatását.

Az osztás közvetlen elvégzése azzal a hátránnyal jár, hogy a kvantálási hiba jelentősen befolyásolja az eredményt, ha a nevezőben lévő $q_n(k, l)$ értékek kicsik. Ennek elkerülésére (2-39) logaritmusát szokták képezni. A logaritmusra való áttérést indokolja az is, hogy mivel véges sok világosságkód-érték létezik, ezek logaritmusai előre kiszámítható és letárolható, így az időigényes osztás helyett a lényegesen gyorsabb táblázatból való keresésre és kivonásra redukálható a művelet elvégzése.

A képsávoknak különböző súlyozó konstansokkal és normalizáló tényezőkkel képzett lineáris kombinációit és hányadosait is lehet definiálni.

A multispektrális képek esetében az SZKI-ban kifejlesztett rendszer [2-3] is tág lehetőségeket kínál a képsávok közötti műveletvégzésre, megengedve a műveletvégzés előtt vagy után különféle világosságkód-transzformációk végrehajtását. Felhívjuk azonban a figyelmet arra, hogy - a lehetséges változatok nagy száma miatt - véletlenszerű alkalmazásuk értelmetlen. Kielégítő eredményeket csak tudatos tervezéssel, a vizsgált objektumok és az értékelőrendszer paramétereinek ismeretében kialakított formulákkal lehet elérni (lásd pl. [A-9], [2-6]).

Végezetül megjegyezzük még, hogy a többsávós képfeldolgozás alapvető problémája a feldolgozandó adatok nagy tömege (egyes rendszerekben már 7-13 sávban készítenek felvételeket). Ezek egyidejű feldolgozása még a ma rendelkezésre álló nagykapacitású számítógépek segítségével is nehézkes. E probléma feloldása úgy lehetséges, hogy egy-egy konkrét feladatra csak 1-2 színösszetevőt használnak. Ezek megfelelő kiválasztásán tulmenően célszerű kiszűrni (és egy leszármaztatott képben tömöríteni) a számításigényes feldolgozások előtt a különböző színössz-

szetevőkben levő hasznos információkat. Ennek szokásos megoldása a *főkomponens-transzformáció* (amit az 5. fejezetben tárgyalunk), azonban gyakran eredményesen alkalmazhatók a szinösszetevők közötti egyszerű alpműveletek is.

3. GEOMETRIAI KORREKCIÓK

A geometriai korrekció a leggyakoribb kép-helyreállítási eljárás. (Ezek célja, mint említettük, a leképezési hibáktól mentes, "ideális" kép létrehozása.) Alkalmazására rendszerint akkor kerül sor, ha a képen helyre akarjuk állítani a *valóságos méretarányokat*. Erre akkor van szükség, ha a képről valódi méreteket akarunk leolvasni, vagy egy tárgy különböző képeit akarjuk összehasonlítani.

A geometriai torzítások három fő oka a következő:

- a képfelvevő rendszer optikailag hibás;
- a háromdimenziós világ kétdimenziós leképezése pontatlan (perspektív torzítás);
- felvétel közben változik a geometriai összefüggés a felvevő eszköz és az objektum között (elmozdulás).

Ellentétben a képjavítási módszerekkel, a geometriai korrekció során *nem* vagyunk tekintettel a torzulás eredetére, ezért a torzítások okaival részletesen nem foglalkozunk. A korrekció általános esetben egy (vagy több, geometriailag összetartozó) *bemenő kép* áttranszformálását jelenti *kimenő kép(ek)*re úgy, hogy a képi információ a megengedettnél nagyobb mértékben *ne* sérüljön, miközben a kívánt geometriai összefüggés *legalább* a megadott mértékben teljesül.

A geometriai korrekciók az esetek többségében globálisan (vagyis a teljes képre vonatkozóan) *nemlineárisak*; ezért, valamint a nagy adatmennyiségek miatt is, jelentős a számításigényük. Megvalósításuk kulcskérdése - a már

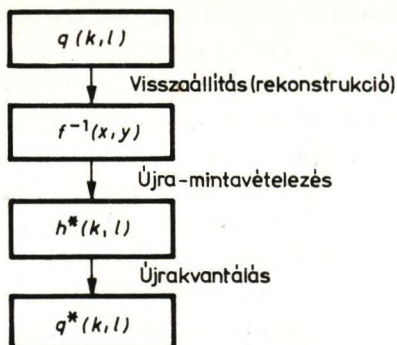
emlitett pontossági követelmény mellett - a végrehajtási sebesség. Szerencsére bizonyos inherens összefüggések felhasználásával (pl.: invertálható korrekciós összefüggések alkalmazása esetén a szomszédos pontok szomszédosak maradnak, lineáris korrekció esetén egyenes képe egyenes stb.) hatékony algoritmusokat lehet kidolgozni. Így egyre nő a jelentőségük, és szélesedik az alkalmazási területük.

Az SZKI-ban folyó képfeldolgozó-rendszerfejlesztési munkák során sokféle geometriai korrekciós feladatot is meg kellett oldanunk (pl.: [3-20], [3-21], [3-22]). Az elméleti és gyakorlati kutatómunka eredményeit az első hazai erőforrás-kutatási célú képfeldolgozó rendszerben [1-31] hasznosítottuk (részletesebben lásd 3.5.2 alpont). Ennek segítségével készült a III. kép, amelyen egy ürfelvétel (a) geometriai korrekcióját (b) mutatjuk be. (A korrigált kép méretarányos.) Bár folytattunk kísérleteket a harmadik dimenzió (magasság) ábrázolására is [3-9] (pl. a IV. képen *ál-háromdimenziós kép* látható), vizsgálatainkat alapvetően a gyakorlatban legnagyobb jelentőségű problémakörre, a *kétdimenziós képek invertálható geometriai korrekcióira* korlátoztuk.

Könyvünk célkitűzéseinek megfelelően a továbbiakban - rövid áttekintés után - elsősorban az egyes részproblémák hatékony megoldását mutatjuk be. Sajnos, a korlátozott terjedelem miatt a javasolt algoritmusok többségét csak vázlatosan tudjuk ismertetni, s nem kerülhetett sor konkrét korrekciós feladat megoldásának ismertetésére sem.

3.1 ELVI ÁTTEKINTÉS

A digitális képalkotás modellje az 1-14. ábrán látható; ennek mintájára a 3-1. ábrán adjuk meg a geometriai korrekciót.



3-1. ábra. A geometriai korrekció modellje

A korrekció elvégzése érdekében először vissza kell állítani az analóg képet a bemenő képnek tekintett digitális képből. A visszaállítással nyert $f^{-1}(x, y)$ folytonos képtől elvárjuk, hogy a digitalizálás előtti, eredeti $f(x, y)$ képet a lehető legjobban megközelítse, az 1.2.3 alponthban leírtak értelmében. (Emlékeztetünk rá, hogy a digitalizálási hibák miatt a visszaállított kép tökéletesen sohasem lehet azonos az eredetivel.) Ezzel a fogással a digitális kép geometriai korrekcióját visszavezettük a folytonos kép újbóli mintavételezésére $[h^*(k, l)]$ és új kvantálásra $[q^*(k, l)]$. Ez utóbbi, geometriailag korrigált képnek (továbbiakban: kimenő kép) lehetőleg ugyanazt a képi információt kell tartalmaznia, mint a korrigálatlan $q(k, l)$ képnek, a kívánt geometriai összefüggés teljesítése mellett.

A korrekció geometriai jellegét az eredeti digitalizálás során alkalmazott, illetve az újbóli mintavételezés *különbsége* adja. Az újbóli mintavételezés rendjét a *korrekciós összefüggés* szabja meg.

Már a digitalizálási folyamat leírásakor (lásd 1.2 pont) említettük, hogy az egyes lépések nem különböznek el

élesen egymástól. Ez most olyannyira igaz, hogy a modellben szereplő visszaállítás csak az új mintavételezési pontokban valósul meg, az újra-mintavételezés pedig egyúttal jár az újbóli kvantálással, s így rögtön az új világosságkódokat kapjuk meg.

A modell mégis jól szemlélteti a geometriai korrekció lényegét. Látható belőle, hogy a korrekció két lényegi részfeladatra bontható:

- *koordinátatranszformáció* (az újra-mintavételezés rendjének meghatározásához), illetve
- *visszaállítás és ujradecimalizálás* (a világosságkód meghatározásához a megadott pozícióban).

Az első témakörrel a következő pontban foglalkozunk részletesebben; a másodikban belül csak az újra-mintavételezés igényel érdemi tárgyalást, ezt a 3.3.8 alpontban tesszük meg.

3.2 KOORDINÁTATRANSZFORMÁCIÓK

Tegyük fel, hogy a digitális képalkotásnál alkalmazott mintavételezés a valós, háromdimenziós világ valamely téreleméhez a $P(x, y)$ képpontot rendelte hozzá. Ha a kapott kép geometriai korrekciójával azt akarjuk elérni, hogy ugyanennek a térelemnek a korrigált képen a $P'(x', y')$ képpont feleljen meg, akkor szükség van az

$$(x, y) \rightarrow (x', y') \quad (3-1)$$

koordinátatranszformációra. A geometriai korrekció végrehajtásakor így határozhatjuk meg, hogy a P' képpont elhelyezéséhez melyik P pontjának környezetében kell visszaállítani a korrigálandó digitális képből az eredeti (többszörösen felbontott) analóg képet, majd ezt ujradecimalizálni.

3.2.1 LINEÁRIS KOORDINÁTATRANSZFORMÁCIÓK

Tekintsük először a lineáris koordinátatranszformációkat. Legyen V és W két vektortér a valós számtest (\mathbb{R}) felett. Az

$$L: V \rightarrow W$$

leképezést akkor nevezzük *lineáris transzformációnak*, ha tetszőleges $\underline{x}_1, \underline{x}_2 \in V$ és $\lambda \in \mathbb{R}$ esetén teljesül az

$$L(\underline{x}_1 + \underline{x}_2) = L(\underline{x}_1) + L(\underline{x}_2),$$

és az

$$L(\lambda \underline{x}_1) = \lambda L(\underline{x}_1)$$

azonosság. Adott bázis (lineárisan független vektorrendszer) esetén az n -dimenziós vektortér lineáris transzformációi és az $n \times n$ -es mátrixok közt egy-egy értelmű megfeleltetés létesíthető. A képfeldolgozásban fontos kétdimenziós esetben a lineáris transzformációkat 2×2 -es mátrixok alkalmazásával lehet megvalósítani.

A $P(x, y)$ és $P(x', y')$ pontok közti lineáris transzformáció

$$[x \ y]_{\underline{A}} = [x' \ y']; \quad \underline{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (3-2)$$

alakban írható fel, ami a mátrixszorzás szabályai szerint (lásd F1.2.1 alpont) megfelel az

$$x' = a_{11}x + a_{21}y,$$

$$y' = a_{12}x + a_{22}y$$

összefüggéseknek.

3.2.1.1 Elemi koordinátatranszformációk

Foglaljuk össze röviden a lineáris transzformációk speciális típusait, az *elemi koordinátatranszformációkat*, a (3-2) összefüggés alapján:

1. Ha $a_{21} = a_{12} = 0$, akkor *léptékváltásról* beszélünk, mégpedig

$|a_{11}| > 1$, illetve $|a_{22}| > 1$ esetén *nagyításról*,

$0 < |a_{11}| < 1$, illetve $0 < |a_{22}| < 1$ esetén *kicsinyítésről* van szó.

Ha $|a_{11}| < 0$, illetve $|a_{22}| < 0$, akkor a léptékváltás egyszersmind *tükrözést* is jelent az y , illetve az x tengelyre nézve.

Ha $|a_{11}| = 1$, illetve $|a_{22}| = 1$, akkor a léptékváltás *mérettartó*; ha $a_{11} = 0$, illetve $a_{22} = 0$, akkor *eltüntető*. Az utóbbi eset a *projekció*, amely a kép elfajulását eredményezi.

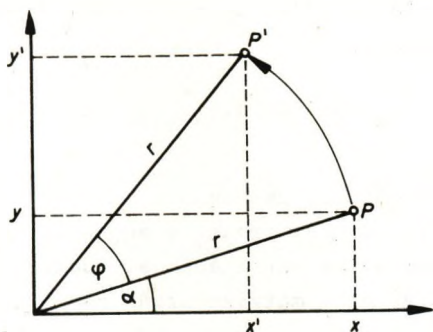
A két koordináta transzformációja egymástól független.

2. Ha $a_{11} = a_{22} = 1$, de $a_{12} \neq 0$, illetve $a_{21} \neq 0$, akkor a transzformációt *nyírásnak* nevezzük.

3. Az

$$\underline{A} = \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix}$$

mátrix φ szögű *elforgatást* eredményez az origó körül. Ez könnyen belátható a 3-2. ábra alapján.



3-2. ábra. Síkbeli forgatás az origó körül

Az ábra jelöléseivel:

$$\begin{aligned} x &= r \cos \alpha, & y &= r \sin \alpha; \\ x' &= r \cos(\alpha + \varphi), & y' &= r \sin(\alpha + \varphi); \\ r &= \sqrt{x^2 + y^2} & & (= \sqrt{x'^2 + y'^2}). \end{aligned}$$

Azonos átalakításokkal az

$$\begin{aligned} x' &= x \cos \varphi - y \sin \varphi; \\ y' &= x \sin \varphi + y \cos \varphi \end{aligned}$$

összefüggésekhez jutunk; ezekből kiolvashatók az A mátrix elemei.

Megjegyezzük még a következőket:

- Az általános 2×2 -es mátrixszal leírt síkbeli lineáris transzformáció egyedül az origót hagyja változatlanul.

- A síkbeli eltolás 2×2 -es mátrixszal nem valósítható meg.
- Tetszőleges T területű síkbeli alakzat transzformációja után az új területet a mátrix determinánsa ismeretében a

$$T' = T |\underline{A}| = T(a_{11}a_{22} - a_{12}a_{21})$$

kifejezésből kapjuk meg.

- A lineáris transzformációk egymásutánját az egyes transzformációk mátrixának *szorzataként* fejezhetjük ki. Mivel a mátrixszorzás nem kommutatív művelet, a tényezők *sorrendje lényeges*. Ennek szemléltetésére lássuk a következő példát:
 - a) A $P(x, y)$ pont 90° -os origó körüli elforgatásának és y tengelyre való tükrözésének eredménye:

$$[x' \ y'] = [x \ y] \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = [-y \ x] \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = [y \ x].$$

- b) Ugyanennek a pontnak előbb az y tengelyre való tükrözése, majd az origó körüli 90° -os elforgatása után láthatóan más eredményre jutunk:

$$[x' \ y'] = [x \ y] \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = [-x \ y] \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = [-y \ -x].$$

3.2.2 HOMOGEN KOORDINÁTÁK

Ha $h \neq 0$, akkor az (x, y, h) számhármaszt a $P(\frac{x}{h}, \frac{y}{h})$ síkbeli pont *homogén koordinátáinak* nevezzük, és ezt a továbbiakban az

$$(x^*, y^*) \sim (x, y, h) \quad (3-3)$$

szimbólummal jelöljük [3-18]. Itt (x^*, y^*) az euklidészi sík tetszőleges pontja. Beláthatjuk, hogy ugyanannak a pontnak végtelen sokféle homogén koordinátás megfelelője van; (3-3) alapján ugyanis

$$(x^*, y^*) \sim (\lambda x, \lambda y, \lambda h) \quad (\lambda \in \mathbb{R})$$

is áll, hiszen

$$(\lambda x, \lambda y, \lambda h) \sim \left(\frac{\lambda x}{\lambda h}, \frac{\lambda y}{\lambda h}\right) = \left(\frac{x}{h}, \frac{y}{h}\right) = (x^*, y^*)$$

Az (x^*, y^*) síkbeli pont *normalizált homogén koordinátáinak* nevezzük az (x, y, h) számhármaszt, ha a homogenitási reláció mellett a $h = 1$ feltétel is teljesül. A (3-3) szerint minden síkbeli pontnak létezik normalizált homogén koordinátás megfelelője:

$$(x^*, y^*) \sim (x^*, y^*, 1).$$

Megjegyezzük, hogy az origótól a síkon az (x^*, y^*) irány mindkét értelmében a végtelen távoli pont homogén koordinátás alakja: $(x^*, y^*, 0)$.

A homogén koordináták transzformációi is leírhatók mátrixokkal. Mielőtt áttekintenénk ennek előnyeit, lássuk

be, hogy az összes lehetőség továbbra is rendelkezésre áll, ami a 2×2 -es mátrixműveletekkel adódott. Ugyanis:

$$(x^*, y^*) \underline{A} \sim (x^*, y^*, 1) \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

hiszen a "homogén 1" a 3×3 -as mátrixművelet után is megmarad.

3.2.2.1 Eltolás és perspektív transzformáció

A sík pontjainak megfeleltetett homogén koordináták 3×3 -as transzformációs mátrixa általános esetben négy fő részre osztható:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} & & a_{13} \\ & \underline{A} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Mint azt korábban láthattuk, az $a_{11}, a_{12}, a_{21}, a_{22}$ elemek segítségével léptékváltást, tükrözést, nyírást és forgatást valósíthatunk meg.

Vizsgáljuk meg a többi rész szerepét külön-külön.

1. Az a_{31} és az a_{32} elem segítségével *eltolást* végezhetünk; ugyanis:

$$(x^*, y^*, 1) \left[\begin{array}{cc|c} \underline{E}_2 & & 0 \\ & & 0 \\ \hline a_{31} & a_{32} & 1 \end{array} \right] = (x^* + a_{31}y^* + a_{32}, 1).$$

(\underline{E}_2 a másodrendű egységmátrixot jelöli.)

2. Ha $a_{33} \neq 1$, akkor

$$(x^*, y^*, 1) \left[\begin{array}{cc|c} \underline{E}_2 & & 0 \\ & & 0 \\ \hline 0 & 0 & a_{33} \end{array} \right] = (x^*, y^*, a_{33}).$$

A homogén koordináták definíciójából következik, hogy ez a választás mindkét koordinátairányban $1/a_{33}$ arányú, egyidejű léptékváltást ("zoom") eredményez.

3. Ha $a_{13} \neq 0$ és $a_{23} \neq 0$, akkor a transzformáció

$$(x^*, y^*, 1) \left[\begin{array}{cc|c} \underline{E}_2 & a_{13} \\ & a_{23} \\ \hline 0 & 0 & 1 \end{array} \right] = (x^*, y^*, a_{13}x^* + a_{23}y^* + 1)$$

alaku, és a transzformált homogén koordináták az

$$\left(\frac{x^*}{a_{13}x^* + a_{23}y^* + 1}, \frac{y^*}{a_{13}x^* + a_{23}y^* + 1} \right)$$

sikbeli pontnak felelnek meg. Belátható, hogy a_{13} , illetve a_{23} növelésével a transzformált pont egyre közelebb kerül az origóhoz, így ez a transzformáció a

centrális projekciónak felel meg; vagyis az a_{13} és a_{23} elem változtatása síkbeli *perspektív transzformációt* eredményez [3-18].

Összetettebb transzformációkra példaként tekintsük a sík tetszőleges pontja körüli forgatást. Legyen a forgatás centruma (u, v) , a forgatás szöge pedig φ . A kívánt transzformációt három lépésben valósíthatjuk meg:

- a forgatás centrumát eltoljuk az origóba;
- forgatást végzünk az origó körül a megadott szöggel;
- az origót eltoljuk a forgatás centrumába.

Ez homogén koordináták felhasználásával így írható fel:

$$(x, y, 1) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -u & -v & 1 \end{bmatrix} \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ u & v & 1 \end{bmatrix}$$

Ha elvégezzük a mátrixszorzásokat, a transzformáció mátrixát a

$$\begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ -u(\cos \varphi - 1) + v \sin \varphi & -u \sin \varphi + v(\cos \varphi - 1) & 1 \end{bmatrix}$$

alakba írhatjuk.

3.2.3 NEMLINEÁRIS KOORDINÁTATRANSZFORMÁCIÓK

A gyakorlati képfeldolgozási problémák jelentős részében a geometriai korrekció *nem valósítható meg* kielégítő pontossággal *linedris* koordinátatranszformációval.

A korrekciókat rendszerint valamilyen fizikai jelenség teszi szükségessé, és a koordinátatranszformációt leíró függvénykapcsolat is ezt a fizikai jelenséget fejezi ki. Olykor bonyolult kétváltozós függvények alkalmazására van szükség, ezek kezelése nehézkes, kiértékelésük lassú. Részletesebb elemzés nélkül két megoldást említünk.

A probléma egyik részének áthidalására általánosan elterjedt módszer a bonyolult függvények *közelítése* egyszerűbb függvényekkel (pl. a különféle sorfejtések, lásd F4.3 pont). Ez arra a megfontolásra épül, hogy az egyszerűbb függvények jobban kezelhetők, miközben a közelítés pontossága tetszőleges lehet. Képfeldolgozási alkalmazásra található példa a 3.4. pontban.

A képfeldolgozás területén különösen nagy jelentősége van annak, hogy a pozíciószámításhoz alkalmazott függvények kiértékelése *gyors* legyen. Ez érthető, hiszen a képi adatmennyiség általában milliós nagyságrendű függvénykiértékelést tesz szükségessé. A gyorsításhoz nyújt segítséget a *szakaszonkénti görbeíves közelítés* (spline approximation, [F-3]). Ekkor az alkalmazott függvények lokális érvényűek és helyről helyre változnak. Köztük természetesen folytonossági és simasági határfeltételeknek kell teljesülniük. A lokálisan alkalmazott függvények alacsony fokszámú (leggyakrabban első vagy harmadfokú) polinomok. A 3.3.5 alpontban leírt felületelemes közelítés esetében a lokális függvénykapcsolat lineáris, és a határfeltételek csak a függvényértékek azonosságát követelik meg.

3.3 A GYAKORLATI MEGVALÓSÍTÁS PROBLÉMÁI

A következőkben a geometriai korrekció megvalósításának részproblémáival foglalkozunk. Egyrészt megkísérel-

jük az eddigiek alapján kiválasztani a célkitűzéseinknek leginkább megfelelő megoldásokat, másrészt új, hatékony eljárásokat javasolunk.

3.3.1 A HATÉKONYSÁGNÖVEDELÉS ELVI LEHETŐSÉGEI

Adott feladat számítógépes megoldásánál a hatékonyság növelésével három lényegesen különböző megközelítésben próbálkozhatunk; vegyük sorra ezeket a geometriai korrekciók szemszögéből.

1. *Hatékonyabb algoritmusok keresésére* kétféle lehetőség kínálkozik:

a) *Ismert algoritmusok felgyorsítása*

- Táblázatok alkalmazása olyan számítások elvégzésére, melyek mind cimtartományban, mind pedig értékkészletben eléggé korlátozottak (pozíció → pozíció számítások, világosságkód-cserék a (2-5) képlet szerint). Az értékkészletet elegendő egyszer, előzetesen feltölteni, a tényleges számítás pedig táblakezeléssé alakul.
- Adott műveletsorozat elvégzésének átszervezése. Pl. rövid ciklusok esetén a ciklusváltó figyelése és döntés helyett célszerűbb a ciklusmag megfelelő számu, egymás utáni végrehajtása (lásd: diszkrét egyenes generalása a 3.3.6.4 alpontban).
- Az algoritmusok átfogalmazása a megengedett hibáktól függő döntéseket tartalmazó formára. A diszkrét pozíciók, valamint a diszkrét (és kisszámu) világosságkódok feleslegessé

teszik a részletszámítások túlzott pontosságát.

- Az algoritmusok átfogalmazása a korábbi rész-eredményektől való eltéréseket tartalmazó alakra. A digitális képek speciális adatszerkezete (pl. a szomszédos képpontok kódja közti korreláció vagy a szokásos geometriai korrekcióknál a pozíciók rendezésének változatlansága) olyan belső összefüggéseket jelent, amelyek felhasználásával csökkenthetők az egyes számításokra eső pontossági követelmények.

b) *Ujszerű algoritmusok kidolgozása*

- Aritmetikai számítások kiváltása logikai döntésekkel. (Lásd pl. interpoláció kiváltása logikai függvények felhasználásával a [3-23]-ban.)
- Integrálisan kellő finomsággal vezérelhető, de túlnyomórészt fixpontos számításokat tartalmazó megoldás alkalmazása a valós aritmetikájú megoldások helyett. (Pl. tetszőleges meredeekségű egyenes generálása fixpontos számításokat tartalmazó, inkrementális algoritmus segítségével, lásd [3-14].)

2. *Optimalizálás a konkrét feladat és a hardver környezet jellemzőinek ismeretében. Optimalizálni lehet pl.:*

- a konkrét feladathoz (pl. [3-6]-ban speciális forgatási szögekre adnak hatékony megoldást);
- a konkrét, alkalmazandó számítástechnikai eszközhöz (pl. [3-5]-ben a lemezes adatkezelés optimalizálását vizsgálják); stb.

3. *Új számítástechnikai eszközök és ezekhez illeszkedő új eljárások kifejlesztése; ilyen pl. a több-*

processzoros, hálózati szervezésű rendszer. (A digitális képek adatszerkezetéhez és a képfeldolgozási feladatok egy részéhez pl. jól illeszkedik a tömbprocesszorok alkalmazása, lásd [3-24].)

A következő alpontokban ismertetett eljárások és megoldások alapjaikban nem tekinthetők újnak, így az 1/a, illetve a 2. csoportba sorolhatók. A 3. csoportba tartozó megközelítések - célkitűzéseink szerint - nem jöhetnek szóba.

3.3.2 A KORREKCIÓS ÖSSZEFÜGGÉS ÉRTELMEZÉSE

Feltételezve, hogy a korrekciós összefüggés invertálható, a következő két megoldás elvileg azonos eredményt ad:

- a korrigálandó képet transzformáljuk a korrigált kép koordináta-rendszerébe, azaz meghatározzuk a korrigálandó kép pontjainak a korrigált rendszerbeli pozícióját (*direkt transzformáció*);
- a korrigált kép rendszerét transzformáljuk a korrigálandó kép koordináta-rendszerébe, azaz a korrigált rendszerbeli pozíciókhoz határozzuk meg a korrigálatlan rendszerbeli pozíciókat (*inverz transzformáció*).

Az utóbbi megoldás a korrekciós összefüggés inverz megfogalmazását igényli ugyan, alkalmazása azonban több előnnyel is jár. A következők szólnak mellette:

- Inverz transzformáció esetén a korrigált rendszerbeli kép minden pozíciója definiált lesz; nincs szükség előzetes értékadásra, illetve utólagos interpolációra.
- Több korrigálandó kép egyazon korrigált képre való illesztése könnyebben valósítható meg (nincs szük-

ség az illesztési határok transzformációjára).

- Inverz transzformáció esetén a korrekt újra-mintavételezés azonnal megvalósítható, míg direkt transzformáció esetén ehhez a pontos pozíciókat átmenetileg tárolni kellene.

Mindezek alapján egyértelműen célszerűbb az inverz transzformáció alkalmazása.

3.3.3 A GEOMETRIAI KORREKCIÓ SZERVEZÉSE

Mivel kizárólag szekvenciális megoldások jöhetnek szóba, van értelme foglalkozni a képpontok feldolgozási *sorrendjével*.

A jelenlegi számítástechnikai eszközök általában nem teszik lehetővé, hogy a korrigálandó (illetve a korrigált) képet teljes egészében a főtárban tároljuk. Ezért használnunk kell háttértárat (mágneslemez, mágnesszalag stb.) is, amittől a hozzáférési idő nagyságrendekkel megnő.

A lemezes adatforgalom gazdaságos megszervezésére a szakirodalomban többféle javaslat található. (Ilyen pl. a [3-5]-ben leírt megoldás szűrési eljárások hatékony megvalósítására kisszámítógépen.)

Kihasználva, hogy a digitális képek (a többsávós úrfelvételek kivételével) rendszerint *sorfolytonosan* helyezkednek el, a korrekciós algoritmust - első közelítésben - úgy célszerű megvalósítani, hogy minél kevesebbszer kelljen több sort felváltva érintő műveleteket végezni.

Ha mindkét kép háttértárolón van, belátható, hogy az optimum eléréséhez egyiküket - mégpedig ha az inverz korrekciót részesítjük előnyben, akkor a kimenő képet - sorfolytonosan kell felhasználni.

Általános esetben a kimenő kép egy sorának összeállításához több bemenő képbeli sort érintő feldolgozásra van

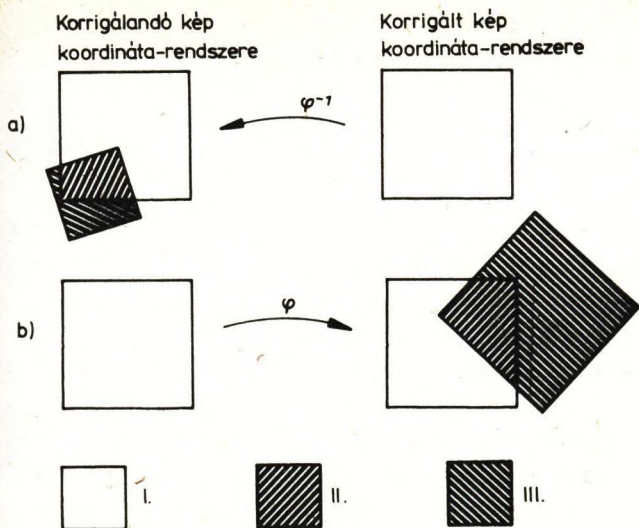
szükség, s ez háttértárkezelést is igényel. Tegyük fel, hogy egy korrigált képbeli sor létrehozásához m -szer kell háttértárhoz fordulni, s ennek időigénye mt_L . A korrekciós összefüggés kissé bonyolultabbá válik ugyan, viszont kis m értékek esetén komoly nyereséget eredményez az *alternáló* megoldás alkalmazása. Ez azt jelenti, hogy a korrigált képsorokat felváltva növekvő, majd csökkenő sorindex szerint határozzuk meg. Ekkor ugyanis a sorok végein eggyel kevesebbszer kell háttértárhoz fordulni, s ezzel a fenti idő $(m - 1)t_L$ értékre csökken ($m > 1$).

Hatékony gyorsítást jelent, ha az adatátviteli időt kihasználjuk műveletvégzésre. (Az átvitel többnyire DMA (direct memory access) megoldással történik, ami csak a processzor által szervezett adatforgalom szüneteiben terheli a sinrendszert.) Ennek alapfeltétele az egynél több adatpuffer használata. Két puffer alkalmazása (*kettős pufferezés*) esetén - ha egy bemenő adathalmaz feldolgozásának időigénye t_k - a nyereség $\min(t_L, t_k)$.

3.3.4 A KORREKCIÓS ADATHALMAZ SZÜKITÉSE

A gyakorlatban mind a korrigálandó, mind pedig a korrigált kép korlátos méretű.

A korrekciós összefüggés általában *nem felelteti meg* egymásnak a két kép határait. Tekintsük pl. a 3-3. ábrát. A bevonalkázott területek az a) esetben a korrigált, a b) esetben a korrigálatlan képnek megfelelő transzformált képet jelentik; az előbbi (az *őskép*) inverz, az utóbbi direkt transzformációval jött létre. Láthatjuk, hogy csak a II. típusu vonalkázással jelölt* részek esnek az eredeti határokon belül.



3-3. ábra. Korlátos méretű képek korrekciója

fsszerűek a következő kivánalmak:

- Ne kelljen külön ellenőrizni, hogy a korrigálatlan kép pontjai a korrigált kép határain belülre kerülnek-e.
- A korrigálatlan kép azon pontjaival, melyek nem kerülnek a korrigált képre, ne kelljen foglalkozni. (Azaz ne a korrekció végrehajtása után derüljön ki, hogy feleslegesek.)

E két kívánalmat az inverz értelmű korrekció automatikusan teljesíti; a korrekció az ábra szerinti a) esetben a korrigálható kép II. jelzésű pontjait érinti. Látható viszont, hogy a korrekciót felesleges a teljes korrigált képre végrehajtani. (Az ábrán a b) eset I. jelzésű korrigált pontjainak ősképe a korrigálható kép határain kívül van.) Ez indokolja a következő kívánalmat is:

- A korrigált kép azon pontjaival, melyeknek ősképe a korrigálatlan kép határain kívülre esik, ne kelljen foglalkozni.

Mindhárom kívánalmat együttesen teljesíti pl. a következő algoritmus:

1. Transzformáljuk a korrigálandó kép határát direkt transzformációval a korrigált kép koordináta-rendszerébe.
2. Képezzük ennek és a korrigált kép határának metszetét. Ez határolja a korrigált kép azon részét, amelynek ősképe a korrigálatlan képen belülre esik, azaz amelyet a korrekció eredményeként kaphatunk meg.
3. Hajtsuk végre az inverz értelmű geometriai korrekciót a metszet határain belüli pontokra.

Az elérhető műveletigény-nyereség durván a korrigált kép teljes és hasznos területének különbségével arányos. Megjegyzendő, hogy e megoldás a korrekciós összefüggés direkt és inverz értelmű megfogalmazását egyaránt igényli. Ha ez valamilyen okból nehézségekbe ütközik, akkor a probléma pl. a képhatárok jobb megválasztásával oldható meg.

3.3.5 FELÜLETELEMES KÖZELITÉS

Az általánosság megsértése nélkül fel lehet tételezni, hogy kis tartományban minden szóba jöhető geometriai korrekció lineárisan viselkedik. Ezért a globálisan nem-lineáris korrekciók megvalósításánál is ináckolt a lineáris korrekciós algoritmusok használata, megfelelő módosításokkal. Ez a módosítás történhet pontról pontra, de "sima" korrekciós összefüggés esetén ennél lényegesen ritkábban is. Elvileg jó megoldás volna, ha a lassan vál-

tozó helyeken ritkábban, a gyorsan változó helyeken sűrűbben kerülne sor módosításra.

A szükséges módosításokat tekintve ugyan nem, de számítástechnikai szempontból optimális a következő algoritmus:

1. Fedjük le az egyik (a továbbiakban: "A") rendszerbeli képet szabályos rácsozatban elhelyezkedő téglalapokkal (ezeket a továbbiakban *felületelem*nek nevezzük), amelyeken belül a korrekciót elegendő pontossággal megvalósíthatjuk lineáris transzformációkkal. (Minél "simább" a korrekciós összefüggés, annál több képpont tartozhat egy felületelemhez.)
2. Számítsuk ki a korrekciós összefüggés segítségével a felületelemek sarokpontjainak a másik (a továbbiakban: "B") rendszerbeli pozícióját. Ezeket a továbbiakban *vezérlő adatoknak* nevezzük.
3. A tényleges geometriai korrekciót hajtsuk végre külön-külön minden egyes felületelemre a sarokpontok "B" rendszerbeli pozíciójának ismeretében, mint általános négypontos transzformációt (lásd 3.3.7 alpont). Invertálható korrekciós összefüggés esetén ezek a "B" rendszerben is diszjunktak, és teljesen lefedik a képet.

E megoldás mellett a következők szólnak:

- A változó sűrűségű módosítás vezérléséhez külön számításokra volna szükség; ez a szabályos esetben elmarad.
- A vezérlő adatok létrehozására ugyan a szabályos esetben többször van szükség, mint ahányszor valóban indokolt volna (hiszen a leggyorsabban változó részre kell felkészülni), de a korrekció tényleges számítási igényét elsősorban ugyis a képpontokkal kapcsolatos, sokkal nagyobb mennyiségű művelet szabja meg.

Az előzőekben (lásd 3.3.2) leírtaknak megfelelően legyen az "A" rendszer a korrigált kép (= kimenő rendszer), a "B" rendszer pedig a korrigálandó kép rendszere (= bemenő rendszer). Ekkor a korrekciós összefüggést inverz értelemben kell meghatározni. A sorfolytonos adatkezelés érdekében a vezérlő adatok segítségével kijelölt általános négyponos korrekciók közül egyidejűleg kell végrehajtani azokat, amelyek ugyanarra a képsorra vonatkoznak. Így a tényleges korrekció a vezérlő adatokból alkotott *vezérlő sáv* segítségével történik.

Becsüljük meg a szakaszonként lineáris közelítéssel elkövetett hibát, az egyszerűség kedvéért egyváltozós esetben, Legyen az $f(x)$ közelítendő függvény konvex az $[a, b]$ intervallumban, és $T_1(x)$ az egyenletesen legjobban közelítő elsőfokú polinom. A Csebisev-tétel felhasználásával levezethető, hogy ez az $(a, f(a))$ és a $(b, f(b))$ ponton keresztül huzott szelő, valamint az ezzel párhuzamos érintő között középben elhelyezkedő, velük párhuzamos egyenes lesz. Így az elkövetett hiba *normájára* (lásd F3.2.1.1 alpont) általában a következő határértékek adódnak [F-3]:

$$E_n(f) = \|f(x) - T_n(x)\| \begin{cases} \leq \max_{[a,b]}(|f^{(n+1)}(x)| \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!}); \\ > \min_{[a,b]}(|f^{(n+1)}(x)| \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!}). \end{cases} \quad (3-4)$$

Esetünkben az egyenletesen legjobb közelítést adó egyenes helyett célszerűbb a szelőt választani. A korrekciós összefüggést ekkor csak a felületelemek sarokpontjaira kell végrehajtani, és nincs szükség egyéb számításon; az elkövetett hiba viszont a (3-4) szerinti értékhatárok kétszerese közé esik.

Ha a teljes $[x_0, x_V]$ intervallumban legfeljebb $E_1(f)$ hibát engedünk meg, akkor (3-4) átalakításával és $n = i$ behelyettesítéssel a maximális osztásköz:

$$(\Delta x)_{\max} = 2 \sqrt{\frac{2E_1(f)}{\max_{[x_0, x_V]} (|f'''(x)|)}}$$

A javasolt közelítés számításigényének vizsgálatához tekintsünk egy $K \times L$ képpontból álló képet s ezen egy $a \times a$ adatot tartalmazó vezérlő rácsot. Legyen a korrekciós összefüggés pontenkénti kiértékeléséhez szükséges idő t_f , az általános négyponos korrekció egy pontra eső számításigénye t_p , ennek módosítása t_{mp} (lásd 3.3.7 alpont). Ekkor a javasolt közelítés, illetve az eredeti korrekció számításigényének aránya:

$$\frac{a^2(t_f + t_{mp}) + KLt_p}{KLt_f} \sim \frac{a^2}{KL} + \frac{t_p}{t_f}$$

feltéve, hogy

$$t_f \gg t_{mp}, \quad \text{és} \quad \sqrt{KL} \gg a.$$

Bonyolult, de közel lineáris korrekciós összefüggés esetén a nyereség elérheti a két nagyságrendet is.

3.3.6 DISZKRÉT EGYENES SZAKASZ LÉTREHOZÁSA

A megadott meredekségű, illetve két tetszőleges pontot összekötő diszkrét egyenest interpoláló algoritmusok-

nak kiterjedt irodalmuk van (lásd pl. [3-2], [3-4], [3-7], [3-14], [3-15], [3-16], [3-19], [3-24], [3-25], [F-5], [F-7] stb.).

Legyen az $e \in g_p$ folytonos, origón átmenő, racionális meredekségű és az első síknyolcadba eső egyenes egyenlete (lásd F3.2.3.2 alpont):

$$e: mx = ny \quad (m \in \mathbb{N}, n \in \mathbb{N}, 0 \leq m \leq n). \quad (3-5)$$

Az egyszerűség kedvéért korlátozzuk vizsgálatainkat az ilyen egyenesek diszkrét megfelelőjére. Ez nem jelenti az általánosság megsértését, mivel tükrözéssel és eltolással tetszőlegesen egyenes ilyen helyzetbe hozható.

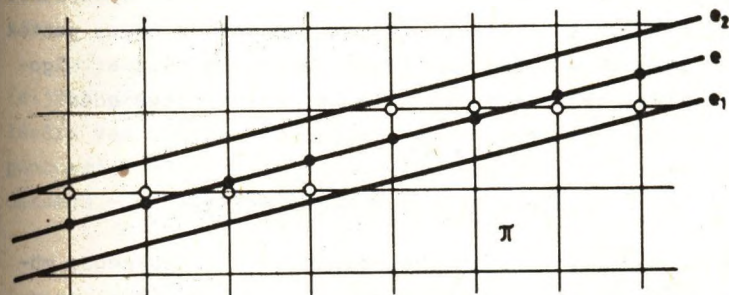
Célkitűzésünk: olyan algoritmusok keresése, melyek hatékonyan állítják elő a megfelelő diszkrét egyenes pontjainak koordinátáit. Az előállítás sorrendjére - egyelőre - nem teszünk megkötést, és az algoritmusok tárigényét - kicsinységük miatt - nem vizsgáljuk. Mivel az előkészítő műveletek időigénye többnyire elhanyagolható, az algoritmusok hatékonyságát az egyes egyenespontok létrehozásának sebességével mérjük.

3.3.6.1 Rácsmetszéses módszer

Keressük a (3-5) egyenlet szerinti folytonos egyenes $\lambda \in \gamma_\pi$ diszkrét megfelelőjét. Utalva az F3.2.3.2 alpontban leírt rácsmetszéses módszerre, ez azt jelenti, hogy keressük a folytonos egyenes és a hálónonalak metszéspontjaihoz *legközelebb eső* rácspontok halmazát. Mivel a koordinátairányban szomszédos rácspontok távolsága 1, egyetlen ilyen metszéspont sem lehet $1/2$ -nél távolabb a legközelebbi rácsponttól. Ebből következik, hogy az e egyenes λ diszkrét megfelelőjéhez az

$$e_1: mx = n(y + \frac{1}{2}) \quad \text{és az} \quad e_2: mx = n(y - \frac{1}{2})$$

egyenes által határolt sávba eső rácspontok tartozhatnak (lásd 3-4. ábra)..



3-4. ábra. Folytonos egyenes diszkrét megfelelője

A hasonló háromszögek tételének felhasználásával be-
látható, hogy - az $m \leq n$ megkötés következtében - a
szóba jöhető rácspontokhoz az ordinátával párhuzamos $x = k$
($k \in \mathbb{N}$) hálónonalakon levő metszéspontok minden esetben kö-
zelebb vannak, mint az abszcisszairányu hálónonalakon le-
vők. Mivel az e_1 és az e_2 egyenes távolsága ordinátairány-
ban 1, a hálónonalszakaszok belsejében legfeljebb egy
rácspont lehet, és ez szükségszerűen a λ diszkrét egyenes-
hez tartozik. Másrészt az e_1 és az e_2 határolta sávon be-
lül ezeken kívül nincs más rácspont, azaz e sávon belül
minden rácspont a λ diszkrét egyenes pontja.

Az előbbiekből következően a λ diszkrét egyenes pont-
jainak egyik legegyszerűbb generálási módja az

$$y = \left\lceil \frac{m}{n} x - \frac{1}{2} \right\rceil, \quad \text{illetve} \quad y = \left\lfloor \frac{m}{n} x + \frac{1}{2} \right\rfloor \quad (3-6)$$

mennyiségek kiszámítása minden $x \in \mathbb{N}$ értékhez. Itt $\lceil x \rceil$ az x -nél nem kisebb; $\lfloor x \rfloor$ az x -nél nem nagyobb legközelebbi egész számot jelöli. (Az $\lfloor x \rfloor$ jelölést szimmetria okokból használtuk az $\text{ent}(x)$ helyett, amivel megegyezik.)

Az elvet megvalósító algoritmus igen rossz hatékonyságu: pontonként egy-egy valós szorzást, valós összeadás és valós-egész átalakitást igényel. Mivel az egyes pontok koordinátáinak kiszámítása független egymástól, az algoritmus alkalmazása során nem léphet fel hibaterjedés.

3.3.6.2 Rekurzív megvalósítás

Hatékonyabb algoritmushoz jutunk, ha a szükséges műveletszámot lecsökkentjük, az egyenes pontjai közti szomszédsági összefüggések kihasználásával.

Tekintsük az előbbi $e \in g_p$ folytonos egyenest, és tegyük fel, hogy a megfelelő $\lambda \in \gamma_\pi$ diszkrét egyenes $P(x, y)$ ($x \in \mathbb{N}, y \in \mathbb{N}$) pontjánál vagyunk.

A (3-5) egyenlet alapján látható, hogy két szomszédos egyenespont ordinátáinak különbsége:

$$\Delta y = \frac{m}{n}.$$

Kézenfekvő egy olyan algoritmus, amely a diszkrét egyenes pontjait sorozatként állítja elő:

$$\begin{aligned} P_0(x_0, y_0) &= P(x, y); \\ P_1(x_1, y_1) &= P_1(x_0 + 1, \text{int}(y_0 + \Delta y)); \\ P_2(x_2, y_2) &= P_2(x_0 + 2, \text{int}(y_0 + 2\Delta y)); \\ &\vdots \\ P_k(x_k, y_k) &= P_k(x_0 + k, \text{int}(y_0 + k\Delta y)) \quad (k \in \mathbb{N}); \end{aligned}$$

ahol $\text{int}(x)$ az x -hez legközelebbi egész számot jelöli. A pontonkénti számításigény: egy-egy valós összeadás, valós-egész átalakítás és egész összeadás.

Ha az m/n meredekség az adott számábrázolás keretei közt kifejezhető véges bináristört alakban (azaz Δy értéke pontos), akkor a számítások során nem halmozódik fel hiba.

Megjegyzendő, hogy a képfeldolgozási gyakorlatban legfeljebb néhány ezer pontból álló diszkrét egyenes generálására van szükség, s így a szokásos valós számábrázolási pontosság mellett egy ilyen algoritmus tetszőleges meredekségű egyenesre kielégítő eredményt ad.

3.3.6.3 Számolás egész értékű ordinátákkal

További javulást érhetünk el a valós-egész átalakítás kiváltásával. Ugyanis az előző alpontban leírt elv szerint - az adott megkötések mellett - mind az abszcissza-, mind pedig az ordinátaértékek monoton növekvő számok sorozataként állnak elő. Az abszcisszák sorozata valóban szigorúan monoton - ti. a természetes számok egymásutánja -, az ordinátaértékek sorozatában azonban az egymás utáni tagok közti különbség 0 vagy 1 lehet. Így az egyenesgenerálási algoritmus szerepe pusztán annak megállapítása, hogy a soron következő ordinátaérték azonos-e a megelőzővel, vagy annál eggyel nagyobb.

Feltéve, hogy a soron következő és az aktuális ordináta nagyságrendje (gépi ábrázolásban a valós szám kitevő része) azonos, a valós-egész átalakítás egyetlen - az "1" helyiértékű - (mantissza) bit vizsgálatára egyszerűsödik. Ha ez megváltozott az előző értékhez képest, akkor az egyenes ordinátája átlépte a következő egységhatárt, vagyis az új ordináta eggyel nagyobb a megelőzőnél. (Ha az algorit-

mus indítása előtt az ordinátát $1/2$ -del eltoljuk, az eredmény pontos lesz.)

Az eljárás során célszerű előre meghatározni a generálandó egyenes ordinátáinak nagyságrendjét, és szükség esetén az egyenest felbontani olyan szakaszokra, amelyekben belül a nagyságrend állandó.

A módosított algoritmus a következő:

1. A kezdeti változók kiszámítása: az egyenes szegmentálása olyan szakaszokra, amelyekben belül az ordináták nagyságrendje változatlan.
2. Kezdőértékek beállítása: $\Delta y = m/n$; $P_0'(x_0', y_0') = P_0(x_0, y_0)$ kezdőpont; $k = 0$.
3. Új koordináták kiszámítása:

$$x_{k+1}' = x_k' + 1 \quad (x_k', x_{k+1}' \in \mathbb{N});$$

$$y_{k+1}'^* = y_k'^* + \Delta y \quad (y_k'^*, y_{k+1}'^* \in \mathbb{R}).$$

4. Korrekció: ha $y_{k+1}'^*$ "1" helyiértékű bitje változott, $y_{k+1}' = y_k' + 1$; különben $y_{k+1}' = y_k'$ ($y_k', y_{k+1}' \in \mathbb{N}$).
5. Ciklusváltás: ha az adott egyenesszakaszra vonatkozó végfeltétel nem teljesül, akkor $k = k + 1$, és ismétlés 3-tól; különben áttérés a következő szakaszra.

A képpontonkénti számításigény csökkent: a valós-egész átalakítás helyett ebben az algoritmusban egy bittől függő döntés szerepel. Az algoritmus egyéb jellemzői azonosak az előzőével.

3.3.6.4 Adaptív megvalósítás

Az algoritmus tovább gyorsítható, ha kihasználjuk, hogy a különböző meredekségű egyeneseknél különböző sűrűséggel van szükség az ordináta megváltoztatására a 4. lépés szerinti vizsgálat után.

Legyen pl.: $\frac{m}{n} = \frac{2}{11}$. Ekkor $S_1 = \text{ent}\left(\frac{n}{m}\right) = 5$ esetben az ordináta "1" helyiértékű bitje nem változik, azaz vizsgálata szükségtelen. A változás S_1 és $S_2 = S_1 + 1$ között lesz.

Az algoritmus módosítása kézenfekvő: a 4. lépés szerinti vizsgálatot csak a szükséges esetben hajtjuk végre. Az előzetes vizsgálatok során meghatározzuk az

$S_1 = \text{ent}\left(\frac{n}{m}\right)$ mennyiséget, majd minden egyes "1" változás után S_1 hosszúságú ciklusokban vizsgálat nélkül hajtjuk végre a 3. és az 5. lépést. A számításigény ezzel irányfüggővé válik: az algoritmus a kisebb meredekségű egyeneseket nagyobb sebességgel generálja.

3.3.6.5 Egyenesgenerálás fixpontos

műveletekkel

Az eddigi algoritmusok valós számok közötti összefüggéseken alapultak, így lebegőpontos számításokat tettek szükségessé. Mivel ezek végrehajtási ideje sokszorosa a fixpontosakénak, ésszerű törekvés, hogy csak az utóbbiakat kelljen alkalmazni. Ehhez olyan, kizárólag egész számok közti összefüggéseket kell keresnünk, amelyek alkalmasak a diszkrét egyenes leírására.

Alakítsuk át a folytonos egyenes (3-5) egyenletét

$$y - \frac{m}{n} x = 0$$

alakura. A rácsmetszéses módszer szerint

$$P(i, j) \in \lambda \iff -\frac{1}{2} \leq j - \frac{m}{n} i \leq \frac{1}{2},$$

vagyis

$$P(i, j) \in \lambda \iff -n \leq 2nj - 2mi \leq n. \quad (3-7)$$

(A \iff szimbólum jelentése: "akkor és csak akkor".)

Vezessük be a

$$\text{RES} = 2nj - 2mi \quad (3-8)$$

mennyiséget. Ezzel a (3-7) feltétel átfogalmazható:

$$P(i, j) \in \lambda \iff |\text{RES}| \leq n. \quad (3-9)$$

A használt mennyiségek (m, n, i, j, RES) egész számok. A RES mennyiségére szabott korlát a megadott $0 \leq m \leq n$ reláció mellett minden abszcisszaértékhez egy és csak egy ordinátaértéket rendel.

A fenti összefüggés felhasználásával különösen hatékonyan valósítható meg a szomszédos egyenespontok egymásutáni generálása, a következő algoritmus alapján:

1. Legyen a $P(x, y)$ kezdőpont a diszkrét egyenes $P(i, j)$ pontja. Határozzuk meg a RES mennyiséget a (3-8) szerint.
2. Növeljük az abszcissza értékét eggyel.
3. Vizsgáljuk az aktuális RES mennyiséget a (3-9) korlát szerint. Ha nagyobb az adott korlátnál,

akkor az ordináta értékét eggyel növelni kell; egyébként változatlan marad.

4. Ha a végfeltétel nem teljesül (azaz még nem értük el a létrehozandó szakasz végpontját), ismétlés 2-től; különben áttérés a következő szakaszra.

Az algoritmus tömbvázlatát a 3-5. ábrán adjuk meg. (Az ábrán DX, illetve DY az egyenes irányvektorának x , illetve y irányu összetevője; a csillag jelentésére a 3.3.7 alpontban térünk vissza.)

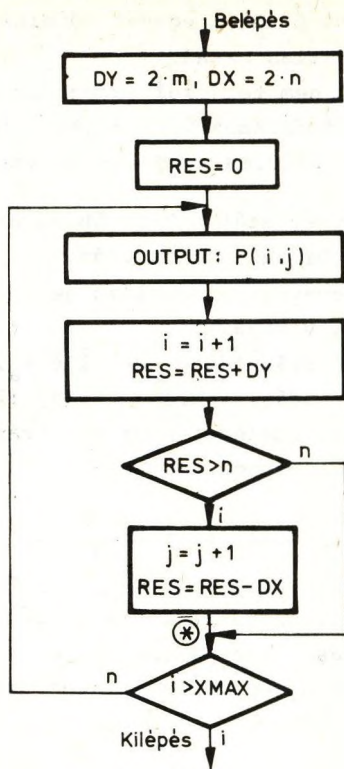
A végrehajtási idő irányfüggő. Jelölje t_a , t_i , illetve t_c a fixpontos összeadás, inkrementálás, illetve összehasonlítás végrehajtási idejét. Ekkor az x irányu, illetve a 45° -os ferde irányu lépéshez

$$t_1 = t_a + t_i + 2t_c; \text{ ill. } t_2 = 2(t_a + t_i + t_c)$$

időre van szükség. (Vö. az F3-5. ábrán a $\overline{P_1B}$, illetve a $\overline{BP_2}$ szakasz megrajzolását.) Az x tengellyel tetszőleges φ szöveget bezáró irányban egy lépés átlagos számítási ideje:

$$\bar{t}_\varphi = t_1(\cos \varphi - \sin \varphi) + t_2 \sin \varphi.$$

Az algoritmus alkalmazása során nem kell hibaterjedéssel számolni. (Kis egész számok közti összeadásokat tartalmaz, és tulcsordulás nem léphet fel.) Az eredmény természetesen csak racionális meredekségű egyenesek esetén pontos.



3-5. ábra. Diszkrét egyenes generálása fixpontos műveletekkel

3.3.6.6 Szakaszok ismétlésének módszere

Az előző alpontbeli algoritmus $\max(|m|, |n|)$ (megkö-
téseink következtében: n) hosszúságu ciklusokban ugyan-
arra a döntésre jut az aktuális RES vizsgálatakor.
Ebből következően:

- ha $P(i, j) \in \lambda$, akkor $P(i + kn, j + km) \in \lambda$ ($k \in \mathbb{N}$);
- ha az algoritmus az első (n hosszúságu) ciklusban az I iránykódsorozattal (lásd F3.1.2 alpont) leírható egyenes szakaszt hozta létre, akkor a λ diszkrét egyenes generálása ennek ismétlésével tetszőleges hosszban folytatható:

$$\lambda: I + I + I + \dots$$

Igy csak az első n pont (vagy iránykód) generálása igényel döntéseket, a többi ponté csak ezen döntések eredményeinek alkalmazását. Hosszu - de rövid ciklusokban generálható - egyenesek létrehozásának pontonkénti számítási igénye közel azonos a két koordináta inkrementálásához szükséges idővel. Minél hosszabb az egyenes, annál kevésbé függ a számítási idő az irányától.

3.3.6.7 Egyéb módszerek

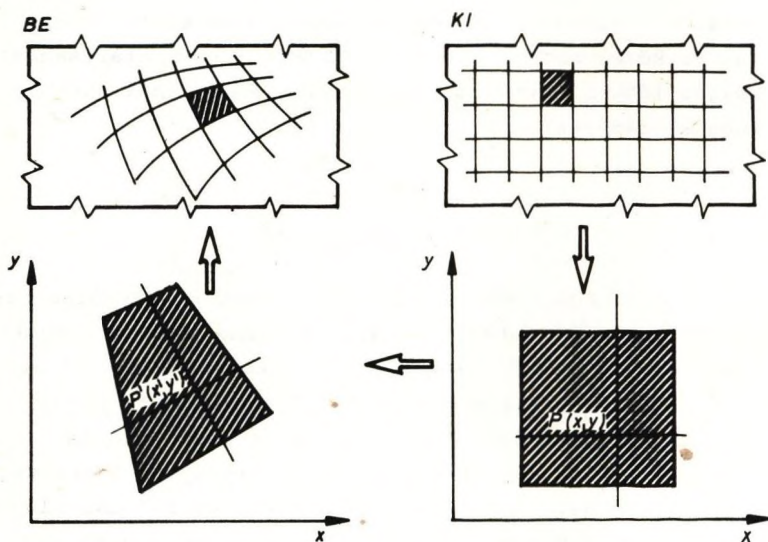
A teljesség kedvéért megemlítjük, hogy a generálás hatékonyságát tovább lehet növelni, ha felhasználjuk a diszkrét egyenes különböző tulajdonságait. (Lásd az F3.2.2 alpontban a Freeman-féle egyenességi kritériumokat; [3-7], [F-10].) E tulajdonságok a diszkrét egyenesek lokális (iránykódra vonatkozó) és kvázilokális (iránykódcsoportra vonatkozó) szabályszerűségeit fejezik ki, és felhasználásukkal az eddigiektől minőségileg eltérő algoritmusok készíthetők. Képfeldolgozási jelentőségük azonban csekély, mivel alkalmazásuk nehézkes.

A diszkrét egyenest létrehozó algoritmusokat eddig egyetlen szempont - egy egyenespont meghatározásához

szükséges átlagos számításigény - szerint értékeltük. A következő alponban emellett a digitális képek korrekciójánál fellépő egyéb szempontokat is figyelembe veszünk és így keressük a számunkra legmegfelelőbb algoritmust.

3.3.7 FELÜLETELEM TRANSZFORMÁCIÓJA

Az előzőekben leírtak szerint a globálisan nemlineáris geometriai korrekció felületelemes közelítését a 3-6. ábrán szemléltetjük.



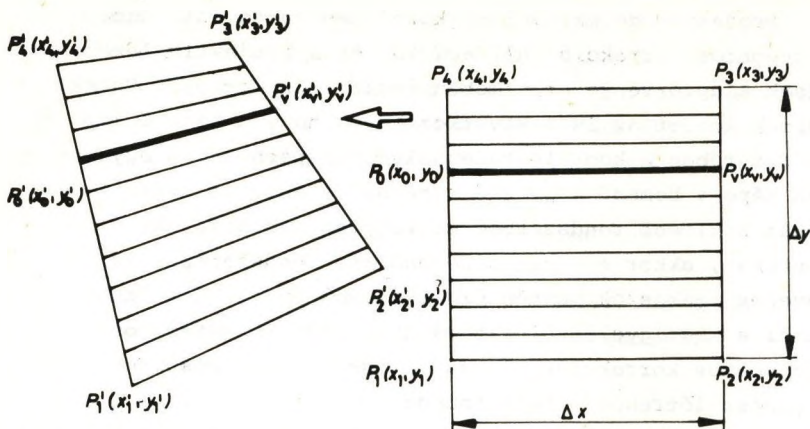
3-6. ábra. Nemlineáris geometriai korrekció közelítése

Projektív geometriai megközelítésben az általános négyponthoz korrekció *kollineáció*, és a projektív leképezések alaptörvénye értelmében *kettősvízytartó*. Ennek többek között az is a következménye, hogy a kimenő kép rendszerében a koordinátatengelyekkel párhuzamos egyenesek képe a bemenő kép rendszerében is egyenes lesz. Ha tehát a kimenő rendszerben sorfolytonosan kezeljük a képpontokat, akkor ezekhez a bemenő kép rendszerében is egyenes szakaszok mentén elhelyezkedő pontok tartoznak. Ettől a megfigyeléstől azt várjuk, hogy az általános négyponthoz korrekciót vissza tudjuk vezetni diszkrét egyenest létrehozó algoritmusokra.

Tekintsük a 3-7. ábrát. A négyszögeken belüli egyenesszakaszok a kimenő rendszerben a képsorokat, a bemenő rendszerben pedig ezek ősképét jelölik. A $\overline{P_0 P_V}$ kimenő rendszerbeli szakasznak a bemenő rendszerben a $\overline{P'_0 P'_V}$ szakaszt kell megfeleltetni. Mivel $y_0 = y_V$ és $x_V - x_0 = \Delta x$ (= a felületelem x irányu mérete), ez a megfeleltetés azt jelenti, hogy a transzformáció elvégzéséhez a $\overline{P'_0 P'_V}$ szakaszt kell egyenletesen Δx részre felosztani.

A teljes felületelem transzformálásához a P_0 pontot a $\overline{P_4 P_1}$ szakaszon, a P_V pontot a $\overline{P_3 P_2}$ szakaszon kell végigléptetni, Δy lépésben; ennek megfelelően a P'_0 illetve P'_V pontok a bemenő rendszerben a $\overline{P'_4 P'_1}$, illetve a $\overline{P'_3 P'_2}$ szakaszon haladnak végig. Kiszámítható, hogy a $\Delta x \cdot \Delta y$ méretű (kimenő rendszerbeli) felületelem összes pontjának pozícióját a fenti módszerrel $4\Delta y + 2\Delta x(\Delta y + 1)$ lépésben kapjuk meg, ahol az első tag a sorok kezdő- és végpontjaira, a második a szakaszpontokra vonatkozik. A teljes műveletszám egyenesen arányos az egyes interpolációs lépések átlagos műveletszámával.

Könnyen belátható, hogy a diszkrét lineáris interpoláció és a diszkrét egyenes létrehozása ekvivalens fel-



3-7. ábra. Felületelem transzformációja

dat, így a nemlineáris geometriai korrekciót ez utóbbi alkalmazására vezettük vissza.

Keressük meg a 3.3.6 alpontban leírt egyenesgenerálási algoritmusok közül a fenti interpolációk elvégzéséhez legalkalmasabbat.

A következő szempontokat kell figyelembe vennünk:

- minél gyorsabb legyen;
- az egyenespontokat (= interpolált adatokat) az eredeti rendezettségnek megfelelően (sorfolytonosan) szolgáltatassa;
- alapvetően diszkrét értékek sorozatát állítsa elő, de szükség esetén könnyű legyen a pontos (valós) érték meghatározása;
- mivel rövid, sűrűn változó meredekségű egyenesek generálására van szükség, az algoritmus ne használja ki a paraméterek közötti speciális összefüggéseket és legyen könnyen átparameterezhető.

A fenti szempontok alapján esetünkben a 3.3.6.5 alponban leirt algoritmus az optimális. A többi eljárás vagy lassu; vagy párhuzamosítást, illetve speciális generálási szekvenciákat alkalmaz; vagy csak hosszú, illetve speciális meredekségű egyenesekre használható hatékonyan. A számításigény szempontjából optimális algoritmusok azért nem alkalmazhatók, mert az egyenespontokat nem szomszédási sorrendben generálják.

Nézzük most részletesebben, hogyan lehet ezt az algoritmust az általános négyponos korrekció megvalósítására felhasználni. Vegyük észre először a következőket:

- Az algoritmusban szereplő RES mennyiség közvetlenül tartalmazza a pontos pozícióhoz szükséges korrekciós tagot. A 3.3.6.5 alponban használt jelölésekkel a pontos pozíciót az

$$y' = y + \frac{RES}{DX} \quad (3-10)$$

képletből lehet kiszámítani, ahol DX az x irányu irányvektor kétszerese. Ezt a korrekciót a 3-5. ábrán csillaggal megjelölt helyen kell elvégezni.

- Ha az algoritmust nem $RES = 0$ beállítással indítjuk, akkor a generált egyenes y irányban

$$y_e = \frac{RES}{DX}$$

mértékben eltolódik.

- Az algoritmus könnyen kiterjeszthető tetszőleges meredekségű esetre. Eredeti megkötéseink szerint a generált egyenes meredekségének 0 és 1 közé kell esnie. Ha ez az egyenes a $P_0(x_0, y_0)$ és a $P_V(x_V, y_V)$

pontot köti össze, akkor ez azt jelenti, hogy

$$\Delta y = (y_v - y_0) \leq (x_v - x_0) = \Delta x.$$

Ilyenkor a $\overline{P_0 P_e}$ szakaszt nagyítjuk $\Delta x / \Delta y$ arányban; itt $P_e = (x_0, y_v)$. Az átalakított algoritmusnak (lásd 3-8. ábra) az $(y_v - y_0) > (x_v - x_0)$ esetre külön ága van: ekkor a $\overline{P_0 P_e}$ szakaszt kicsinyítjük $\Delta x / \Delta y$ arányban. A nagyítást végző jobb oldali ágban a pontos pozíció kiszámításhoz szükséges korrekciót a (3-10) képlet, a kicsinyítést végző (bal oldali) ágban az

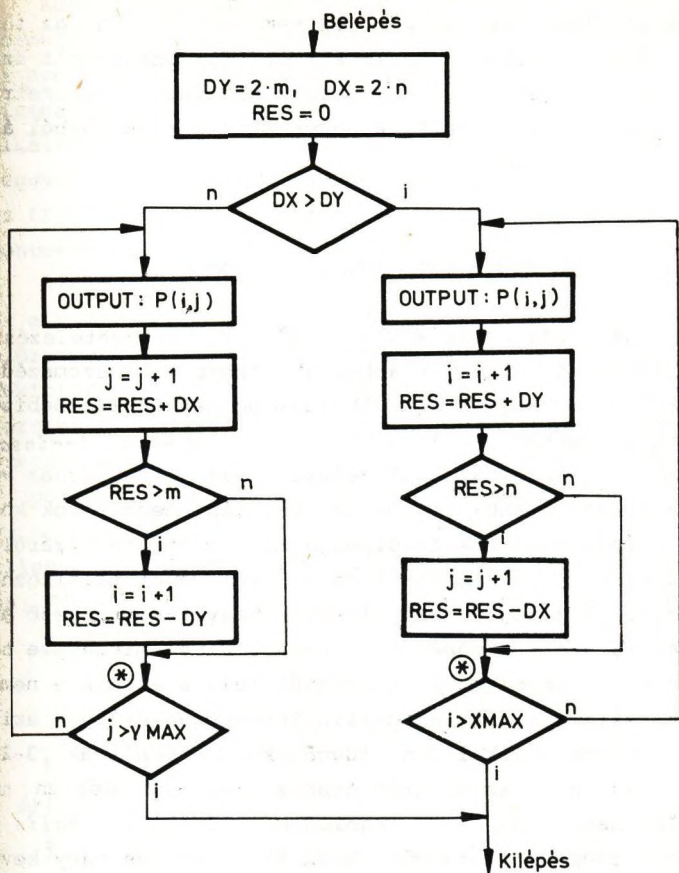
$$y' = y - \frac{RES}{DX} \quad (3-10/a)$$

összefüggés adja.

Feltéve, hogy a meredekség valószínűségi eloszlásának maximuma egynél van, és az eloszlás szimmetrikus, a két ág számításigénye azonos. Ha a 3.3.6.5 alpont algoritmusában egy pontra eső átlagos számításigényt \bar{t}_φ jelöli, akkor a kiterjesztett algoritmussal a $\Delta x \times \Delta y$ méretű felületelem korrekciójának számításigénye (az újra-mintavételezéstől és a kiegészítő műveletektől eltekintve):

$$\bar{t} = [4\Delta y + 2\Delta x(\Delta y + 1)] \bar{t}_\varphi.$$

Egy felületelem korrekciójához a fenti algoritmust megvalósító rutint hat független példányban kell egyidejűleg működtetni. Ebből kettő végzi az aktuális sorban a pozíciók meghatározását, négy pedig az aktuális sor paramétereinek kiszámítására szolgál.



3-8. ábra. Általánosított egyenesgenerálási algoritmus tömbvázlata

Mint már említettük - mivel a sorfolytonos adatkezelést nemcsak lokálisan (egy felületelemen belül), hanem globálisan (a teljes kimenő rendszerbeli képsoron keresztül) is érvényesítenünk kell - egyidejűleg annyi felületelem korrekcióját kell végezni, ahány részre osztot-

tuk a korrigált kép sorait. Ez nem befolyásolja az interpolációhoz szükséges műveletszámot, viszont növeli az algoritmus tárigényét. A globálisan nemlineáris geometriai korrekciót tehát általános négyponos korrekciókból álló *sávval* közelítjük.

3.3.8 UJRA-MINTAVÉTELEZÉS

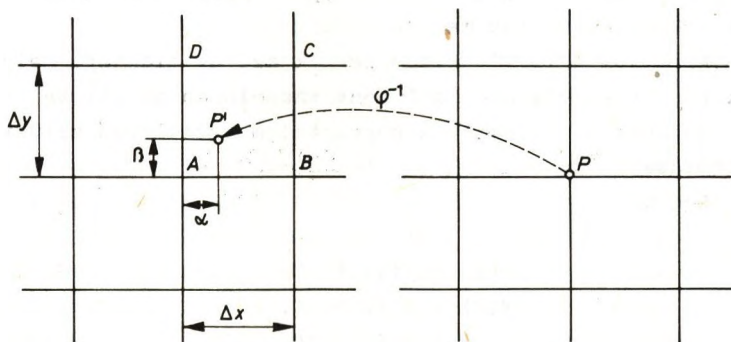
A szakirodalomban a különböző újra-mintavételezési megoldások közt leggyakrabban a legközelebbi szomszéd módszer alkalmazására találhatunk példát [3-3]. Bebizonyítható azonban ([A-14]; 75-83. old.), hogy ez *alul-mintavételezéshez*, s ezzel az alulszelektáltság jelenségéhez vezet.

A bonyolultabb újra-mintavételezési megoldások közül (még elfogadható számításigénye miatt) szinte kizárólag a *bilineáris interpolációs* formulával lehet találkozni (lásd 1.2.3 alpont). Nagy lokális nagyítási tényező esetén ez is hamis eredményt ad [3-1], mivel mindössze négy szomszédos képpontértéket használ fel, s ez eleve nem lehet körszimmetrikus. Speciális interpolációt (nem aritmetikai, hanem logikai összefüggéseket) használnak [3-23]-ban, sajnos ez az eljárás többszintes kép esetén nem alkalmazható. Egy másik megoldásban [3-17] a *lokális nagyítási tényezőtől* függően használnak pontos vagy kevésbé pontos interpolációs formulát.

*Homogén képrészek*en a különböző interpolációs formulák a pontos pozíciótól függetlenül azonos eredményt adnak, mivel az interpolációt közel azonos számértékek között végezzük. Az interpolációs formulák különbözősége a gyorsan változó képrészeken, az élek közelében jelentkezik. Ebből következően a pontosabb formulák alkalmazása csak ezeken a helyeken indokolt.

Keressünk olyan megoldást, amelyben az interpolációs formulák megválasztását a képtartalom vezérli úgy, hogy az eredmény a képtartományban sehol se legyen rosszabb, mintha mindenütt a pontosabb képlettel számoltunk volna. Alkalmazzuk a homogén képrészeken a legegyszerűbb (és leggyorsabb) megoldást, a legközelebbi szomszéd módszert; az [F-3]-ban leírt hibamérték alapján pedig válasszuk pontosabb formulának a bilinéáris interpolációt.

Alapvetően az egyszerűbb megoldást akarjuk használni, és a pontosabb megoldásra csak akkor kerül sor, ha az előre megadott küszöbnél nagyobb hibát követnénk el. Tekintsük a 3-9. ábrát. Rendelje a φ^{-1} inverz értelmű korrekciós összefüggés a kimenő képbeli koordináta-rendszerben felvett P pozícióhoz a bemenő kép rendszerében a P' pozíciót. Az aktuális rácselemet a bemenő kép rendszerében az az egymással szomszédos négy képpont (az ábrán $A, B, C,$ illetve D) határozza meg, amelyek a P' pozíciót közrefogja; P' -nek a legközelebbi rácsponthoz (az ábrán az A pont),



3-9. ábra. Interpolációs paraméterek meghatározása

képest (α, β) a relativ pozíciója. (Legyen $0 \leq \alpha, \beta \leq 0,5$ és $\Delta x = \Delta y = 1$.) Keressük a P pozícióhoz tartozó képfüggvényértéket az A, B, C, D pozícióhoz tartozó $f(A), f(B), f(C), f(D)$ függvényértékek, valamint az (α, β) relativ pozíció függvényeként.

A legközelebbi szomszéd módszer (α és β ilyen megválasztása mellett) egyértelműen az

$$f(P) = f(A)$$

hozzárendelést adja. Ennek kiszámítása a P' pont koordinátáinak egész értékre kerekítését jelenti.

A bilineáris interpolációs formula alkalmazása valós számításokat igényel (vö. (1-23) képlet):

$$f(P) = \beta[\alpha f(C) + (1 - \alpha)f(D)] + \\ + (1 - \beta)[\alpha f(B) + (1 - \alpha)f(A)]. \quad (3-11)$$

(Megjegyzendő, hogy az $f(A), \dots, f(D)$ értékek általában nem sikot határoznak meg.)

Keressük a (3-11) kifejezésnek azt az alakját, melynek független változói a P pont rácselemen belüli relativ pozíciója, valamint a képtartalom rácselemen belüli változása.

Legyen

$$f(B) = f(A) + \Delta B, \\ f(C) = f(A) + \Delta C, \\ f(D) = f(A) + \Delta D.$$

Ekkor a (3-11) kifejezés a következőképpen írható fel:

$$f(P) = f(A) + \alpha(1 - \beta)\Delta B + \beta(1 - \alpha)\Delta D + \alpha\beta\Delta C.$$

A kapott kifejezést átrendezéssel célszerű tovább alakítani:

$$f(P) = \underbrace{f(A)}_{\text{I.}} + \underbrace{\alpha\Delta B + \beta\Delta D}_{\text{II.}} + \underbrace{\alpha\beta(\Delta C - \Delta B - \Delta D)}_{\text{III.}}. \quad (3-12)$$

Ezzel a következő előnyökhöz jutunk:

- szeparáltuk a relatív pozícióktól, valamint a szorzatuktól való függést;
- meg tudjuk határozni azokat a korlátokat, melyek alapján - a szomszédos képpontok világosságkódjának különbségétől függően - döntünk az egyes tagok szükségességéről;
- a szükségesség eldöntéséhez használt részletszámítások eredménye felhasználható a szükséges tagok értékének kiszámításánál.

A (3-12) kifejezés I. tagja azonos a legközelebbi szomszéd értékével. A $0 \leq \alpha, \beta \leq 0,5$ megkötés következtében

$$0 \leq \alpha\beta \leq 0,25,$$

azaz a II., illetve a III. tag elhagyása akkor nem befolyásolja az eredményt, ha tengelyirányban egynél, átlós irányban pedig négnél nem nagyobb a világosságkódok különbségének abszolútértéke. Ha a tengelyirányban elkövethető hibát h_1 , az átlós irányban elkövethetőt pedig h_2 jelöli, akkor a II., illetve a III. tag elhagyásának feltétele:

$$|\Delta B| \leq 2h_1 \text{ és } |\Delta D| \leq 2h_1,$$

illetve

$$|\Delta C - \Delta B - \Delta D| \leq 4h_2.$$

Mivel a III. tag kiszámításához kell a legtöbb idő, célszerű a $h_2 > h_1$ választás.

3.3.9 ADAPTIV GEOMETRIAI KORREKCIÓ

Általános tapasztalat, hogy széles feladatosztályra kisebb hatékonysággal lehet optimalizálni, mint konkrét feladatra. (Pl. az univerzális megoldások az egyszerűbb esetekhez szükségtelen vizsgálatokat tartalmaznak.)

Ésszerű ezért a következő szervezés:

- a) állítsunk össze a feladatok szűkebb körére optimalizált programokból olyan programkönyvtárat, amely minden szóba jövő feladathoz tartalmaz hatékony megoldó programot;
- b) adott feladat megoldásánál egy elemző modul
 - döntse el, hogy a rendelkezésre álló erőforrások figyelembevételével melyik program alkalmazása az optimális;
 - jelölje ki, paraméterezze és futtassa le a kiválasztott feladatmegoldó programot.

Geometriai korrekció esetében a feladatelemző modul pl. a következő szempontok figyelembevételével dönthet:

- a kijelölt korrekció lineáris-e vagy bonyolultabb;
- egyszerűbb vagy pontosabb újra-mintavételezésre van-e szükség;
- a korrekciós adathalmaz szűkített-e vagy teljes.

3.4 A KORREKCIÓS ÖSSZEFÜGGÉS MEGHATÁROZÁSA

A geometriai korrekciós összefüggéseket kétféle módszerrel határozhatjuk meg:

- *Elméleti uton*: ha ismerjük a függvénykapcsolatot a bemenő és a kimenő koordináta-rendszer között; (pl. le tudjuk írni a képfellevő rendszer torzításait, adva vannak az ábrázolt objektum elmozdulását leíró összefüggések stb.); továbbá ismerjük ezen függvénykapcsolatok összes paraméterét.
- *Kísérleti uton*: ha - mint általában az esetek nagy részében - nem áll rendelkezésünkre egzakt geometriai korrekciós összefüggés. Ekkor a korrekciót olyan pontok pozíciójának megadásával definiálhatjuk, amelyeknek mind a bemenő, mind a kimenő képen *ugyanaz a jelentésük*; ezek az *azonosítási pontok* (GCP's: ground control points). Mivel ezeket az adatokat legalább az egyik kép koordinátarendszerében mérésekből kapjuk, *korlátozott pontosságúak*. (Még a feltűnő képi alakzatok kiemelt pontjai sem azonosíthatók eléggé egyértelműen, másrészt a pontosságuknak elvi korlátot szab a pozíció kvantálása.)

Az elméleti és a kísérleti mód *keveredhet* egymással: pl.: az elméleti uton meghatározott korrekciós összefüggést paraméterezhetjük azonosítási pontok felhasználásával is.

Ezek után észszerű, hogy a korrekciós összefüggést a következő *feltételek* alapján keressük:

- az azonosítási pontok összerendelésénél elkövetett hiba legyen minimális;
- a meghatározott korrekciós összefüggés a teljes képtartományban a lehető "legsímábban" viselkedjék;

- az eleve szükségesnél nagyobb számú azonosítási pont megadásával a korrekció legyen pontosítható.

A korrekciós összefüggést előállító *algoritmussal* kapcsolatos igények:

- az azonosítási pontok pozícióin kívül ne legyen szükség más összefüggés ismeretére, de ha ilyenek rendelkezésre állnak, azokat fel lehessen használni (pl. elméleti uton meghatározott korrekció paraméterezése azonosítási pontok segítségével);
- egyaránt tartalmazzon egyszerűbb és bonyolultabb közelítéseket, és a szükséges bonyolultságot elő lehessen írni;
- az azonosítási pontok összerendelésénél elkövetett hiba nagy számú azonosítási pont esetén is eléggé gyorsan konvergáljon zérushoz;
- legyen lehetőség az elért pontosság, illetve az elkövetett hiba kvantitatív ellenőrzésére.

Legyen N pontból álló azonosítási ponthalmazunk, amely az (A) és (B) koordináta-rendszer között kölcsönösen egyértelmű leképezést definiál:

$$P_s: (x_{sA}, y_{sA}) \rightarrow (x_{sB}, y_{sB}) \quad (s = 1, 2, \dots, N); \quad (3-13)$$

ahol P_s az s -edik azonosítási pont; $x_{sA}, y_{sA}, x_{sB}, y_{sB}$ pedig az (A), illetve (B) rendszerbeli pozíciók koordinátái. Mivel ezeket mérésekből ismerjük, tegyük fel, hogy az elkövetett hibák normális eloszlásúak, zérus várható értékkel.

Keressük külön az

$$f_x: \{(x_{sA}, y_{sA}) \mid s = 1, 2, \dots, N\} \rightarrow x_{sB}$$

és az

$$f_y: \{(x_{sA}, y_{sA}) \mid s = 1, 2, \dots, N\} \rightarrow y_{sB}$$

folytonos függvényeket, amelyek a megfeleltetést a lehető legkisebb hibával valószínűsítik meg a P_s ($s = 1, 2, \dots, N$) pontokban, és elegendően "simák". (Igy az elkövetett hiba nem izotrop, viszont a szeparálás jelentősen csökkenti a számítási időt.)

Az [F-3]-ban a hozzárendelést Kálmán-szűrő felhasználásával állapítják meg. Célszerűbb azonban az a megoldás, amelyben a hozzárendelő függvények típusa szabadon megválasztható. (A képfeldolgozás területén fellépő geometriai transzformációs, illesztési, illetve torzítási probléma többnyire ismert függvénykapcsolatként jelentkezik, csak a konkrét paraméterek ismeretlenek.)

Keressük ezért a megoldást megválasztható függvények lineáris kombinációjaként. Tételezzük fel, hogy a keresett f_x és f_y függvény közelíthető m számú lineárisan független folytonos függvény összegével:

$$f(x_{sA}, y_{sA}) = \sum_{i=1}^m a_i g_i(x_{sA}, y_{sA}),$$

ahol f vagy az f_x vagy az f_y függvényt közelíti. (A továbbiakban az egyszerűség kedvéért elhagyjuk az argumentumokat.) Ekkor az azonosítási pontokra felírhatjuk, hogy

$$f = \sum_{i=1}^m a_i g_i, \quad (3-14)$$

ahol az f , illetve a g_i (N elemű) oszlopvektorok elemei az f , illetve a g_i függvények helyettesítési értékei az (x_{sA}, y_{sA}) pontokban (az előbbiek nyilván az x_{sB} vagy az y_{sB} függő változó közelítő értékei); a_i pedig az alkalmasan megválasztott \underline{a} együtthatóvektornak a g_i függvényhez tartozó eleme.

A korrekciós összefüggés keresésének *első* (pontosági) feltétele kielégíthető, ha az eltéréseket megfelelő módon leíró

$$D(\underline{a}) = \sum_{s=1}^N \varphi_s D_s(\underline{a}) \quad (3-15)$$

funkcionált (lásd F4.1.1 alpont) *minimalizáljuk* a_1, a_2, \dots, a_m szerint. A képletben φ_s a súlytényezőket jelöli; és pl.

$$D_s(\underline{a}) = (f(x_{sA}, y_{sA}) - \sum_{i=1}^m a_i g_i(x_{sA}, y_{sA}))^2 \quad (3-16)$$

lehet. A φ_s súlyokat célszerű a P_s pontok eloszlási sűrűségétől függően választani: ahol a P_s pontok sűrűbben helyezkednek el, ott φ_s legyen kisebb.

A korrekciós összefüggés keresésének *második* (simasági) feltétele értelmében célszerű az $m \ll N$ választás. Ekkor ui. az f függvénynek viszonylag kevés paramétere van, és megfelelő g_i függvények esetén várható, hogy a közelítés a mérési pontok között is javul.

A simaságot kellően választott *regularizációval* is elő lehet segíteni. Ekkor az a_i együtthatókat $D(\underline{a})$ helyett $D'(\lambda, \underline{a})$ minimumából lehet meghatározni a

$$\min_{\underline{a}} [D'(\lambda, \underline{a})] = \min_{\underline{a}} [D(\underline{a}) + \lambda \Psi(\underline{a})]$$

összefüggés szerint. Itt a $\Psi(\underline{a})$ funkcionál a függvény simaságát fejezi ki. A λ paraméter megválasztásától függ, hogy a közelítés pontosabb lesz-e a megadott pontokban vagy simább.

Keressük a megoldást $\varphi_s = 1$ ($s = 1, \dots, N$) és $\lambda = 0$ esetre; a közelítés pontosságát és simaságát pedig a P_s pontok egyenletes elosztásával és sima g_i függvények használatával segítsük elő. Ekkor a (3-15) a következőképpen alakítható át:

$$D(\underline{a}) = \underline{f}^T \underline{f} - 2 \sum_{i=1}^m a_i (\underline{f}^T \underline{g}_i) + \sum_{i=1}^m \sum_{j=1}^m a_i a_j (\underline{g}_i^T \underline{g}_j), \quad (3-17)$$

ahol a \top felső index a transzponált vektort jelöli.

Vezessük be a

$$\underline{z} = [\underline{f}^T \underline{g}_i],$$

$m \times 1$ -es oszlopvektort, illetve a

$$\underline{G} = [\underline{g}_i^T \underline{g}_j] \quad (i, j = 1, \dots, m)$$

$m \times m$ -es mátrixot; ezekkel (3-17) a

$$D(\underline{a}) = \underline{f}^T \underline{f} - 2 \underline{a}^T \underline{z} + \underline{a}^T \underline{G} \underline{a} \quad (3-18)$$

alakba írható.

A minimalizáláshoz fejtsük sorba $D(\underline{a})$ -t tetszőleges m -dimenziós térbeli \underline{a}^* pont körül:

$$D(\underline{a}) = D(\underline{a}^*) + \left[\frac{\partial D(\underline{a})}{\partial \underline{a}} \right]_{\underline{a}=\underline{a}^*} (\underline{a} - \underline{a}^*) +$$

$$+ \frac{1}{2} (\underline{a} - \underline{a}^*)^\top \frac{\partial}{\partial \underline{a}} \left[\frac{\partial D(\underline{a})}{\partial \underline{a}} \right]_{\underline{a}=\underline{a}^*} (\underline{a} - \underline{a}^*) + \dots$$

ahol a parciális deriváltak (3-18) alapján:

$$\left[\frac{\partial D(\underline{a})}{\partial \underline{a}} \right] = -2\underline{z}^\top + 2\underline{a}^\top \underline{G};$$

$$\frac{\partial}{\partial \underline{a}} \left[\frac{\partial D(\underline{a})}{\partial \underline{a}} \right] = 2\underline{G};$$

és - mivel \underline{G} nem függ (\underline{a}) -tól - az összes magasabbrendű parciális derivált eltűnik. Így (3-18)-at végül a

$$D(\underline{a}) = D(\underline{a}^*) + (-2\underline{z}^\top + 2\underline{a}^{*\top} \underline{G}) (\underline{a} - \underline{a}^*) +$$

$$+ \frac{1}{2} [(\underline{a} - \underline{a}^*)^\top 2\underline{G} (\underline{a} - \underline{a}^*)]$$

alakra hoztuk. A $D(\underline{a})$ függvényre előírt szélsőérték teljesülésének feltétele (3-15) értelmében az, hogy

$$\left[\frac{\partial D(\underline{a})}{\partial \underline{a}} \right] = -2\underline{z}^\top + 2\underline{a}^\top \underline{G} = 0$$

legyen, amiből az $\underline{a}^\top \underline{G} = \underline{z}^\top$, vagyis a

$$\sum_{j=1}^m a_j \underline{g}_i^\top \underline{g}_j = \underline{f}^\top \underline{g}_i \quad (i = 1, \dots, m) \quad (3-19)$$

végeredmény adódik.

A közelítés hibája is könnyen kifejezhető:

$$E(\underline{a}) = \underline{f}^\top \underline{f} - \sum_{i=1}^m \underline{f}^\top \underline{g}_i. \quad (3-20)$$

A leirt megoldás műveletigénye erősen függ m -től, mivel (3-19) szerint $- m$ közelítő függvény alkalmazása esetén $- m \times m$ -es mátrixot kell invertálnunk. Így $-$ nagy m esetén $-$ a g_i bázisfüggvények cserélgetése, illetve a közelítésben részt vevő függvények számának változtatása nehézkes, lassu.

A probléma áthidalására használjuk fel az m tagu és az $(m + 1)$ tagu közelítés közti rekurzív kapcsolatot [3-13]. Ezen kívül az adott m tagu közelítésben részt vevő függvényeket egy $M \gg m$ függvényt tartalmazó *bázis-függvény-könyvtár*ból válasszuk ki, az adott választáshoz tartozó $E(\underline{a})$ hiba minimális értékének megfelelően. Így egyrészt nem kell előre meghatározni a közelítő függvény tagszámát, másrészt minden lépésnél a legjobb eredményt adó bázisfüggvények kerülnek felhasználásra.

A rekurzív algoritmustól megkivánjuk, hogy az $(m + 1)$ tagu közelítés számításainak nagy része támaszkodjék az m tagu számítások részeredményeire. (Természetesen a szám-ábrázolás véges pontossága és a kerekítési hibák miatt a magasabb tagszámú közelítés alkalmazási lehetősége korlátozott.) Az algoritmustól elvárjuk azt is, hogy a hibát jól kézben lehessen tartani.

Írjuk fel a (3-19) egyenlet felhasználásával a különböző tagszámú közelítéseket.

$$\boxed{m = 1} \quad (g_1^T g_1) a_{11} = f^T g_1;$$

$$a_{11} = \frac{f^T g_1}{g_1^T g_1}.$$

Bevezetve a $B_{11} = g_1^T g_1$ jelölést:

$$a_{11} = \frac{f^T g_1}{B_{11}}.$$

A bázisfüggvénykönyvtár M függvényéből az lesz a választott g_1 függvény, melyre a (3-20) hiba minimális, az $m = 2$ szinten pedig a maradék $M - 1$ számú függvény közül lehet választani.

$$\boxed{m = 2} \quad \begin{bmatrix} g_1^T g_1 & g_1^T g_2 \\ g_2^T g_1 & g_2^T g_2 \end{bmatrix} \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} = \begin{bmatrix} f^T g_1 \\ f^T g_2 \end{bmatrix}$$

A B_{ij} jelölés bevezetésével a megoldás:

$$B_{12} = \frac{g_1^T g_2}{B_{11}}, \quad B_{21} = g_2^T g_1,$$

$$B_{22} = g_2^T g_2 - B_{21} B_{12},$$

$$a_{22} = \frac{f^T q_2 - a_{11} B_{21}}{B_{22}}, \quad a_{21} = a_{11} - B_{12} a_{22}.$$

Látható, hogy az $m = 1$ szintről a_{11} és B_{11} felhasználható.

Általánosítva, a rekurzív algoritmus a következő:

$$m \geq 2$$

$$\left. \begin{aligned} B_{i1} &= q_i^T q_1; & B_{ji} &= \frac{B_{ij}}{B_{jj}}; \\ a_{ii} &= \frac{f^T q_i - \sum_{n=1}^{i-1} a_{nn} B_{in}}{B_{ii}}; \\ a_{i, i-j} &= a_{i-j, i-j} - \sum_{n=1}^j a_{i, i-j+n} B_{i-j, i-j+n}; \end{aligned} \right\} \begin{aligned} &(i = 2, \dots, m); \\ &(j = 1, \dots, i-1); \end{aligned}$$

$$B_{ij} = q_i^T q_j - \sum_{n=1}^{i-1} B_{nj} B_{in}; \quad \begin{aligned} &(i = 2, \dots, m); \\ &(j = 2, \dots, i). \end{aligned}$$

Az adott m szinten a korábbi szinteken fel nem használt $M - m + 1$ függvény közül a (3-20) feltételek szerint optimális függvényt választjuk.

3.5 GYAKORLATI MEGOLDÁSOK

A digitális képek geometriai korrekciójának megvalósítására a szakirodalomban számos gyakorlati megoldás található. Ezek többsége a diszkrét egyenes ábrázolásá-

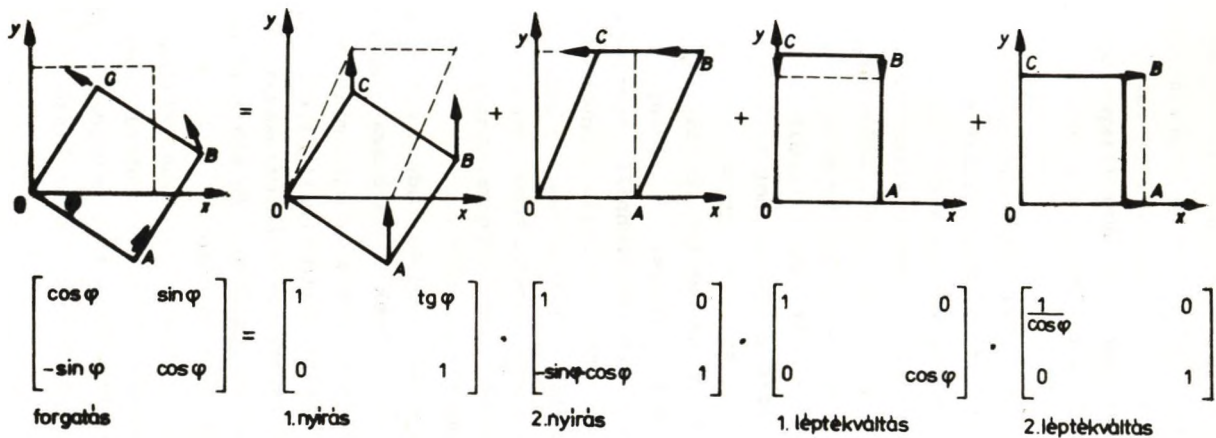
nak, illetve létrehozásának (lásd 3.3.6 alpont) törvényszerűségein alapul. A következő alpontban áttekintünk és értékelünk az irodalomban található néhány megoldást az általános, nemlineáris geometriai korrekcióra. Végül a 3.5.2 alpontban ismertetjük az SZKI-ban elért eredményeket.

3.5.1 IRODALMI ÁTTEKINTÉS

A számítógépes grafika, illetve a számítógéppel segített tervezés (CAD) területén a geometriai transzformációkat elsősorban mátrixformalizmussal fogalmazzák meg ([3-18], [F-10]). Az egységes tárgyalás és adatkezelés érdekében homogén koordinátás reprezentációt alkalmaznak (lásd 3.2.2 alpont). Az általános perspektív transzformáció homogén koordinátás reprezentációjának (két- és háromdimenziós esetben, mind direkt, mind inverz megvalósításban) korrekt vizsgálata található [3-8]-ban, de a számítási igény elemzése nélkül. Valójában a mátrixformalizmus hátrányos a digitális képek transzformációjánál, mivel az egyes képpontok koordinátáinak kiszámításához sok műveletet kell végezni.

A mátrixműveletek felbonthatóságát kihasználva a lineáris transzformációkat vissza lehet vezetni kizárólag a koordinátatengelyekkel párhuzamos nyírásra és léptékváltásra [3-24].

Példaként tekintsük az origó középpontu, φ szögű elforgatás felbontását (lásd 3-10. ábra). Az eljárás a forgatást két nyírásra és két léptékváltásra bontja. A leírt megoldásban alkalmazott újra-mintavételezési eljárás (több szomszédos képpont értékének összeadása, majd osztás a tagok számával) megfelelő vizuális hatást kelt, és elkerüli az *aliasing* (aliasing) problémáját,



3-10. ábra. Forgatás visszavezetése nyírásra és léptékváltásra

de erősen módosítja a kép statisztikáját és jelentős hibákat okoz. Másrészt az így végrehajtható korrekció túl speciális, nemlineáris esetre való általánosítása pedig nehézkes. Ezért előnyei (fixpontos műveletek alkalmazása) ellenére általános módszerként nem jöhet szóba.

Sík képek perspektív transzformációját tárgyalják [3-17]-ben. Speciális dekompozíció segítségével a korrekciót három lépésre bontják: előbb *véttdimenziós forgatásos eltolást* alkalmaznak szűrés nélkül, majd *kétdimenziós projektív transzformációt szűréssel*, és végül újra *szűrés nélküli forgatásos eltolást*. Céljuk: a lehető legkevesebb számításal elkerülni az alulszelektált-ságot. Ennek érdekében a képből (interaktívan) megállapítják a felül-mintavételezés mértékét, és ebből a lokális nagyításra vonatkozó küszöbszámot számolnak. Ettől függően döntenek el, hogy adott helyen szükség van-e szűrésre vagy sem. Számítástechnikai megfontolásból az inverz irányú korrekciót tartják előnyösebbnek. Az alkalmazott felbontás általános esetben nem használható, viszont jó a szűrésre vonatkozó *alapötlet* (ti. az, hogy a precíz megoldást lehetőség szerint csak a szükséges helyeken alkalmazzák).

Mások speciális esetekre adnak gyors megoldást. *Kizárólag bináris* képek skálázására alkalmazható a [3-23]-ban leírt eljárás; az újra-mintavételezéshez szükséges interpolációt aritmetikai jellegű számítások helyett *logikai függvények* alkalmazásával oldják meg. Sajnos, a nagy variációs lehetőség miatt többszintes képek esetén ez az út nem járható.

Hatékony *forgatási algoritmus* található [3-6]-ban, amelynek az a korlátja, hogy csak bizonyos forgatási szögekre működik. Az alkalmazott megoldás a koordináták szögfüggvénnyel (azaz előzetesen kiszámított konstanssal

való megszorzása helyett megfelelő számú *léptetést és összeadást* tartalmaz. Az egyes képpontok pozícióját egymástól függetlenül számolják ki.

Elsősorban számítógépes grafikai megközelítésben javasolt a görbült vonalak, illetve görbült felületek ábrázolása *szakaszosan értelmezett polinomokkal* (spline-okkal). (E témakörrel jó áttekintés található pl. [F-10]-ben.) A módszer - kétségtelen előnyei és eleganciája ellenére - geometriai korrekciók esetében *gazdaságtalan*, mivel minden képpont pozíciójának kiszámításához tetemes mennyiségű, valós számokkal végzendő műveletet igényel.

Ugyancsak nem gazdaságosak az [F-5]-ben, illetve a [3-14]-ben található módszerek sem. Ezek a legfeljebb másodfoku polinommal megadható görbék generálását kizárólag fixpontos aritmetikai műveletekkel oldják meg. Képek geometriai korrekciója esetén ehhez külön számítás-igényként járulna a pontok vonalmenti eloszlásának, valamint a pontos pozíció kiszámításához szükséges korrekciók tagnak a meghatározása.

Igen hatékony a [3-3]-ban található megoldás. Gyors egyenesgenerálási algoritmus a lényegében azonos a 3.3.6.5 alpontbeli algoritmussal; újra-mintavételezéshez pedig a *legközelebbi szomszéd* módszert alkalmazza. Mivel direkt irányú korrekciót hajt végre (azaz a bemenő kép pozícióihoz határozza meg a kimenő kép rendszerében a megfelelő pozíciókat), nagyítás esetén elkerülhetetlen az *alulszelektáltság*. Ezt *képpontismétléssel* kompenzálják, ami nem korrekt, de gyors és egyes esetekben kielégítő megoldás. A leírás a lineáris korrekciókat tárgyalja részletesen. Az eljárás alkalmazása forgatást, eltolást, kicsinyítést, illetve kismértékű nagyítást tartalmazó korrekciókra javasolt; s mivel fixpontos aritmetikai műveletekre épül, működése gyors. (Pl. egy 512×512 pontból álló kép elforgatásához TPA-1140 számítógépen

- leszámítva a háttértárkezelés idejét - kb. 40 s szükséges, szemben a valós számokkal végzett mátrixszorzás kb. 45 perces időigényével - lebegőpontos számológépi használata nélkül).

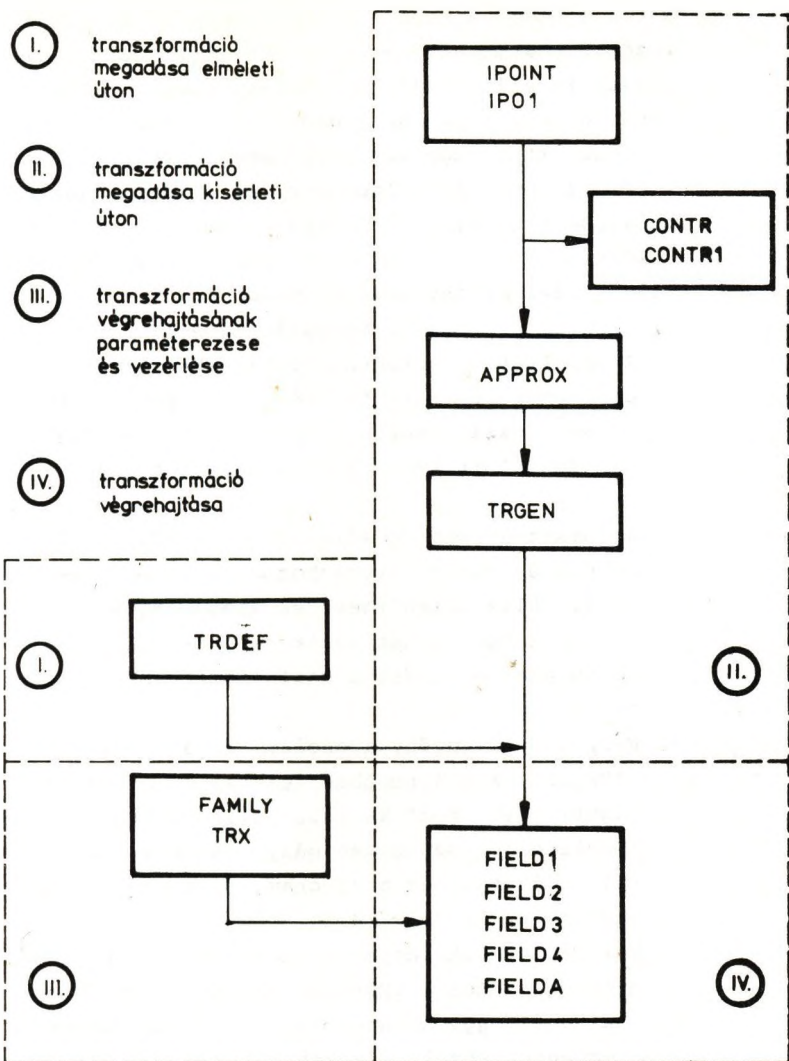
Az általános négyponthoz való korrekció megvalósításához szükséges paraméterek levezetésénél megállapítják, hogy a nemlineáris korrekció az affin korrekció nemlineáris taggal való kiegészítésével lenne megoldható, erre azonban nem adnak érdemi javaslatot.

3.5.2 MEGVALÓSÍTÁSI EREDMÉNYEK

A geometriai korrekciós algoritmusokat és a szükséges segédprogramokat az SZKI-ban a MIP rendszer részeként valósítottuk meg [1-31]. A fejlesztés TPA-1140 kiszámítógépen, RSX-11M V3.2 operációs rendszer alatt történt; továbbfejlesztésként több MIP-adaptáció (TPA-1148, SZM-4) született. A geometriai korrekciós programok a MIP rendszernek kb. tízed részét alkotják (kb. 15000 programsor, többsége assembly nyelven). A képi megjelenítés az SZKI-ban kifejlesztett képfeldolgozó berendezés segítségével történik ([1-9], [1-13]).

A különböző, jól behatárolható feladatokat különálló programmodulok oldják meg, amelyek közt a kapcsolat adat-szinten automatikus (lásd 3-11. ábra).

1. Az elméleti uton meghatározott geometriai korrekciók a TRDEF program segítségével adhatók meg. Explicit módon legfeljebb ötödfoku racionális törtefüggvénnyel leírható, inverz értelmű korrekciók paraméterezhetők. A program a lineáris korrekció paramétereit implicit módon (forgatási szög, skálázási tényező stb.) is elfogadja. Bemérő adatok még a felületelemes közelítés adatai és



3-11. ábra. Kapcsolatok a geometriai korrekciós programok között

- a kép adatai. Kimenő adat a geometriai korrekciót vezérlő adatok hálózata, amely szükség esetén vizuálisan is ellenőrizhető. (A 14. képen a kimenő rendszerbeli képablak megadott rácsállandóval létrehozott ösképe látható a bemenő rendszerben kijelölt képablak belsejében. Az ábrázolt korrekciós összefüggés nemlineáris és nem invertálható.)
2. A korrekciót kísérleti uton több program egymás utáni alkalmazásával lehet megadni:
- Azonosítási pontok meghatározása (IPO1, IPOINT). A két program legfeljebb 255 db pozíció numerikus, illetve képi (fénypontos) megadását teszi lehetővé. (Lásd 15. kép: a fényponttal megjelölt képpont világosságkódja leolvasható.)
 - Azonosítási pontok ellenőrzése (CONTR, CONTR1). Lehetőség van az egymáshoz rendelések numerikus és vizuális ellenőrzésére. (Lásd 16. kép: a megadott pozíciókat szálkeresztek jelzik, elmozdulásaikat pedig a szálkeresztekből kiinduló szakaszok.)
 - Folytonos függvénykapcsolat meghatározása (APPROX). A 3.4 pontban leírtaknak megfelelő algoritmust ötödfoku racionális törtfüggvényes közelítéshez szükséges bázisfüggvény-könyvtárral valósítja meg a program. A közelítés összetettsége vezérelhető.
 - Vezérlő adatok létrehozása (TRGEN). A program szerepe azonos a TRDEF programéval úgy, hogy a kívánt függvénykapcsolat az APPROX program eredményei közül választható.
3. A FAMILY és TRX program segítségével a feladatok meghatározása és a végrehajtásuk szétválasztható. Így lehetőség van a geometriai korrekciók

kötegetelt jellegű végrehajtására. A TRX vezérlő program a felhasználói paraméterek elemzésével a 3.3.9 alpontban leírt szervező tevékenységet is ellátja.

4. A korrekciót ténylegesen végrehajtó programok (FIELD1...4, FIELDA) a 3.3.2-3.3.8 alpontban leírt algoritmusokat valósítják meg. Az egyes programok egy-egy szűkebb feladatosztályra optimálisak. A FIELD1-4 program képek geometriai korrekciójára, a FIELDA program a korrekciós adathalmaz szűkítéséhez szükséges *másodrendű kép* (lásd [1-31]) transzformációjára szolgál. A transzformációk leállíthatók és újraindíthatók. A programok tetszőleges képmérettel működnek, a korlátokat a háttértárolók kapacitása szabja meg.

A háttértárhoz fordulások számától, illetve - bilineáris interpolációt alkalmazó ujramintavételezés esetén - a kép összetettségétől függően egy 256×256 képpontból álló kép geometriai korrekciója kb. 40 s, 1024×1024 képpontból álló képé kb. 8-10 perc alatt készül el.

A korrekció végrehajtási ideje közel független a korrekciós összefüggés bonyolultságától, és a pontonkénti végrehajtásnál több nagyságrenddel kisebb, gyakorlatilag azonos eredmény mellett.

A 16. képen kitűzött azonosítási pontok felhasználásával példaképpen illesztettük a 17/a képet a 17/b kép szerinti felvételhez. A transzformáció eredménye a 17/c képen látható.

A felületelemes közelítés és az alkalmazott újra-mintavételezési eljárás eredményeképpen az erősen nemlineáris korrekciók is nagy pontossággal valósíthatók meg. Példaképpen egy szabályos négy-

zetrácsot transzformáltunk a 14. kép vezérlő adataival; az eredményt a 18. képen adjuk meg.

A bemutatott programrendszer futtatása során szerzett tapasztalataink alapján megállapíthatjuk, hogy a digitális képek geometriai korrekcióit egyrészt nagy pontossággal lehet megvalósítani még bonyolult függvénykapcsolattal leírható összefüggések esetén is; másrészt megfelelően választott gyors algoritmusokkal kisszámítógépen is el lehet érni elfogadható futási időket, speciális hardvereszközök nélkül. Az alkalmazott algoritmusok ([3-10], [3-11], [3-12]) segítségével pl. képpontonként 3-5 fixpontos műveletre sikerült leszorítani a korrekció átlagos műveletigényét! Ez a korábbi eredményeinkhez ([3-20], [3-21], [3-22]) képest több nagyságrend (!) gyorsítást jelent, és lényegesen már csak új elvű hardvereszközök (pl. tömbprocesszor stb.) igénybevételével javítható.

4. SZEGMENTÁLÁS

Egy digitális kép szegmentálásán a képpontok valamilyen sajátságvektor(ok) szerinti osztályozását, majd a kapott osztályozásra nézve *összefüggő* képponthalmazok, azaz *tartományok* (lásd F3.1.3 alpont) *megkeresését* értjük.

Szegmentálásról csak a *fizikai* (jel) *szintű* feldolgozás esetén beszélünk (lásd 1-1. és 1-2. ábra), bár az itt alkalmazott módszerek sok hasonlóságot mutatnak az alakfelismerés során használtakkal. Az alapelvük az, hogy a digitális kép végeredményben a jelenetben szereplő - valamilyen értelemben rendezett - *valós világ* *leképezése*, valamilyen *meghatározott szabályok* szerint. Így elvben lehetséges, hogy a szegmentálással nyert objektumok a valóságos tárgyakat határoló sík vagy görbült felületeknek vagy ezek metszészímvonalainak, illetve a texturaelemeknek síkbeli leképezései lesznek; annak ellenére, hogy a sajátságvektorok kialakításánál ezekre tulajdonképpen nem voltunk tekintettel.

Gondot jelenthet a képekben található *redundancia*, különösen többsávos képek esetében. Kiküszöbölésére különböző transzformációkat és rekurzív eljárásokat alkalmaznak. Ha viszont a szegmentálást transzformált adaton végezzük, ez meghamisíthatja az eredményt (tartományok elveszhetnek, vagy hamisan jelenhetnek meg), ezért a transzformációs eljárásokat gondosan kell kiválasztani és az eredményeket kritikával kell értékelni.

Az alapvető probléma a digitális képből *hiányzó valódi* (térbeli) *harmadik dimenzióra* vonatkozó információk kinyerése, mivel a leképezési transzformáció inverze általában nem egyértelmű. (Gondoljuk meg, hogy két, különbözően megvilágított, de különböző visszaverőképességű, ezért azonos fénysűrűségűnek látszó felület (pl. a felhőzet és a hótakaró képe egy meteorológiai ürfelvételen) *egy* homogén tartományt eredményezhet; vagy különböző térbeli helyzetű élek képe egy egyenesbe eshet stb.)

Az erre vonatkozó tudásunk ez idő szerint csekély. A kísérletek fő célja rendszerint a felületek irányított-ságának meghatározása, pl. a világosságkódok gradienséből [4-35] vagy az árnyékhatásokból ([4-5], [4-24]). Egyes eljárásokban különféle kiegészítő információkat használnak (pl. a tárgyak lehetséges helyzeteinek és tulajdonságainak leírását [4-56], hasonló képek elemzésével nyert tudásbázist [4-33] stb.). A térbeliséget is tekintetbe vevő eljárásokkal a továbbiakban nem foglalkozunk.

4.1 ELVI ÁTTEKINTÉS

A szegmentálás folyamatát 3 lépésre lehet bontani.

1. A *definíciós* lépésben meg kell határozni az osztályozáshoz szükséges paramétereket. Ezek a következők:

- a) *Sajátságvektorok*. Legelőször is elő kell állítani az egyes képpontok sajáttságvektorát. A szegmentáláshoz rendszerint 1, de legfeljebb 3-4 sajáttságot veszünk figyelembe, így a sajáttságtér sokszor csak egydimenziós. Legtöbbször a világosság-, illetve szinkódokat elemezzük, de gyakran használjuk a hisztogramot (lásd 2.2.1 alpont) vagy a gradienst (lásd F4.2.1 al-

- pont) is. Esetenként más, a konkrét módszerhez illeszkedő jellemzőkkel (pl. súlyozott átlagok, rangszámok stb.) számolunk.
- b) *Távolságfüggvény*. A döntés alapjául a képpontok *szegmentálási távolsága* - a továbbiakban röviden *s-távolsága* - szolgál. Matematikailag ez a sajáttságvektorok végpontjának távolsága az n -dimenziós sajátstágtérben, amelyet valamilyen - rendszerint előre megadott - távolságfüggvénnyel (metrika; lásd F3.2.1 alpont) mérünk. Ebben a tekintetben a szegmentálás sok hasonlóságot mutat a távolságméréseken alapuló alakfelismeréssel (lásd 5.2.4 alpont). Megjegyezzük, hogy bár az alkalmazott metrika tetszőleges, megválasztásánál körültekintően kell eljárni, mert döntően befolyásolja a szegmentálás "jóságát".
- c) *Döntésfüggvény(ek)*. Az osztályozást legtöbbször valamilyen (konstans) küszöb(ök) szerint véghezvük; bár esetenként bonyolultabb összefüggések kiértékelésével hozunk döntést. A küszöb(ök)et vagy előre felvesszük tapasztalati alapon, vagy a szegmentálási eljárás során automatikusan határozzuk meg. Mindkét esetben megkülönböztetünk statikus, illetve dinamikus küszöb(ök)et; az előbbi esetben a küszöb az egész képre állandó, az utóbbiban képrészekenként - szélső esetben képpontonként - változik.
- d) *Osztályjellemzők*. Az osztályok számát általában előre megadják; ezen kívül *névazonosító* és *sajátstílus* kód is megadható. A szegmentálás során keletkezett tartományok felveszik osztályuk nevét és (a képvágás során) rendszerint a sajátstílusát is. (Az utóbbi nem szükséges a további

feldolgozáshoz; elegendő megfelelően feltölteni egy Q-táblát; lásd 2.2.3 alpont.)

e) *Tanítók.* Az osztályokat *tanítókkal*, vagyis az osztályt jellemző minta-sajátságvektorokkal is meghatározhatjuk. Ez esetben minden osztályhoz legalább egy tanítót meg kell adni.

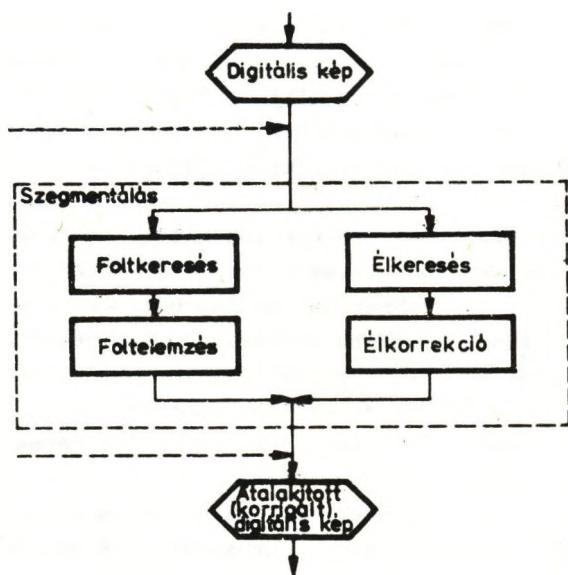
2. Az *osztályozási* lépésben rendeljük hozzá a képpontokat az egyes osztályokhoz. A döntési kritérium az "*elég kicsi s-távolság*", amit igen sokszor a megadott vagy kiszámított küszöb(ök) alapján, esetenként pedig matematikai minimumszámítással ítélünk meg. Más szóval: két képpontot akkor sorolunk ugyanabba az osztályba, ha

- s-távolságuk elég kicsi, illetve
- az adott osztály (egyik) tanítójától elég kis s-távolságra vannak.

(Ha léteznek tanítók, és egy képpont többitől van elég kis s-távolságra, abba az osztályba soroljuk, amelyiknek tanítójától a legkisebb a távolsága. Ha többitől van egyenlő távolságra, bármelyikükbe besorolható.) A mérési hibák, az ellentmondó követelmények vagy az algoritmusok közelítő jellege miatt gyakran előfordul, hogy egyes képpontokat az első menetben nem tudunk besorolni egyik osztályba sem. Kedvelt módszer ezeket külön összegyűjteni, majd egy második menetben a környezetük elemzésével többségi alapon vagy valamilyen súlyfüggvény szerint besorolni.

3. Végül az *összefüzési* lépésben megkeressük az egyes osztályokba sorolt képpontok összefüggő halmazait, a tartományokat. Emlékeztetünk rá, hogy a tartományok ekvivalencia-osztályokat alkotnak (lásd F3.1.3 alpont).

A szegmentálással *foltokhoz*, illetve *élekhez* jutunk (lásd 4-1. ábra), attól függően, hogy a figyelembe vett sajátságok *hasonlósági*, illetve *különbözőségi* jellemzőket mérnek-e. Ideális esetben ezek egybeesnek a képen található objektumok határoló felületeivel, illetve konturvonalalaikkal. Mindkét feladat megoldására nagyon sok módszer létezik, amelyek sokszor még céljukat tekintve is erősen különböznek [1-26]. Teljes körű áttekintésük nemcsak számuk miatt lehetetlen: arról van szó, hogy - az elméleti alapok már többször említett hiánya miatt - egy-egy konkrét feladatkör kapcsán alakultak ki, és ritkán vizsgálták az általános érvényességüket. Ezért az anyagunkat is a módszerek szerint rendeztük.



4-1. ábra. A szegmentálás modellje.

A 4.2 pontban globális, lokális, illetve hibrid *foltkereső* módszereket ismertetünk. A globális módszerek a geometriai összefüggőség, a lokálisak a határvonalak "simasága" tekintetében hagynak kívánnivalókat; előnyüket a hibrid módszerek kísérik meg egyesíteni. A különböző osztályokba tartozó foltokat további csoportokba sorolhatjuk különböző geometriai jellemzőik szerint; a felmerülő problémák a *foltelezés* témakörébe tartoznak (lásd 4.3 pont).

Az *élkeresés* az "élgyanus" pontok *kitűzését*, majd a megtalált élek *korrekcióját* jelenti. Az előbbivel a 4.4.1 alpontban foglalkozunk; itt a gradiens nagyságát, illetve irányát figyelembe vevő módszereken kívül az éleket vonalszakaszokból összeállító, illetve texturális jellemzőkből meghatározó módszereket tekintjük át. Az élkorrekció, vagyis az "élgyanus" pontok elemzése a 4.4.2 alpont témája; végső célja a "sima" lefutásu, folytonos élvonalak kialakítása. A felmerülő és egymásnak ellentmondó követelményeket nehéz egyidejűleg kielégíteni; ezért az analitikus és a grafikus (vonalvékonyító) módszerek mellett a környezeti hatásokat is figyelembe vevő relaxációs, illetve a tapasztalaton alapuló heurisztikus módszereknek is van létjogosultságuk.

A képek fizikai szintű leírásának, így végső soron értelmezésének és a képi (vizuális) információ megértésének alapját a foltok és az élek képezik, ezért a szegmentálás alapvető fontosságú lépés a számítógépes látás megvalósításában.

Felhívjuk még a figyelmet arra a fontos körülményre, hogy a mért, illetve a számított sajátságok sohasem egzakt értékek: "*sajosak*". Ugyancsak nehézségeket okoz az "elég kicsi s-távolság" pontos megfogalmazása is.

Ezek miatt a szegmentálás, illetve az osztályozás mindig bizonyos hibával jár. A következőkben ezekre a problémákra még részletesebben visszatérünk.

4.2 FOLTKERESÉS

Az egy osztályba tartozó, (4)-összefüggő tartományokat *foltnak* nevezzük, ha a sajátságvektorok *hasonlósági* jellemzőket mérnek. Ez más szóval azt jelenti, hogy az egy osztályba sorolt képpontok *homogének* a vizsgált sajátásra nézve. A továbbiakban *értékesnek* nevezzük azokat a képpontokat, amelyeket besoroltunk valamelyik osztályba. A többi pont *háttérpont*; ezeket mindig összefüggőnek tekintjük, vagyis csak egy "háttérosztályt" veszünk figyelembe.

A szegmentálás "jóságának" egyik ismérve az, hogy milyen mértékben teljesül a homogenitási feltétel. Ideális esetben a tartomány *belső* pontjainak (amelyeknek minden (4)-szomszédja a tartományhoz tartozik) sajátságvektorai közel azonosak, vagyis nincsenek "lyukak" a tartomány belsejében. További kíváncsi, hogy a szomszédos tartományok jelentősen *különbözzenek* azon sajátságuk tekintetében, amelyre nézve homogének. Végül fontos szempont, hogy a határpontok (amelyeknek legalább egy (4)-szomszédjuk nem tartozik a tartományhoz) "*kellően sima*" *görbék*eken helyezkedjenek el.

A gyakorlatban alig lehet maradéktalanul teljesíteni minden követelményt. Minél szigorubbak ugyanis a homogenitási kritériumok (vagyis minél kisebb az "elég kicsi" s-távolság), annál több lesz a kisebb-nagyobb "lyuk" a tartományok belsejében, és annál zezugosabbak lesznek a határvonalak. Továbbá minél nagyobb különbségekhez ragaszkodunk a különböző szomszédos tartományok között, annál

inkább ki vagyunk téve annak, hogy a határpontok egy része elveszik, és egyes tartományok egybeolvadnak. Kielégítő szegmentálási eredményeket általában csak kompromisszumok árán lehet elérni, és csak akkor, ha mind az objektumok, mind a háttér eleve kellően homogének. Evégett a szegmentálás megkezdése előtt - szükség szerint - háttérkiegyenlitő, élkiemelő és simító transzformációkat kell végezni (lásd 2.3 és 2.4 pont).

A problémákat az eljárások jellege (globális, lokális, illetve hibrid) szerint tekintjük át.

4.2.1 GLOBÁLIS FOLTKERESŐ ELJÁRÁSOK

A szegmentálás szempontjából gyakorlatilag figyelembe vehető globális képjellemző a *hisztogram*. A vizsgált sajtáság (többsávós képek esetén sajtáságvektor) a világozósgkódok relatív gyakorisága.

Az eljárások olyan képek esetében adnak jó eredményt, amelyekeken *kevés féle, diszjunkt* objektum található, közel homogén háttéren. Az ilyen képek hisztogramján jellegzetes csucok: *módusok* található, amelyeket jól meghatározható *völgyek* választanak el.

A hisztogram módusai a különböző típusu objektumoknak, illetve a háttérnek felelnek meg. Az osztályozáshoz először meg kell határozni a csucok (módusok), illetve a völgyek helyzetét, ezekből állapítjuk meg az osztályok sajtászinét, illetve a küszöbököt. A szegmentálás kép-vágással történik (lásd 2.2.3 alpont).

A módszer hatékonyan csak egysávós képekre alkalmazható, bár többsávósokra is általánosítani lehet (lásd 4.2.1.6 alpont).

4.2.1.1 Hisztogramparticionálás

Ha a kép hisztogramja kétmódusu, a vágási küszöb automatikus meghatározására a gyakorlatban igen jól bevált a hisztogram particionálása a *szórásnégyzetek minimalizálásával*. Az eljárást [4-40] alapján dolgoztuk ki.

Az eddigieknek megfelelően jelöljük az egyes világosságkódok relatív gyakoriságát $\chi(q)$ -val, várhatóértéküket μ_H -val, maximális értéküket J -vel; ekkor nyilván

$$\sum_{q=0}^J \chi(q) = 1, \quad \mu_H = \sum_{q=0}^J q\chi(q). \quad (4-1)$$

Először keressük a κ küszöböt két szintre vágáshoz úgy, hogy a $q \leq \kappa$ világosságkódu képpontok a ω_1 , a többiek az ω_2 osztályba tartozzanak. Az osztályok statisztikai jellemzői a következők:

- előfordulási valószínűség (a *tapasztalati eloszlásfüggvény* értéke a κ helyen; lásd F2.2 pont):

$$s_1 = \sum_{q=0}^{\kappa} \chi(q), \quad s_2 = \sum_{q=\kappa+1}^J \chi(q) = 1 - s_1 \quad (4-2)$$

- feltételes várhatóérték:

$$\mu_1 = \sum_{q=0}^{\kappa} \frac{q\chi(q)}{s_1}, \quad \mu_2 = \sum_{q=\kappa+1}^J \frac{q\chi(q)}{s_2}; \quad (4-3)$$

- feltételes szórásnégyzet:

$$\sigma_1^2 = \sum_{q=0}^{\kappa} \frac{(q - \mu_1)^2 \chi(q)}{s_1}, \quad \sigma_2^2 = \sum_{q=\kappa+1}^J \frac{(q - \mu_2)^2 \chi(q)}{s_2}. \quad (4-4)$$

Vezessük be a következő mennyiségeket:

- az osztályok *átlagos* együttes szórásnégyzete:

$$\begin{aligned}\sigma_A^2 &= \sigma_1^2 S_1 + \sigma_2^2 S_2 = \\ &= \sum_{q=0}^{\kappa} (q - \mu_1)^2 \chi(q) + \sum_{q=\kappa+1}^J (q - \mu_2)^2 \chi(q); \quad (4-5)\end{aligned}$$

- az osztályok *különbözőségi* szórásnégyzete:

$$\sigma_D^2 = (\mu_H - \mu_1)^2 S_1 + (\mu_H - \mu_2)^2 S_2; \quad (4-6)$$

- a *teljes hisztogram* szórásnégyzete:

$$\sigma_H^2 = \sum_{q=0}^J (q - \mu_H)^2 \chi(q). \quad (4-7)$$

Kis számolással belátható, hogy

$$\sigma_A^2 + \sigma_D^2 = \sigma_H^2. \quad (4-8)$$

A vágási küszöböt akkor tekinthetjük "jónak", ha σ_A^2 a lehető legkisebb, σ_D^2 pedig a lehető legnagyobb. Mivel (4-8) szerint ezek nem függetlenek egymástól, olyan küszöböt keresünk, amelyre az

$$\eta = \frac{\sigma_D^2}{\sigma_A^2 + \sigma_D^2} = \frac{\sigma_D^2}{\sigma_H^2} = \text{maximum}. \quad (4-9)$$

Figyelembe véve, hogy σ_H^2 egy adott hisztogram esetében állandó, elegendő megkeresni azt a κ küszöböt, amelyre σ_D^2 maximális. Helyettesítsük be (4-8)-ba (4-7)-et és

(4-5)-öt, ekkor négyzetreemelés és összevonás után a

$$\sigma_D^2 = \mu_1^2 S_1 + \mu_2^2 S_2 - \mu_H^2$$

összefüggéshez jutunk. Használjuk még ki, hogy (4-3) és (4-1) alapján

$$\mu_1 S_1 + \mu_2 S_2 = \sum_{q=0}^{\kappa} q \chi(q) + \sum_{q=\kappa+1}^J q \chi(q) = \mu_H,$$

így végül is a megoldandó szélsőérték-feladat:

$$S_1 S_2 (\mu_2 - \mu_1)^2 = \text{maximum.} \quad (4-10)$$

Könnyű belátni, hogy (4-6) olyan diszkrét (kétértékű) valószínűségi változó szórásnégyzete (lásd F2.3 pont), amely S_1 valószínűséggel veszi fel a μ_1 , illetve S_2 valószínűséggel a μ_2 értéket, és várhatóértéke:

$$\mu = \mu_1^2 S_1 + \mu_2^2 S_2.$$

Szemléletesen ezt úgy értelmezzük, hogy az optimális válgás esetében az osztályok várhatóértékének különbsége a lehető legnagyobb, miközben a szórásnégyzetük minimális. Az $|\mu_1 - \mu_2|$ különbséget azonban önmagában *nem elég* vizsgálni, mert a küszöb változtatásával az osztályok előfordulási valószínűsége is megváltozik.

Megjegyezzük még, hogy a (4-9) összefüggéssel definiált η az osztályok szétválaszthatóságának mértékéül is szolgálhat ($0 \leq \eta \leq 1$):

$\eta = 0$ esetén a hisztogram: $\chi(q) = \text{állandó}$;

$\eta = 1$ a tiszta kétmódusu hisztogramot jellemzi.

Az eredmények könnyen általánosíthatók kettőnél több módusu hisztogramok particionálására is [4-32].

Az eljárást a 19/a képen mutatott tesztképre hajtottuk végre, amelynek hisztogramját a 19/b képen adtuk meg. A $\kappa = 25$ küszöbnél két szintre vágott eredménykép a 19/c képen látható.

4.2.1.2 Közelítés normális sűrűségfüggvényekkel

Tudjuk, hogy digitális képek esetében a hisztogram az analóg képek sűrűségfüggvényének megfelelője (lásd 2.2.1 alpont), így megközelíthető sűrűségfüggvények lineáris kombinációjával [4-30]. Kézenfekvőnek látszik a feltételezés, hogy a világosság kódok eloszlása közel normális; ezen a feltevésen alapul az általunk kidolgozott eljárás [4-31].

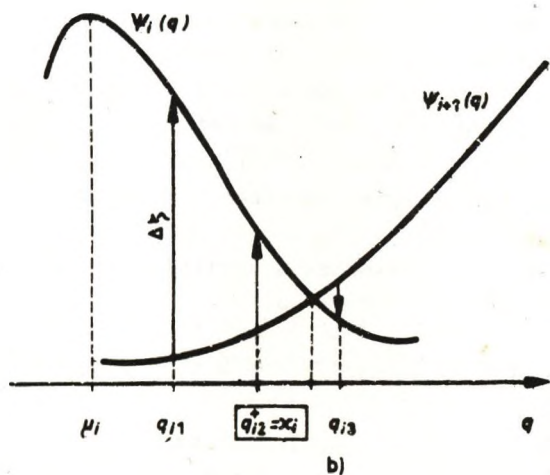
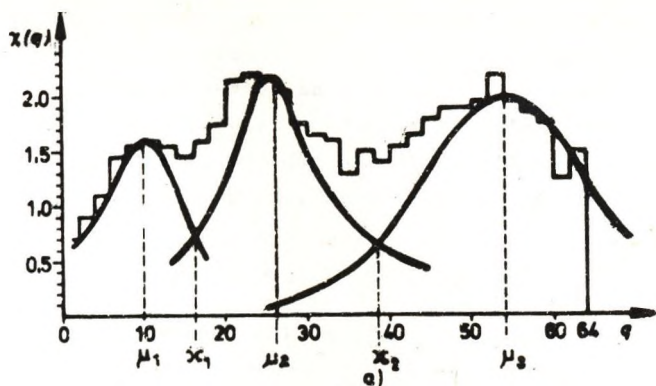
A 4-2/a ábrán hárommódusu hisztogram vázlatát látjuk, amelyet három normális sűrűségfüggvénnyel, illetve Gauss-görbével kívánunk megközelíteni.

Általánosságban tegyük fel, hogy a világosság kódokat $h \geq 2$ osztályba lehet sorolni (vagyis a hisztogram legalább kétmódusu). Jelöljük az osztályokat d_1, d_2, \dots, d_h -val az i -edik osztály sűrűségfüggvényét $s_i(\zeta)$ -val, a közelítendő hisztogramot $\chi(q)$ -val, a közelítő sűrűségfüggvényt $s(\zeta)$ -val. Ekkor a feltevéseink alapján felírhatjuk:

$$\chi(q) \approx s(\zeta) = \sum_{i=1}^h s_i(\zeta) = \sum_{i=1}^h s(\zeta | d_i) P(d_i), \quad (4-11)$$

ahol $s(\zeta | d_i)$ az i -edik osztály feltételes (a priori) sűrűségfüggvénye, $P(d_i)$ pedig annak valószínűsége, hogy egy képpont világosság kódja az i -edik osztályba esik.

A feladat a d_1, d_2, \dots, d_h osztály ("döntési tartomány"), vagy, ami ezzel egyenértékű, a $\kappa_1, \kappa_2, \dots, \kappa_{h-1}$ küszöb meghatározása. A megoldásra a Bayes-féle döntés-



4-2. ábra. Híszogram felbontása: a) hárommódusu híszogram, b) küszöbérték meghatározása

függvényt használtuk, amelyet részletesen ismertetünk az 5.2.2.1 alpontban. Eszerint a $q \sim \zeta$ ($0 \leq q \leq J$; esetünkben $J = 63$) világosságkódot abba az osztályba soroljuk, amelyre az (F2-6) alapján számított $P(d_i|\zeta)$ feltételes besorolási valószínűség a legnagyobb. Be lehet bizonyítani (vö. (5-16) és (5-18) képlet), hogy

$$P(d_m|\zeta) > P(d_i|\zeta) \Leftrightarrow s_m(\zeta|d_m)P(d_m) > s_i(\zeta|d_i)P(d_i). \quad (4-12)$$

Tekintve, hogy az egyes osztályok $P(d_i)$ előfordulási valószínűsége azonos, (4-12) értelmében a döntési intervallumok úgy helyezkednek el, hogy felettük a megfelelő osztály sűrűségfüggvénye a legnagyobb. Elvileg egy döntéshez (4-12) szerint $h - 1$ összehasonlítást kellene elvégezni. A gyakorlatban azonban két szomszédos módus közötti küszöb körüli tartományban a többi sűrűségfüggvény olyan kis értékű, hogy a döntést nem befolyásolja. Így végül is a feladatot a szomszédos sűrűségfüggvények metszéspontjának meghatározására vezethetjük vissza.

Az $s_i(\zeta)$ függvényeket a hisztogram módushelyeinek és ezek környezetébe eső értékeknek a segítségével határozzuk meg. A közelítő Gauss-görbét a

$$\psi_i(\zeta) = \alpha_i \exp[-\beta_i(\zeta - \mu_i)^2] \quad (i = 1, 2, \dots, h) \quad (4-13)$$

alakban vesszük fel, mivel *eggyel több* szabad paramétere van, mint a normális sűrűségfüggvénynek, s így jobb közelítést érhetünk el. Helyettesítsük (4-13)-ba a $\zeta = \mu_i$, illetve a $\zeta = \mu_i + u$ értékeket, ekkor kis számolással az

$$\alpha_i = \psi_i(\mu_i),$$

$$(i = 1, 2, \dots, h) \quad (4-14)$$

$$\beta_i = \frac{1}{u^2} [\ln \psi_i(\mu_i) - \ln \psi_i(\mu_i + u)]$$

összefüggésekhez jutunk. Nyilván $\mu_i \in E63$, és feltesszük, hogy $\mu_1 < \mu_2 < \dots < \mu_h$. A közelítő függvény alakját döntően befolyásoló β_i paraméterre az $u = \pm 1, \pm 2, \dots, \pm s$ értékek (vagyis a módus körüli hisztogrampontok) felhasználásával a

$$\bar{\beta}_i = \frac{1}{2s} \left(\sum_{j=-s}^{-1} \beta_j + \sum_{j=1}^s \beta_j \right) \quad (4-15)$$

átlagot használjuk.

A (4-12)-ből következik, hogy a $[0, \mu_1]$ intervallumban a ψ_1 , a $[\mu_h, \infty]$ intervallumban a ψ_h , a μ_i módus környezetében (legalábbis a $\mu_i \pm 1$ értékre) a ψ_i függvény a legnagyobb. Így a hisztogram elején és végén nem kell számításokat végeznünk.

A κ_i küszöbre (amely a ψ_i és ψ_{i+1} közelítő görbe metszéspontjának abszcisszája) a (4-13) alapján egzakt számítási képletet kaphatunk. Jobb eredményekre vezet azonban az alábbi eljárás (lásd 4-2/b ábra):

1. Induljunk ki a μ_i módushelytől, és jobbra haladva képezzük a

$$\Delta\zeta = \psi_i(q_{i,j}) - \psi_{i+1}(q_{i,j}) \quad (q_{i,j} = \mu_i + j; \quad j = 1, 2, \dots)$$

különbségeket, és figyeljük, hogy $\Delta\zeta$ mikor vált előjelet.

2. Legyen $\kappa_i = \bar{q}_{ij}^+$, ahol q_{ij}^+ az előjelváltást megelőző világosságkód.
3. Ismételjük meg az eljárást $i = 1, 2, \dots, h-1$ értékre, ezáltal meghatároztuk a d_1, d_2, \dots, d_h osztályt

A szegmentálást - a kapott küszöbök alapján - h szintre vágással végezzük. Az egy osztályba sorolt képpontokat (amelyeknek világosságkódja az osztályt határoló két küszöb közé esik) az osztálynak μ_i által meghatározott saját színére színezzük át.

Az eljárást egy mikroszkópi sejt felvételre hajtottuk végre (lásd 20/a kép). Hisztogramját a 20/b képen adjuk meg; ennek alapján a fenti módszerrel meghatározott módusok és küszöbök:

$$\mu_1 = 10; \mu_2 = 36; \mu_3 = 52; \text{ illetve } \kappa_1 = 15; \kappa_2 = 43.$$

A szegmentálás eredménye a 20/c képen látható. A módszer előnye, hogy kevésbé számításigényes; gyengéje, hogy lapos lefutású hisztogramok vagy közeli csúcok esetén a csúcspont meghatározása bizonytalan.

4.2.1.3 Küszöbbebecslés entrópiafüggvénnyel

Mivel a módusok helyét elég nehéz pontosan kijelölni, számos más elvű módszert dolgoztak ki a szegmentálási küszöbök meghatározására. Az alábbiakban az eloszlásfüggvény entrópiafüggvényének elemzésén alapuló módszert ismertetjük [4-27] alapján.

A szokásos jelölések megtartásával tegyük fel, hogy a hisztogram két módusú, és legyen κ a keresett küszöb. Vezessük be a hisztogramból összegzéssel képzett, diszk-

rét tapasztalati eloszlásfüggvényre, illetve ennek entropiájára az

$$S_j = \sum_{q=0}^j \chi(q), \quad (4-16/a)$$

$$H_j = - \sum_{q=0}^j \chi(q) \ln[\chi(q)] \quad (4-16/b)$$

jelöléseket. Ezeket felhasználva a normalizált entrópiafüggvényt

$$\Phi(j) = - \sum_{q=0}^j \frac{\chi(q)}{S_j} \ln \left[\frac{\chi(q)}{S_j} \right] - \sum_{q=j+1}^J \frac{\chi(q)}{1 - S_j} \ln \left[\frac{\chi(q)}{1 - S_j} \right] \quad (4-17)$$

alakban írhatjuk fel. Kihasználva, hogy S_j , illetve $\ln(S_j)$ az összegzés szempontjából állandónak tekinthető és (4-1)-et is figyelembe véve, (4-17)-et a következő alakra hozhatjuk:

$$\Phi(j) = - \frac{1}{S_j} \left[\sum_{q=0}^j \chi(q) \ln[\chi(q)] - \ln(S_j) \sum_{q=0}^j \chi(q) \right] - \frac{1}{1 - S_j} \left[\sum_{q=j+1}^J \chi(q) \ln[\chi(q)] - \ln(1 - S_j) \sum_{q=j+1}^J \chi(q) \right]. \quad (4-18)$$

Jelöljük (4-16) alapján a teljes hisztogram entrópiáját H_j -vel, ekkor (4-18)-ra a

$$\phi(j) = \ln[S_j(1 - S_j)] + \frac{H_j}{S_j} + \frac{H_J - H_j}{1 - S_j} \quad (4-19)$$

végleges alakot nyerjük. A $\phi(j)$ függvényt néhány jellegzetes képtípusra a 4-3. ábrán mutatjuk be, [4-27] alapján. A keresett küszöböt az a $j = \kappa$ érték adja, amelyre $\phi(\kappa)$ maximális. Ezt $j = 1$ -től indulva fokozatos közeli-téssel találhatjuk meg, amihez minden lépésben csak S_j és H_j értékét kell kiszámítani:

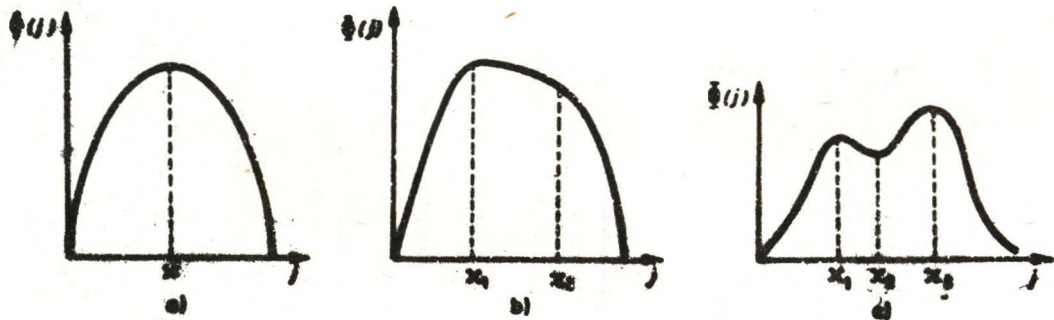
$$S_j = S_{j-1} + \chi(j),$$

$$H_j = H_{j-1} + \chi(j) \ln[\chi(j)],$$

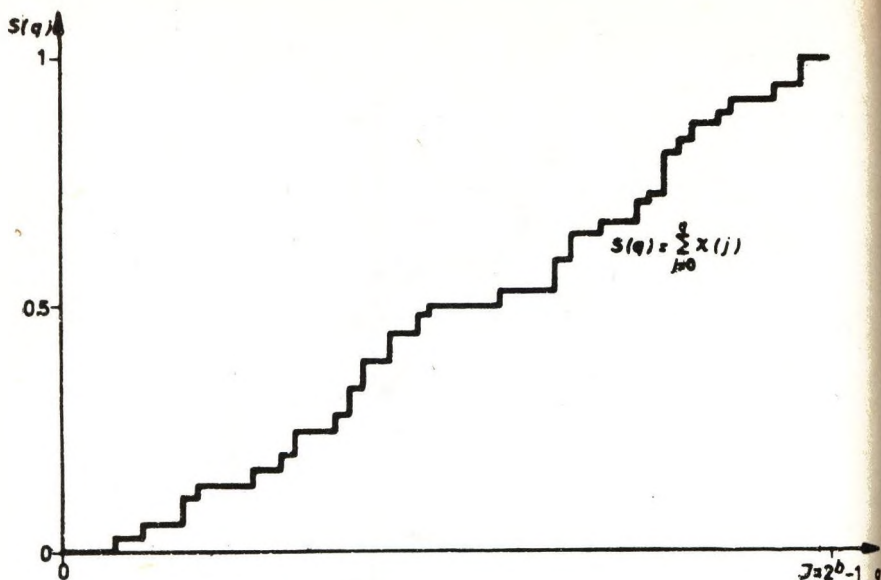
mivel H_j állandó. A módszer hátránya, hogy j értékét mindig végig kell futtatni a teljes $[0, J]$ intervallumon, ezért meglehetősen számításgényes. A (4-19) kifejezés könnyen általánosítható kettőnél több módusu hisztogram esetére is.

4.2.1.4 Szegmentálás az eloszlásfüggvény alapján

Lapult vagy közeli csucsockal rendelkező hisztogram esetén az eddigi eljárások nem adnak kielégítő eredményt. Ilyenkor a hisztogram helyett a (4-16/a)-ban felírt tapasztalati eloszlásfüggvény elemzése vezet célra [4-7]. Mint tudjuk (lásd F2.2 pont), ez lépcsősfüggvény, amelyben a q -adik lépcső magasságát a $\chi(q)$ érték adja (lásd 4-4. ábra).



4-3. ábra. Jellegetes entrópiafüggvények



4-4. ábra. Diszkrét eloszlásfüggvény

Folytonos $y = f(x)$ függvény görbületét az x helyen az

$$\frac{1}{\Upsilon(x)} = \frac{f''(x)}{[1 + f'^2(x)]^{3/2}} \quad (4-20)$$

képlet adja. Mivel $S(j)$ lépcsősfüggvény, ezt az

$$\frac{1}{\Upsilon(j)} \approx \frac{S(j+2n) + S(j-2n) - 2S(j)}{(1 + [S(j+u) - S(j-u)]^2)^{3/2}} \quad (4-21)$$

kifejezéssel közelítjük (lásd F3.2.3.1 alpont). Az $u = 1$ eset a differenciálhányados szokásos szimmetrikus közelítésének felel meg; $u > 1$ választással az $S(j)$ függvényt "kisimitjük".

A könnyebb számolás kedvéért a görbületfüggvényt folytonos függvénnyel közelítjük:

$$\frac{1}{\gamma(x)} \approx g(x) = \sum_{r=0}^m a_r T_r(x), \quad (4-22)$$

ahol $T_r(x)$ az r -edfoku Csebisev-polinom, és m a közelítés fokszáma. (A Csebisev-polinomok választását - azon túl, hogy ortogonális rendszert alkotnak - az indokolja, hogy így a közelítés a teljes intervallumban egyenletes.)

Az a_r együtthatókat úgy választjuk meg, hogy az

$$\epsilon^2 = \sum_{j=j_1}^{j_2} \left[g(j) - \sum_{r=0}^m a_r T_r(j) \right]^2 \quad (4-23)$$

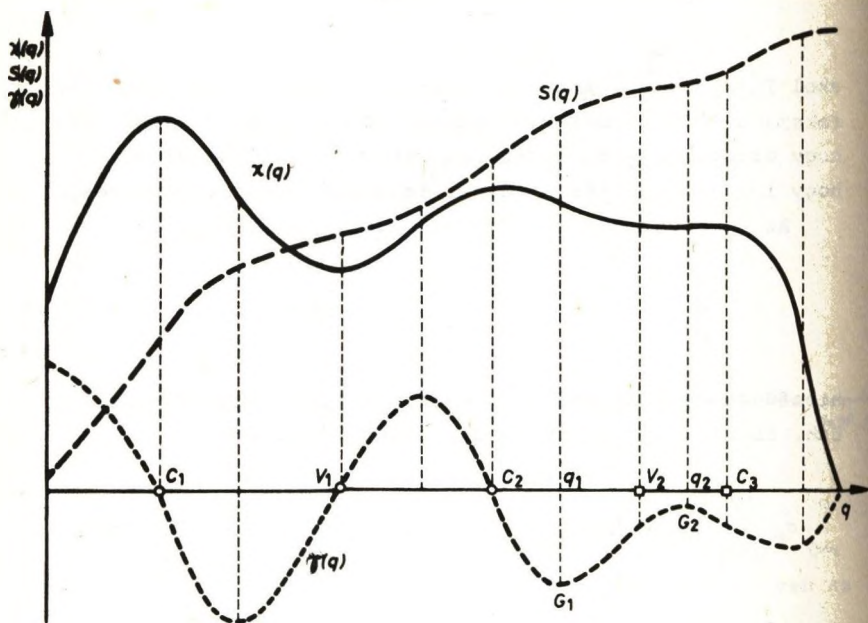
hibafüggvény minimális legyen a $[j_1, j_2]$ intervallumban. Ebből az a_i szerinti differenciálással a

$$\sum_{r=0}^m a_r \sum_{j=j_1}^{j_2} T_r(j) T_i(j) = \sum_{j=j_1}^{j_2} g(j) T_i(j) \quad (i = 0, 1, \dots, m) \quad (4-24)$$

összefüggést kapjuk, ami $m + 1$ lineáris egyenletet jelent az a_i együtthatók kiszámítására. (Az egyenletrendszert pl. a Gauss-Seidel-iterációval lehet megoldani.)

A 4-5. ábrán egy fiktív, folytonos hisztogram esetére rajzoltuk fel az eloszlás- és a görbületfüggvényt, hogy megvizsgáljuk a módszer használhatóságát. (A könnyebb áttekintés kedvéért a függőleges tengelyen különböző léptéket választottunk mindhárom görbéhez.) Láthatjuk, hogy a görbületfüggvénynek ott vannak zérushelyei, ahol a hisztogramnak szélsőértékei. A zérushelyek $(C_1 -) V_1 - C_2 - V_2 - \dots$ sorrendben követik egymást, és köztük a görbületfüggvénynek váltakozó előjelű szélsőértékei he-

lyezkednek el. (A hisztogram jellegétől függően a sorozat V-típusú zérushellyel is kezdődhet, ha a görbületfüggvény a negatív oldalon indul.)



4-5. abra. Fiktív (folytonos) hisztogram eloszlás- és görbületfüggvénye

Ha a hisztogram lapos, előfordulhat, hogy a görbületfüggvénynek két szélsőértéke között nincs zérushelye. (Lásd az ábrán a (q_1, G_1) , illetve a (q_2, G_2) pontot.)

Ilyenkor a fiktív zérushelyet, (amelynek típusa az előtte levőtől függ) lineáris interpolációval lehet meghatározni a

$$\left. \begin{array}{l} V_n \\ C_n \end{array} \right\} = q_1 + \frac{G_1}{G_1 + G_2} (q_2 - q_1) \quad (4-25)$$

képletből.

A szegmentálási küszöböket a görbületfüggvény V-típusú zérushelyei adják.

Az eljárással kapcsolatban meg kell még említenünk két számítástechnikai megfontolást. Az egyik az az ismert tény, hogy a Csebisev-polinomok csak a $[-1, 1]$ intervallumban vannak értelmezve. A görbületfüggvény előállítására érdekében ezért a világosságkód-tartományt a $[0, 1]$ intervallumba kell transzformálni; a küszöböket pedig a megfelelő zérushelyek visszatranszformálásával számíthatjuk ki.

A másik nehézséget az okozza, hogy a hisztogram kezdeti és végső lapos szakaszait rendszerint hirtelen ugrások választják el a követő, illetve az előző lépcsőtől. Ezekre a görbület nem értelmezhető, és a görbületfüggvény közelítését is megnehezítik. Ez az oka, hogy a (4-23) és a (4-24) képletben a határokat nem terjesztettük ki a teljes világosságkód-tartományra. A $[j_1, j_2]$ intervallumot úgy szokás megválasztani, hogy legfeljebb a hisztogramgörbe alatti terület 5%-a maradjon figyelmen kívül.

Az eljárás egyetlen választható paramétere a közelítés fokszáma. Minél magasabbra választjuk, annál pontosabbak lesznek a becsléseink, de - nyilvánvalóan annál inkább érzékelhetővé válnak a hisztogram szabálytalan és szélső értéként nem értékelendő ingadozásai.

A gyakorlatban 18-20-adfokú polinomokkal 5-7 küszöböt lehet meghatározni. A kapott küszöbök szerinti vágást végrehajtottuk a 20/a képen látható sejtfelvételen. Az eredményeket 10-edfokú közelítés esetén a 21/a, 25-ödfokú közelítés esetén a 21/b képen adtuk meg.

Az eljárás igen jó a hisztogram rejtett szélső értékeinek megkeresésére. A szegmentálás viszont - az irodalmi adatok szerint - csak a kevésbé részletgazdag képek esetében ad kielégítő eredményt.

4.2.1.5 Iteratív globális eljárások

Az eddig ismertetett eljárások számításigénye meglehetősen nagy. Másik probléma, hogy - ha az egyes tartományok összterülete jelentősen eltérő - a nagy területű tartományok a hisztogramban lefedhetik a kisebbeknek megfelelő módusokat. Mindkét probléma megkerülhető a következő szellemes rekurzív algoritmussal [4-39]:

1. Határozzunk meg egy maszkot, amelyen belül vizsgáljuk a képpontokat. Kezdetben az aktuális maszk az egész képet tartalmazza, és a maszklista legyen üres.
2. Határozzuk meg az aktuális maszkba eső képpontok hisztogramját. Keressük meg elemi úton a legmélyebb völgye(ke)t. Ha nem találunk völgyet, folytatás 4-től, különben folytatás 3-tól.
3. Osztályozzuk a képpontokat a kapott küszöb(ök) szerint. Képezzünk olyan maszkokat, amelyek az egyes osztályokhoz tartozó összefüggő tartományokat választják ki, és helyezzük ezeket a maszklistába.
4. Ha a maszklista nem üres, vegyük ki a soronkövetkező maszkot, és ismételjünk 2-től, különben az eljárás véget ért.

A völgyeket könnyen megkereshetjük egyszerű algoritmusokkal. Ha pl. a völgy "mélységét" a legmélyebb pont-

hoz és ennek alacsonyabbik szomszédjához tartozó hisztogramérték különbségeként definiáljuk, könnyű megkeresni a "legmélyebbet". Egy küszöb bevezetésével előírhatjuk a legkisebb kiértékelendő mélységet is. Az eljárást természetesen más módszerekkel is kombinálhatjuk (lásd pl. 4.2.3.2 alpont).

4.2.1.6 Foltkeresés többsávú képeken

Más jellegű nehézségek merülnek fel többsávú ($n \geq 2$) képek esetében. Egyrészt a többdimenziós hisztogramokat nehéz kezelni. Másrészt az n -dimenziós térben a saját-ságvektoroknak rendszerint több kombinációjuk lehetséges, mint ahány képpontból áll a kép, s így a hisztogram elveszti statisztikai jellegét. Célszerű megoldás lehet, ha az n -dimenziós teret kisebb dimenziószámú alterekre bontjuk, s a szegmentálást ezekben végezzük el. A részeredményekből már kevesebb számolással lehet visszatérni az n -dimenziós térbe.

Egyszerű példaként tegyük fel, hogy 4-sávú ürfelvételünk van. Osztályozzuk a képet először az 1. és a 2., majd a 3. és a 4. sáv felhasználásával, külön-külön. Legyenek a kapott osztályok az első esetben a, b és c , a másodikban A és B . Képezzük ezekből az összes lehetséges 4-dimenziós osztályt, ezek: $(a A), (a B), (b A), (b B), (c A), (c B)$. Egy (q_1, q_2, q_3, q_4) saját-ságvektoru képpontot akkor sorolunk pl. a $(b B)$ osztályba, ha a (q_1, q_2) redukált saját-ságvektor a b , a (q_3, q_4) pedig a B osztályba esik.

Az alábbiakban - speciális esetként - bemutatunk egy eljárást *valódiszínes* képek előállítására a *prototípusmódszer* alapján [4-4]. Tulajdonképpen távolságméré-

sen alapuló, tanítóval segített osztályozásról van szó, amivel az 5.2.4 alpontban még részletesen foglalkozunk.

Induljunk ki abból, hogy rendelkezésünkre áll a kép 3 színösszetevője, amelyek a TV-technikában szokásos R-, G-, illetve B-sávban készültek. Így minden képpont-hoz 3 "világosságkódunk" van. Tekintsük ezeket egy *szinvektor* (\underline{V}_j) komponenseinek a 3-dimenziós *szintérben* (lásd 4-6/a ábra):

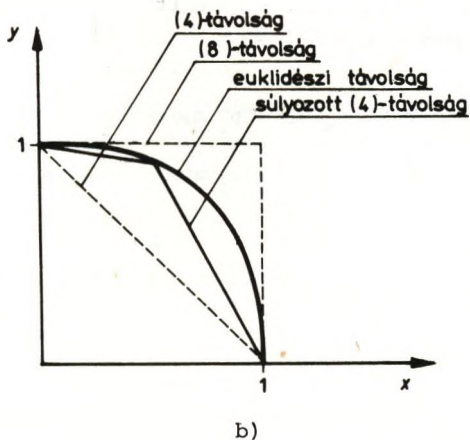
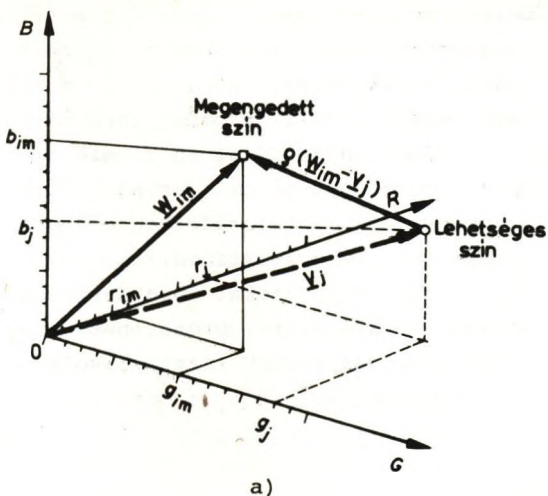
$$\underline{V}_j = (r_j, g_j, b_j) \quad (4-26)$$

(Az egyszerűség kedvéért most a képpontokat sorfolytonosan számoztuk, vagyis $j = 1, 2, \dots, N = K \cdot L$).

Legyenek megadva tanítók, amelyek meghatározzák az eredményképen *megengedett* szinosztályokat.

A *tanítókat* pl. az operátor jelölheti ki úgy, hogy (bármelyik színösszetevőben) mintaterületeket kerít körül a fényponttal, és megmondja, hogy ezek milyen színek legyenek az eredményképen. A mintaterületek minden pontja egy prototípusvektort határoz meg a szintérben, amelynek komponensei a mindhárom színösszetevőben azonos geometriai helyzetű képpontok világosságkódjai. Legyen megengedett szinosztályok száma h , az i -edik szinosztál-hoz tartozó mintaterületben levő képpontok halmaza $\{Q_i\}$ világosságkódjuk, az R, G, illetve B színösszetevőben r_{im}, g_{im} , illetve b_{im} és számuk legyen c_i . Ekkor a kijelölés a

$$\underline{W}_{im} = (r_{im}, g_{im}, b_{im}) \leftarrow (\rho_i, \gamma_i, \beta_i) \quad \begin{aligned} (m = 1, 2, \dots, c_i; \\ i = 1, 2, \dots, h). \end{aligned} \quad (4-27)$$



4-6. ábra. Távolságok a szintérben: a) prototípusvektor és színvektor távolsága, b) súlyozott (4)-es metrika

hosszárendelést jelenti, ahol $(\rho_i, \gamma_i, \beta_i)$ az i -edik osztályhoz megadott szinkód. (A mintaterületek egy pontból is állhatnak. Számuk tetszés szerinti, de minden osztályhoz legalább egyet ki kell jelölni. Ha azonos szinvektoru képpont több mintaterületben is előfordul, a megfelelő prototipusvektort ahhoz a szinosztályhoz rendeljük, amelyhez tartozó mintaterületben a legnagyobb a komponenseinek relativ gyakorisága.)

Ezután az egyes képpontokat abba a szinosztályba soroljuk, amelynek (valamelyik) prototipusvektorától a legkisebb s -távolságra vannak. Ezt a távolságot *minimax* módszerrel határozzuk meg:

$$d_i^* = \min_i (\max_m (\rho(\underline{w}_{im} - \underline{v}_j))) \quad \begin{array}{l} (i = 1, 2, \dots, h; \\ m = 1, 2, \dots, c_i; \\ j = 1, 2, \dots, N). \end{array} \quad (4-28)$$

A (4-28)-ban ρ szimbólummal jelölt távolságfüggvényt a következőképpen definiáltuk:

$$\rho(\underline{w}_{im} - \underline{v}_j) = \mu_1 \max(d_r, d_g, d_b) + \mu_2 \text{med}(d_r, d_g, d_b) + \mu_3 \min(d_r, d_g, d_b); \quad (4-29)$$

ahol

$$\begin{aligned} d_r &= \lambda_r |r_{im} - r_j|; \\ d_g &= \lambda_g |g_{im} - g_j|; \\ d_b &= \lambda_b |b_{im} - b_j|. \end{aligned} \quad (4-30)$$

A (4-29) képletben max, med, illetve min a számhármias legnagyobb, középső, illetve legkisebb elemét jelöli. Be lehet bizonyítani [4-55], hogy a (4-29) kifejezés nemnegatív definit, szimmetrikus, reguláris és teljesül rá a háromszög-egyenlőtlenség, vagyis *metrikaként* kezelhető.

A szem nemlineáris működése miatt a szintér *görbült* (lásd 1.1.3.3 alpont). Ezt a (4-29) alapján szakaszonként lineáris metrikával közelítjük meg, a súlytényezők megfelelő megválasztásával ($\mu_1 + \mu_2 + \mu_3 = 1!$). A 4-6/b ábrán kétdimenziós esetben megmutatjuk, hogy ily módon pl. az euklidészi metrika is jobban megközelíthető, mint a (4/8)-metrikával. Hasonló a szerepük a $\lambda_r, \lambda_g, \lambda_b$ korrekciós tényezőknek a (4-30) képletben: feladatuk, hogy kiegyenlítsék a leképező eszközök nemlineáris működéséből adódó torzításokat.

A konkrét megvalósítás során 6-bites világosság-, illetve szinkódokkal dolgoztunk, így a lehetséges szinek száma (4-26) szerint $2^{18} = 64^3$ volna. A számítások csökkentése érdekében először *adatzsugorítást* hajtunk végre. Ezt úgy végezzük, hogy kiszámoljuk mindegyik színösszetevő hisztogramját, majd 15 küszöbvel 16 közel egyenlő területű részre osztjuk ($\kappa_0 = 0$ és $\kappa_{16} = 64$ felhasználásával).

Ezután meghatározzuk az

$$r_j^* = \{t | \kappa_t \leq r_j < \kappa_{t+1}\} \quad (t = 0, 1, \dots, 15; \quad (4-31) \\ j = 1, 2, \dots, N);$$

redukált világosságkódokat, amelyek már 4-bitesek.

Az eljárást a G és a B komponensre is elvégezzük, itt természetesen más-más küszöbértékeket kapunk.

Igy a lehetséges szinek száma $2^{12} = 4096$ -ra csökken. Most a (4-28) képlet alapján megkeressük mindegyik lehet-

séges színhez a legközelebbi prototípusvektort. Az ehhez tartozó (6-bites) osztályszinkódot beírjuk egy 4096 elemű átszinező táblának abba az elemébe, amelynek indexe az illető színnek a (4-31) szerinti komponensek egymás mellé helyezésével nyert szinkódja. Végezetül átszinezük a képet a kitöltött táblázat szerint.

Az V/a képen egy állatszivről készült eredeti hőkamerafelvételt láthatunk. (A képet az Országos Ér- és Szívsebészeti Intézet szives hozzájárulásával közöltük.) A digitalizált kép a főtárban 4096 színű, az V/b képen azonban - technikai okokból - csak 64 szín látható. A képet a leírt eljárással 10 színre ($i = 10$) szinezük át (V/c kép; az azonos színű területek azonos hőmérsékletűek: a leghidegebbek a kék és a zöld, középhőmérsékletűek a lilá és a piros, legmelegebbek a sárga és a fehér tartományok). Az értékes területet az operátor által felrajzolt fekete konturvonalak határolják. A felesleges képrészek levágása után (V/d kép) az azonos színű foltok területét a 4.3.3.1 alpontban leírt eljárással számítottuk ki. Ebből a referencia-hőmérséklet (példánkban sárga 37,5 °C) alapján határoztuk meg az egyes részek felszíni hőmérsékletét. Az eredményeket a 4-1. táblázat tartalmazza.

4-1. táblázat. Hőfénykép kiértékelési eredményei

SZINKÓD	HCM	ISCHEMIAS RÉSZ		NORMAL RÉSZ		TELJES SZIV	
		TFP	TER*HOM	TER	TER*HOM	TER	TER*HOM
FEHER	38	371	33098.0	4074	154312.0	4945	187910.0
SÁRGA	37.5	1523	57112.5	2453	91987.5	3975	149100.0
NARANCS	37	2322	104414.0	7255	268435.0	10077	372849.0
PIROS	36.5	3504	127896.0	14173	517314.5	17577	645210.5
LILA	36	4221	151956.0	11499	413064.0	15720	565920.0
POLYA	35.5	136	4828.0	299	10514.5	433	15442.5
TURKIZ	35	2136	284760.0	1146	40110.0	9282	324370.0
ZÖLD	34.5	3451	119059.5	97	3346.5	3548	122405.0
KÉK	34	138	1692.0	217	7378.0	355	12070.0
ENCIAN	33.5	82	2747.0	137	4589.5	219	7336.5
ÖSSZESEN		24384	890563.0	41350	1512551.5	66234	2403114.5
ÁTLAG HŐMÉRSÉKLET			35.79		36.58		36.28

4.2.2 LOKÁLIS FOLTKERESŐ ELJÁRÁSOK

A globális eljárások eredendő közös hibája, hogy a folytonossági követelményeket nem tudják figyelembe venni. Így az említett hibák gyakoriak: a homogén tartományok "kilyukadnak", a határvonalak pedig "zajosak" és kiszélesednek. Ezeket a hiányosságokat kevésbé tapasztaljuk a lokális módszerek alkalmazása esetén. Nagy számuk miatt csak a legjellemzőbbeket tekintjük át, és példaképpen ismertetünk egy-egy eljárást.

Itt jegyezzük meg, hogy a lokalitás viszonylagos fogalom: a lokális eljárásokban a figyelembe vett környezet legtöbbször 3×3 vagy 5×5 , de sohasem több, mint 15×15 képpont.

Ennél nagyobb környezetek elemzése esetén *regionális* eljárásokról beszélünk; ezekben a környezet 128×128 képpontig is terjedhet. Ilyenkor tulajdonképpen felbontjuk a képet olyan *régiókra*, amelyeken belül valamilyen globális tulajdonság - pl. a háttér - kielégítően homogén (ami az egész képre nem teljesül). A tulajdonképpeni lokális eljárást ezután az egyes régiókon belül végezzük, régióként változó paraméterekkel.

Jellegzetes példaként megemlítjük a *rangszűréseket* (rank filtering), amelyekben a geometriai környezet elemeiből képzett *rendezett* halmazból választunk ki valamilyen *logikai* környezetet a lokális eljárásokhoz. (Így pl. elérhető, hogy bizonyos méretnél kisebb objektumokat ne vegyünk figyelembe.) A regionális eljárásokkal a továbbiakban nem foglalkozunk.

4.2.2.1 Összefüzési módszerek

Az egyszerű összefüzési módszerek a képpontokból indulnak ki. Két szomszédos képpontot akkor kapcsolnak össze, ha s -távolságuk (pl. világosságkódjuk különbsége) "elég kicsi". A foltokat a lehető legnagyobb összefüztött képponthalmazok alkotják. A módszerek előnye, hogy gyorsak és elég pontosan megtalálják a határvonalakat. Viszont rendkívül érzékenyek a képhibákra is (pl. szakadás a határvonalban): egyetlen téves "kapocs" elég ahhoz, hogy két különböző foltot egyesítsenek.

A kombinált összefüzési módszerekben - a zajérzékenység csökkentésére - környezetelemzést is végeznek: csak akkor kapcsolnak össze két szomszédos pontot, ha $m \times n$ -es környezetük s -távolsága is elég kicsi. Ezen az elven alapszik az az eljárás, amelyben csak a "kölosönösen jó szomszéd" képpontokat kapcsolják össze [4-34]. Ennek során először minden képponthoz meghatározzák az $m \times n$ -es környezetébe esők világosságkódjának átlagát. Legyen két szomszédos képpont világosságkódja q_i , illetve q_j , és környezetük átlaga \bar{q}_i , illetve \bar{q}_j . A két pontot akkor kapcsoljuk össze, ha a

$$d = s_1 |q_i - q_j| + s_2 |q_i - \bar{q}_j| + s_3 |q_j - \bar{q}_i|$$

képlettel definiált s -távolság kisebb egy adott küszöbértéknél. (Az alkalmasan megválasztott s_n súlyok összege 1.)

Egy másik eljárás a *fazettamodellen* alapul [4-44]. (Fazettának nevezzük a képsík egy tetszés szerinti, rendszerint szabályos sokszög alakú részletét. Lényeges, hogy a fazetták halmaza átfedés- és hézagmentesen fedje le a képet). Először tetszőleges fazettákra bontjuk a képet (szélső esetben minden képpontot egy fazettának

tekinthetünk), és mindegyik fazettához meghatározunk egy sajátságvektort. Ekkor iteráció kezdődik, amelynek során mindegyik fazetta sajátságvektorát helyettesítjük a vele szomszédosak sajátságvektorából meghatározott módon képzett értékkel. Ezután egyesítjük azokat a fazettákat, amelyeknek sajátságvektora elég kis s -távolságra van. Az eljárást ismételve fokozatosan durvább felbontáshoz jutunk. Célszerű az egyesítés elé bizonyos korlátokat állítani; pl. nem egyesíteni két fazettát, ha az eredő szórása túl nagy lenne stb.

4.2.2.2 Területnövesztés

Területnövesztésnek nevezzük azokat a módszereket, amelyek során a még nem osztályozott képpontokat a már kialakított, de még nem befejezett tartományokhoz fűzzük hozzá. A növesztési eljárást rendszerint soronként hajtják végre, ilyenkor *soronkénti növesztésről* beszélünk. Evégett meg kell állapítani a bejárési irányt, pl. balról jobbra és felülről lefelé.

Az *alapelv* a következő: a soron következő képpont sajátságvektorát összehasonlítjuk a bejárési irány szerint előtte (példánkban a felette és balra) levő tartományokéval. Ekkor három eset lehetséges:

1. ha egyik tartományéhoz sincs elég közel, a ponttal új tartomány kezdődik;
2. ha csak egyik tartományétól van elég kis s -távolságra, ehhez hozzáfűzzük;
3. ha mindkét tartományétól elég kis s -távolságra van, további két alesetet különböztetünk meg:
 - ha a két tartomány s -távolsága nagy, ahhoz fűzzük, amelyikhez közelebb van;

- ha viszont a két tartomány is elég közel van egymáshoz, ezeket egyesítjük, és a pontot az új tartományhoz fűzzük.

Példaként tekintsük azt az esetet, amikor a sajáttságvektor egyetlen összetevője a *világosságkód*. Jelölje T a vizsgált tartományt, amely N képpontból áll. Számítsuk ki a világosságkódok átlagát, illetve szórását a

$$\bar{q}_T = \frac{1}{N} \sum_T q(k, l);$$

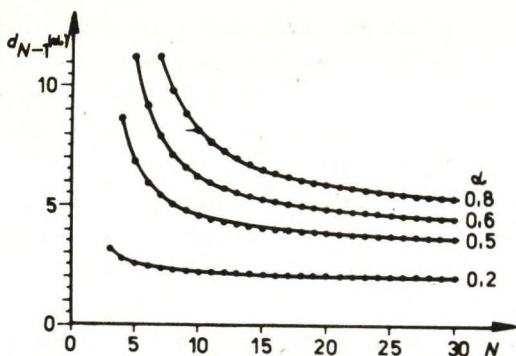
illetve a

$$\sigma_T^2 = \frac{1}{N^2} \sum_T (q(k, l) - \bar{q}_T)^2$$

képletből. Az s -távolságot pl. a

$$d = \left[\frac{(N - 1)N}{N + 1} \frac{(q - \bar{q}_T)^2}{\sigma_T^2} \right]^{1/2} \quad (4-32)$$

képlettel definiálhatjuk [4-23]; ahol q az aktuális képpont világosságkódja. Hogy ez a távolság elég kicsi-e, az $N - 1$ szabadságfokú T próbára vonatkozó szignifikancia-tesztel dönthetjük el. A 4-7. ábrán a $d_{N-1}(\alpha)$ függvényeket ábráztuk, ahol $\alpha \leq 1$ a felhasználó által szabadon megválasztható szignifikanciaszint [4-23]. Ha a (4-32) képlettel számított d érték nagyobb, mint az ábráról leolvasható $d_{N-1}(\alpha)$, akkor az eltérés *szignifikáns*. Ez esetben $1 - \alpha$ valószínűséggel állíthatjuk, hogy az aktuális képpont nem tartozik a vizsgált tartományhoz.



4-7. ábra. A T-próba szignifikanciafüggvényei

A területnövesztő módszerek előnye, hogy kis átmeneti meredekségeket (lásd 2-12. ábra) is képesek érzékelni. Ha kombináljuk az összefüzési módszerekkel, nagy pontosságú és a hibákra kevésbé érzékeny eljárásokhoz jutunk. Megjegyezzük, hogy a képhibák hatásának csökkentése érdekében az eredményt szokás két ellenkező irányú bejárásból nyert részeredmények metszetéből meghatározni.

4.2.2.3 Szétvágás és egyesítés

A módszer alapelve a következő (split and merge):

- vágjuk részekre az aktuális képszegmenst, ha nem elég homogén;
- egyesítsük azokat a szomszédos részeket, amelyek eléggé homogének.

A megvalósított eljárásokban különféle értékelőfüggvényeket használnak. A kiinduló szegmens mindig a teljes kép. A könnyebb programozás érdekében rendszerint 4 egyenlő területű részre vágunk.

Az eljárást fastrukturával lehet szemléltetni: a fa csomópontjai a képrészecskék felelnek meg. Egy csomópont-hoz annyi él tartozik, ahány részre vágjuk a megfelelő képrészletet. A fa azonos szintű, különböző ösökhöz tartozó csomópontjait egyesíthetjük.

A figyelembe vett sajátság rendszerint a világosságkód. A homogenitást egyszerű esetben az átlagtól való eltéréssel mérik, finomabb módszerekben a szórást is számítják.

A módszer hátránya, hogy lépcsős határvonalakat eredményez. A megvalósítás legfőbb nehézsége, hogy a fa leírásához szükséges tárkapacitás a képmérettel négyzetesen nő.

4.2.2.4 A ferde sikkok ("palatető")

módszere

Bontsuk fel a képsikot téglalap alakú *fazettákra* (lásd 4.2.2.1 alpont):

$$\Phi = \Phi_1 \cup \Phi_2 \cup \dots \cup \Phi_m;$$

ahol Φ_j a j -edik fazettát és Φ a teljes képsikot jelöli.

Közelítsük a digitális képfüggvényt minden fazetta felett egy ferde sikkal [4-21]. (Szemléletesen szólva: fedjünk le minden fazettát egy-egy "palatető"-lappal.) Más szóval: ha a (k, l) képkoordinátájú képpont a j -edik fazettába esik, vagyis $(k, l) \in \Phi_j$, akkor legyen

$$q(k, l) \approx \alpha_j k + \beta_j l + \gamma_j. \quad (4-33)$$

A (4-33)-mal meghatározott ferde sík gradiensének nagyságát és az x tengellyel bezárt irányszögét az

$$|\nabla q(k, l)| = \sqrt{\alpha_j^2 + \beta_j^2}, \quad (4-34/a)$$

illetve a

$$\operatorname{tg}(\vartheta_j) = \frac{\beta_j}{\alpha_j} \quad (4-34/b)$$

összefüggések adják (lásd F4.2.1 alpont).

Az egyes fazetták feletti ferde síkok általában V- (háztető) alakban metszik egymást. Közelíthetjük azonban a képet "visszintes" síkokkal (lépcsőkkel) is, ez esetben a metszésvonalak helyén ugrások keletkeznek. Vegyük fel a képkoordináta-rendszer origóját a kép középpontjában. (Ha a sor és az oszlopszám páratlan, ez valamelyik képpontba, különben két, illetve négy képpont középpontjába esik.) Jelöljük a j -edik fazettához tartozó képpontok x koordinátáinak halmazát K_j -vel, y koordinátáit L_j -vel; továbbá tegyük fel, hogy a képhez zaj is keveredett, amelynek várhatóértéke 0, szórása σ . Ekkor (4-33)-at általánosságban a

$$q(k, l) = \alpha_j k + \beta_j l + \gamma_j + \eta(k, l)$$

alakban írhatjuk. Keressük ennek közelítését olyan ferde síkkal, amelyre a négyzetes eltérés minimális:

$$\epsilon_j^2 = \sum_{k \in K_j} \sum_{l \in L_j} [\alpha_j k + \beta_j l + \gamma_j - q(k, l)]^2 = \text{minimum.}$$

Ebből - differenciálás és átrendezés után - figyelembe véve, hogy feltevéseink értelmében

$$\sum_{k \in K_j} k = \sum_{l \in L_j} l = 0,$$

a módosított együtthatókra némi számolás után a következő kifejezéseket kapjuk:

$$\alpha'_j = \frac{\sum_{k \in K_j} \sum_{l \in L_j} kq(k, l)}{\sum_{k \in K_j} \sum_{l \in L_j} k^2}; \quad (4-35/a)$$

$$\beta'_j = \frac{\sum_{k \in K_j} \sum_{l \in L_j} lq(k, l)}{\sum_{k \in K_j} \sum_{l \in L_j} l^2}; \quad (4-35/b)$$

$$\gamma'_j = \frac{\sum_{k \in K_j} \sum_{l \in L_j} q(k, l)}{\sum_{k \in K_j} \sum_{l \in L_j} 1}. \quad (4-35/c)$$

Ennek alapján lehetőség van arra (lásd [4-21]), hogy a közelítés jóságát a 4.2.2.2 alpontban említett T-próbával ellenőrizzük. A modell alkalmazása a következő algoritmushoz vezet:

1. Határozzuk meg minden képpontban - egy tetszőszerinti környezetének felhasználásával - a (4-35) képletekkel a közelítő sík paramétereit.
2. Rendeljük hozzá a vizsgált környezet minden pontjához a kapott $(\alpha'_j, \beta'_j, \gamma'_j, \varepsilon_j^2)$ számnégyest. Így minden ponthoz annyi számnégyes fog tartozni, ahány képpont környezetének eleme.

3. Válasszuk minden képpont sajátságvektorául azt a számnegyest, amelyik a legjobb illeszkedést adja, vagyis amelyben ϵ_j^2 a legkisebb.

Az így kapott sajátságvektorok alapján osztályozzuk a képet: vagyis kapcsoljuk össze azokat a képpontokat, amelyeknek sajátságvektora a T-próba szerint elég kis s-távolságra van, a megadott szignifikancia-szinten. (Ez más szóval azt jelenti, hogy a rájuk legjobban illeszkedő ferde síkok gradiensének hajlásszöge jó közelítéssel azonosnak tekinthető.)

4.2.3 HIBRID MÓDSZEREK

A hibrid módszerekben felváltva alkalmaznak globális és lokális eljárásokat egy probléma megoldására.

Már említettük a globális eljárások azon hiányosságát, hogy a folytonosságra nem tudnak tekintettel lenni. Ezért kézenfekvőnek látszik az a módszer, hogy egy globális eljárással szegmentált képre valamilyen simitó vagy mediánszűrést (lásd 2.3.2 alpont), esetleg egyszerűen többségi szűrést végzünk: minden képpont világosságkódját helyettesítjük a környezetében legtöbbször előforduló kóddal. Az ilyen eljárásokat *utófeldolgozásnak* nevezzük [4-6].

A 22. képen a 19/c képre végrehajtott többségi szűrés eredményét adjuk meg. Nem túl zajos képek esetében utófeldolgozással kielégítő eredményt lehet elérni.

Az igényesebb hibrid eljárások többségének az az alapelve, hogy a képpontokat először valamilyen lokális eljárással osztályozzák, majd az így kapott képponthalma-

zokat külön-külön kiszámított hisztogram alapján vágják szét, s ezeket a részhalmazokat újraosztályozzák az erede-
ti szempontok szerint.

4.2.3.1 Foltkeresés a gradiensmező felhasználásával

Bár nem szükségszerű, a gradienst általában lokálisan, a képpontok egy kis környezetének vizsgálatával határozzák meg (az egzakt definíciókat lásd az F4.2.1 alpontban). Közelítő kiszámítására egy sereg módszer létezik (lásd 4.4.1.1 alpont), amelyekből a gradiens iránya is meghatározható; gyakran használják azonban helyette az ismert Laplace-operátort.

Emlékeztetünk rá, hogy a gradiens értéke a közel homogén tartományok belsejében kicsi (közel 0), a foltok határainál viszont nagy, mégpedig a "sötét" oldalon negatív. Ezen alapul a "szelektív hisztogram" módszer [4-58]. Eszerint a hisztogramot csak azokra a képpontokra határozzák meg, amelyekben a gradiens abszolút értéke kisebb egy küszöbnél. A módszer akkor alkalmazható sikerrel, ha a képen világos színű objektumok vannak sötét háttéren. Ilyenkor - ha a határok elég élesek - a hisztogram jó közelítéssel kétmódusú (bimodális) lesz, mert a figyelembe vett pontok között közel azonos arányban lesznek értékes és háttérpontok. A szegmentálási küszöböt a völgypont megkeresésével nyerjük; a szegmentálás képvágással (lásd 2.2.3 alpont) történik.

A hisztogram módusainak kiemelésére gyorsan konvergáló eljárást használhatunk [4-49]. E célból számítsuk ki azoknak a képpontoknak *szelektív hisztogramját*, amelyekre a (4-44) szerinti *Roberts-operátorral* nyert gradiens értéke kicsi, vagyis

$$\max(|\hat{q}(k, l) - q(k + 1, l + 1)|, |q(k + 1, l) - q(k, l + 1)|) \leq \kappa_1,$$

ahol κ_1 megfelelően választott küszöbérték. Ez a feltétel csak a közel homogén tartományok belső pontjaira teljesül. (Igy kihagytuk vizsgálatainkból az átmeneti tartományok pontjait, amelyeknek értéke véletlenszerűen ingadozik a tartományok (közel állandó) szintjei között. Eloszlásuk és az átmeneti tartomány szélessége a képfellevő eszköztől, a beállítás élességétől stb. függ, a képtartalomtól nem.)

Jelöljük a q világosságkódu képpontok relatív gyakoriságát $\chi(q)$ -val, és legyen $q = 2, 3, \dots, J-2$, azaz hagyjuk figyelmen kívül a szélső értékeket. Számítsuk ki a

$$\bar{s} = \frac{1}{7} \sum_{j=-3}^{+3} \chi(q + j)$$

és az

$$a(q, j) = \frac{\chi(q) - \chi(q + j)}{\chi(q + j)} \quad (j = \pm 1, \pm 2, \pm 3)$$

mennyiségeket. Ha $\chi(q) > \bar{s}$, és $|a(q, j)| < \kappa_2$, (ahol κ_2 előre adott küszöb), akkor végezzük el a következő "szincserét" minden érintett képpontban:

$$\begin{aligned} q &\leftarrow q + j, & \text{ha } \chi(q + j) > \chi(q); \\ q &\leftarrow j \leftarrow q, & \text{ha } \chi(q) \geq \chi(q + j). \end{aligned}$$

Az eljárást addig ismételjük, amíg a megváltoztatott világosságkódok száma nagyobb egy κ_3 küszöbnél.

A gyakorlatban

$$3 \leq \kappa_1 \leq 10; \quad 0,6 \leq \kappa_2 \leq 0,8 \quad \text{és} \quad 0,05N \leq \kappa_3 \leq 0,1N$$

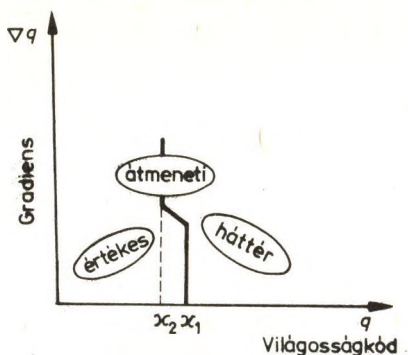
értékhatárokkal számolhatunk (N az összes képpont száma).

A "halmazott gradiens" módszerben minden képpont gradiensét meghatározzák [4-57]. Ezután összegzik az azonos világosságkódu képpontok gradiensét, és küszöbül azt az értéket választják, amelyre ez az összeg a legnagyobb. Jelölje $\{Q_i\}$ a q_i világosságkódu képpontok halmazát, és $\nabla q_i(k, l)$ a $\{Q_i\}$ -ba tartozó és (k, l) képkoordinátájú képpont gradiensét. Ekkor a küszöböt a

$$\kappa = \{q_i \mid \sum_{\{Q_i\}} \nabla q_i(k, l) = \max_j \left(\sum_{\{Q_j\}} \nabla q_j(k, l) \right)\} \quad (j = 0, 1, \dots, J)$$

kifejezés kiértékelésével kapjuk.

Kettős küszöböt használunk a szegmentálásra a "gradiens-hisztogram" módszerben [4-41]. Először is minden képpont gradiensét meghatározzuk, majd olyan kétdimenziós hisztogram készül, amelynek alapsíkjaiban az x tengelyen a világosságkódok, az y tengelyen a gradiensértékek vannak felmérve, míg a harmadik dimenzióban az ordináta a relatív gyakoriság. (Ezt úgy képzeljük el, hogy annyi hisztogramot készítünk, ahány különböző gradiensértéket találtunk. Egy hisztogram az azonos gradiensű képpontokra vonatkozik, s ezeket - a gradiensértékek növekvő sorrendjében - "egymás mögé" illesztjük.) A gradiens-hisztogram értelmezése a következő (lásd 4-8. ábra):



4-8. ábra. Gradienshistogram értelmezése

- Azok a képpontok, amelyeknek gradiense kicsi, homogen tartományok belsejében vannak, tehát vagy az objektumokhoz, vagy a háttérhez tartoznak. Az ezekre képzett hisztogram kétmódusu, bár a csucok nem azonos magasságúak (több a háttérpont, mint az értékes). Így jól szétválaszthatók a κ_1 küszöbvel.
- A nagy gradiensű képpontok az előbbi módszerrel kapcsolatban mondottak szerint az átmeneti tartományokba esnek. Ha az átmenetek nem túl élesek (a gyakorlatban általában ez a helyzet), akkor ezek hisztogramja egymódusu (unimodális), és legjobb szétválasztásuk a csucsnak megfelelő küszöbvel lehetséges.

A 19/a. képen látható tesztkép gradienshistogramját a 23/a. képen mutatjuk. (Az ordináták nagyságát a hisztogrampontok világosságával szemléltetjük; figyeljük meg a hisztogram jellegzetes "patkószerű" alakját.) A szegmentálás eredményét a 23/b képen adjuk meg. A kis gradiensű képpontokat a $\kappa_1 = 28$ küszöbvel, a nagy gradiensűket $\kappa_2 = 42$ küszöbvel osztályoztuk; a háttérpontokat fehérre, az értékeseket feketére színtettük át.

4.2.3.2 Sikklaszterezés

Ebben a módszerben a sajáttságvektorok szerinti osztályozást a területnövesztéssel kombinálják.

Az eddigiek alapján kézenfekvő módszer, hogy itt is a *hisztogramból* induljunk ki [4-18]. Az algoritmus a következő:

1. Keressünk meg minden csucst, és helyezzük ezeket egy listába (c-lista). Ha a c-lista üres, folytatás 6-tól; különben válasszuk ki a soron következő csucst.
2. Keressük meg mindazokat a képpontokat, amelyeknek világosságkódja megegyezik a csucséval; ezek lesznek az aktuális szegmenscsoport generátorpontjai. Helyezzük mindegyiket a szegmenslistába.
3. Ha a szegmenslista üres, folytatás 1-től; különben válasszuk ki a soron következő generátorpontot.
4. Vizsgáljuk meg az aktuális képpont (4)-szomszédságát. Keressünk ebben olyan képpontot, amelyiket még nem fűztünk az aktuális szegmenshez. Ha van ilyen, és elég kis s -távolságra van a generátorponttól (vagyis világosságkódjuk különbsége kicsi, és relatív gyakorisága nem nagyobb a generátorponténál), fűzzük hozzá.
5. Ha van olyan pontja az aktuális szegmensnek, amelynek szomszédságát még nem vizsgáltuk, legyen ez a soron következő aktuális pont, és folytatás 4-től, különben folytatás 3-tól.
6. Vizsgáljuk meg az összes szegmenst, és egyesítsük azokat, amelyekben azonos képkoordinátájú képpontok vannak. Ezzel az eljárás véget ért.

Az eljárás meglehetősen számításigényes. Valamit javít a helyzeten, ha az összetartozó szegmenseket már menet közben egyesítjük.

Egészen más a *piramiselvű sikklaszterezés* [4-8].

A "piramis" csomópontok hierarchikus rendszere, amelyek minden szinten négyzet alakú tömböt alkotnak. A legalsó (0-dik) szint csomópontjai a képpontok alkotta négyzetrács négyzeteinek középpontjában helyezkednek el; számuk $2^n \times 2^n$, ahol n a piramis rendszáma. Ez a szám felfelé haladva szintenként a negyedére csökken; az i -edik szinten $2^{n-i} \times 2^{n-i}$. A legalsó szint kivételével egy csomópontnak az eggyel alacsonyabb szinten az a 4×4 csomópont a *lehetőleges utódja*, amelyeknek geometriai középpontjában helyezkedik el; az eggyel magasabb szinten viszont az a 4 csomópont, a *lehetséges őse*, amelyik közrefogja (vagyis amelyiknek 4×4 -es környezetébe tartozik).

A tényleges hozzárendelés többlépéses iterációval történik. A legalsó szint csomópontjainak kezdeti saját-sága a megfelelő képpontok világosságkódja; majd - felfelé haladva a piramison - minden csomópont megkapja 16 lehetséges utódjának átlagértékét. Ezután ciklus indul. Ebben először alulról felfelé haladva minden csomópontot hozzárendelnek a lehetséges 4 őse közül ahhoz, amelyiknek saját-ságvektorától az övé a legkisebb s -távolságra van. (Ha többitől van azonos s -távolságra, a hozzárendelés véletlenszerű, illetve megtartják a korábbi iteráció eredményét. A széleken levő csomópontok külön kezelendők.) Ismét felfelé haladva minden őse felveszi a kijelölt utódai saját-ságának súlyozott átlagértékét; a súly az utód utódainak száma. Végül felülről lefelé haladva minden utód megkapja őse saját-ságvektorát. Ekkor a legalsó szinten egy osztályba sorolják azokat a csomópontokat, amelyeknek saját-ságvektora azonos. Az utóbbi három

lépést addig ismétlik, amíg a változás elég kis mértékű nem lesz. Az algoritmus gyorsan konvergál.

Az eredmény befolyásolható azzal, hogy a világosságkódok helyett más kiinduló sajátságokat választanak. Az eljárás egy módosított változatában például *texturális* jellemzők alkották a sajátságvektorok összetevőit [4-42]. Ezeket úgy határozták meg, hogy a képet kezdetben 8×8 képpontból álló blokkokra bontották, és ezek a "texturalelemek" képezték a 0-dik szint csomópontjait. A figyelembe vett texturális sajátság a *kontraszt* volt, amelyet az *együtt-előfordulási mátrix* alapján, az (5-48) összefüggés szerint lehet kiszámítani [4-20]. (Ehelyett - a számítások egyszerűsítése érdekében - a *különbségi hisztogram*-ból [4-19] származtatták le, amelyet egyszerűen a szomszédos képpontok világosságkódja különbségének négyzetéből állítottak elő.) Mivel az eljárás nagyon érzékeny a zajokra, a szegmentálás megkezdése előtt 3-5 lépésben mediánszűrést végeztek a képen, 3×3-as ablakkal.

Az eredmény lényegesen javult, amikor - további módosításként - beiktattak egy *felülről lefelé* történő *összekapcsolási* lépést; ezáltal kihasználták a piramis felső szintjein rendelkezésre álló globális információkat az alacsonyabb szinteken végbemenő lokális szegmentálásra.

Attól függően, hogy az őseket milyen legmagasabb szintig követjük vissza, az eljárás a képet különböző számú, homogén tulajdonságú (nem szükségképpen összefüggő) pontthalmazra fogja bontani; ezek összefüggő részthalmazai alkotják az értékes tartományokat.

4.3 FOLTELEMZÉS

A továbbiakban a szegmentált képből indulunk ki. Az értékes képpontok egy vagy több osztályhoz tartozhatnak, ezeket esetleges nevükkel és sajátságkódjukkal azonosít-

juk. (A kép vagy át van szinezve az osztályok sajátsszine szerint, vagy adott egy Q-tábla, amelyből kiolvasható, hogy a különböző világosságkódu képpontokat melyik osztályba soroltuk.)

4.3.1 FOLTSZÉTVÁLASZTÁS

A különböző képhibák miatt a szegmentálás után a foltok nem felelnek meg pontosan a valódi objektumoknak, illetve ezek határoló felületeinek. Legtöbbször arról van szó, hogy a térben közeli vagy a felvétel irányából egymást takaró tárgyak képe egybeolvad, esetleg csak a különböző árnyékhatások miatt. Az is előidézhet leképezési torzulást, hogy a térben különböző irányítású és visszaverő-képességű felületek adott megvilágítási irány esetén azonos fénysűrűségűnek látszanak.

Ha ismertek az objektumok lehetséges alakjai és várható legnagyobb területük, a további feldolgozás előtt korrigálhatjuk a szegmentálás eredményét.

4.3.1.1 Hámozás (erózió)

Ha tudjuk, hogy a valóságos objektumok konvexek, a túl nagy területű konkáv foltokat a hámozási technikával vághatjuk szét. Az eljárás neve onnan ered, hogy minden lépésben "lehámozzuk" a foltok legkülső rétegét, a határpontokat. (Határpontnak nevezzük a foltok azon pontjait, amelyeknek legalább egy szomszédjuk háttérpont.)

Az eljárás eredményessége nagy mértékben függ a szomszédság, illetve a távolságfüggvény megválasztásától. Jó eredményre vezet, ha az egymás utáni lépésekben felváltva vesszük figyelembe a (4)-, illetve a (8)-szomszédságot,

és az ennek megfelelő metrikát alkalmazzuk (lásd F3.1.1 és F3.2.1 alpont). Így a zsugorodás közel egyenletes, mivel ez a kombinált *nyolcszögtávolság* jól közelíti az euklidészt. (Lásd 4-9. ábra; itt a páratlan sorszámú pontok az elődeik (4), a párosak (8)-szomszédai.)

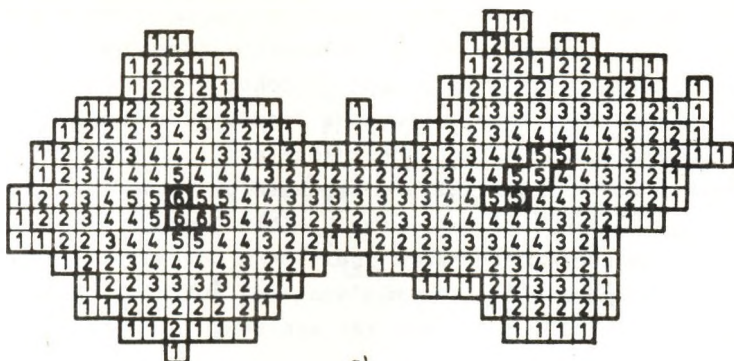
			5	5	5	5	5			
		5	4	4	4	4	4	5		
	5	4	4	3	3	3	4	4	5	
5	4	4	3	2	2	2	3	4	5	
5	4	3	2	2	1	2	2	3	4	5
5	4	3	2	1	●	1	2	3	4	5
5	4	3	2	2	1	2	2	3	4	5
5	4	4	3	2	2	2	3	4	4	5
	5	4	4	3	3	3	4	4	5	
		5	4	4	4	4	4	5		
			5	5	5	5	5			

4-9. ábra. Nyolcszögtávolság szemléltetése

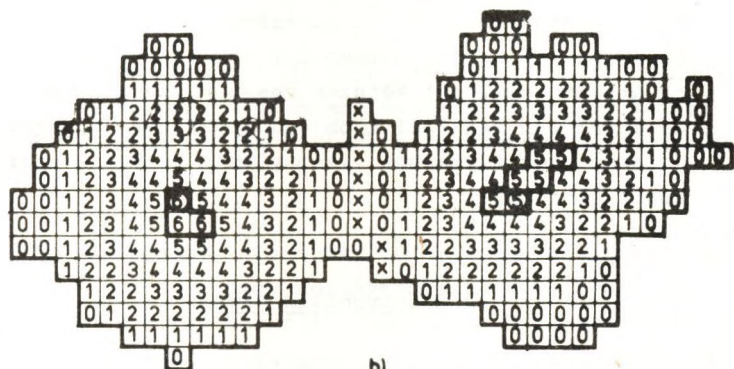
Hámozásra a 4-10. ábrán mutatunk példát. A 4-10/a ábrán az egyes lépésekben keletkező határpontokat a lépésszámmal sorszámoztuk.

A zsugorítás akkor ér véget, ha egyedülálló pontokhoz vagy olyan pontcsoportokhoz jutunk, amelyek csupa határpontból állnak, s így a következő lépésben eltűnnének. Ezeket *csomópontoknak* nevezzük, és az ábrán vastagon bekereteztük. A csomópontok közül a koordinátáik súlyozásával kijelölhetők a *szignifikánsak*.

A csomópontokból kiindulva most foltnövesztést vég-



a)



b)

4-10. ábra. Hámozás (a) és foltnövesztés (b)
nyolcszögtávoltság szerint

günk (vagyis minden ponthoz hozzákapcsoljuk a nyolcszög-szomszédját) annyi lépésben, ahányban eljutottunk a csomópontig. A számozatlanul maradt pontokat 0-val jelöljük meg (lásd 4-10/b ábra). Ezután a 0 jelű pontok azon halmazából, amely a két csomópont közé esik, *vonalvékonyító* eljárással (lásd 4.4.2.2 alpont) határozzuk meg a *vágóvonalat* (a 4-10/b ábrán az x-szel jelölt pontok). (Ha csak a szignifikáns csomópontokból indítjuk a növesztést, egyes esetekben kedvezőbb alakú vágóvonalhoz jutunk.) További megfontolásokra van szükség két esetben:

- Ha a csomópontok távolsága kisebb, mint az elérésükhöz szükséges lépésszámok összege, akkor a két folt határvonalait az 1. sorszámú pontok adják. A határvonalak a két folt között átmetszik egymást, az átfedő területet kétszer vesszük számításba.
- Ha a foltban kettőnél több csomópont van, a vágóvonalakat ezek egymáshoz viszonyított helyzete szerint korrigálni kell.

4.3.1.2 Szintvonalas vágás

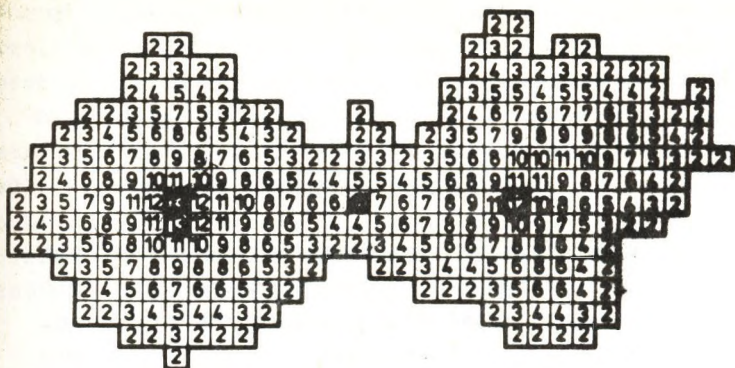
Az eróziós módszer általunk kidolgozott, némileg módosított változatában először meghatározzuk a folthoz tartozó pontok legkisebb távolságát a háttértől [4-14].

A metrikát a

$$\rho((k, l), (i, j)) = 2 \max(|k - i|, |l - j|) + \min(|k - i|, |l - j|)$$

képlet definiálja, ahol (k, l) , illetve (i, j) a vizsgált két képpont képpontkoordinátái.

Az azonos távolságu pontokat "*szintvonalakkal*" kötjük össze. (A távolságértékeket a 4-11. ábrán adjuk meg.) A szintvonalak *lokális maximumokat* zárnak körül (az ábrán



4-11. ábra. Távolságszámítás a szintvonalas vágáshoz

vastagon bekeretézve). Ha a maximum több pontból áll, megkeressük a *szignifikáns* maximumot (az ábrán ezt körrel jelöltük). Ezután minden maximumra elvégezzük a következő eljárást:

1. A maximumból kiindulva fokozatosan növesztjük a foltot úgy, hogy hozzávesszük a következő szintvonal által határolt területet. A növesztés addig tart, míg *nyeregpont*hoz nem érünk: amelynek legalább egyik (8)-szomszédjához nagyobb távolságkód tartozik, mint a sajátja. (A 4-11. ábrán a bal oldali maximumból kiindulva a körrel jelölt 6 távolságkódu pont.)
2. A megtalált nyeregpontból 2 irányban kiindulva (8)-összefüggő *vágóvonalat* keresünk úgy, hogy mindig a lehető legkisebb távolságkódu (8)-szomszédokon át haladunk addig, amíg háttérponthoz nem jutunk. Ha egy nyeregpontból kiindulva vágóvonalhoz érünk, töröljük a hozzá tartozó vágási pontokkal együtt.

Példaképpen végrehajtottuk az eljárást a 23/b képre; a kapott szintvonalakat a VI. képen különböző színnel jelöltük. Az eredmény a 24. képen látható.

4.3.2 MINTAILLESZTÉS

A geometriai és/vagy topológiai jellemzők szerinti osztályozás a foltelemzésnek egyik fontos részterülete; a figyelembe vehető sajátságok száma szinte korlátlan. Az alábbiakban csak a (geometriai) *alak szerinti osztályozással* foglalkozunk.

Igen gyakori eset, hogy rendelkezésünkre áll a képen lehetséges vagy várt objektumok "ideális" képe: *mintája*; ezeket egy *mintatárban* tároljuk. Az osztályozás összehasonlítással történik, innen az elnevezése: *mintaillesztés* (template matching). Természetesen nem tételezhetjük fel, hogy a minta pontosan illeszthető a kép egy részére; a valódi célunk csak a mintához legjobban "hasonlító" alakzat(ok) megtalálása lehet. A hasonlóság megállapításánál - a leképezési hibákból adódó eltéréseken túlmenően (amelyek hatását előzőleg célszerű minimalizálni a 2. fejezetben leírt módszerekkel) - az okozza a fő nehézséget, hogy a képen az objektumok mérete és állása tetszőleges lehet, a mintákat viszont - természetesen - meghatározott méretben és állásban tároljuk.

A mintaillesztési módszerek talán egyik legrégebbi alkalmazási területe a *karakterfelismerés*. Ha korlátozzuk a felismerendő karakterek típusát (pl. egy írógép- vagy sornyomtatótípus stb.) vagy számát (pl. számjegyek), egyszerűbb módszerekkel is nagy biztonságu azonosítást érhetünk el. (Az utóbbi esetben pl. kézzel írt számjegyek felismerése is megoldható; így működnek egyes postai irányítószám-leolvasó automaták.) A feladat általános megol-

dása (tetszőleges betűtípus, teljes abc; nem is szólva arról, ha kézírást is megengedünk) azonban *csak* mintaillesztési módszerekkel *reménytelen* vállalkozás.

Hasonló problémákhoz vezethet, ha ugyanarról a jeletről (tárgyról, területről) különböző nézőpontból készült felvételeink vannak. Ez esetben nem előre adott minták, hanem a képeken található *azonosítási pontok* illesztéséről van szó. A szükséges geometriai transzformáció(ka)t definiálhatjuk (lásd 3.4 pont)

- ismert vagy levezethető analitikus összefüggésekkel, illetve
- elegendő számú azonosítási pont interaktív megadásával vagy automatikus meghatározásával.

A képen előforduló élek, vonalak, sávok, görbék stb. megtalálása ugyancsak rokonítható feladat, hiszen lényegében előre ismert szerkezetű képrészeket keresünk.

4.3.2.1 Numerikus eljárások

Az illeszkedés "jóságát" gyakran mérik a képrészlet és a minta képpontjai közötti távolsággal. Jelölje $q(k, \bar{l})$ szokás szerint a digitális képfüggvényt és $M(s, t)$ a $(2m+1) \times (2n+1)$ méretű ablakon értelmezett mintát. Ekkor a

$$D(k, \bar{l}) = \sum_{s=-m}^{+m} \sum_{t=-n}^{+n} \rho(q(k + s, \bar{l} + t) - M(s, t)) \quad (4-36/a)$$

mennyiség jellemző az illeszkedés pontosságára. A képletben ρ valamilyen (diszkrét) metrika (lásd F3.2.1 alpont). Megjegyezzük, hogy (8)-távolság esetén az összegzés helyett csak a legnagyobb különbséget vesszük figyelembe; használata akkor célszerű, ha a minta minden pontjában kis eltérést engedünk csak meg. Euklidészi távolság esetén a négyzetes,

(4)-távolság alkalmazásakor az abszolút eltérést minimalizáljuk. Az eljárás pl. a következő lehet: minden képpontban kiszámítjuk a (4-36/a) kifejezés értékét, és ott fogadunk el illeszkedést, ahol ez kisebb egy adott küszöbnél.

Másik igen gyakori mérték az

$$N(k, l) = \frac{\sum_{s=-m}^{+m} \sum_{t=-n}^{+n} q(k+s, l+t) M(s, t)}{\left[\sum_{s=-m}^{+m} \sum_{t=-n}^{+n} q^2(k+s, l+t) \right]^{1/2} \left[\sum_{s=-m}^{+m} \sum_{t=-n}^{+n} M^2(s, t) \right]^{1/2}} \quad (4-36/b)$$

képlettel definiált *normalizált keresztkorrelációs együttható*. Ismeretes, hogy $|N(k, l)| \leq 1$; továbbá $N(k, l) = 1$ akkor és csak akkor, ha $q(k+s, l+t) = cM(s, t)$; vagyis - egy konstans szorzótól eltekintve - a képrészlet teljesen megegyezik a mintával (lásd F2.3.3 alpont). Ha az illeszkedést a (4-36/b) szerint számított értékkel mérjük, egyezést ott fogadhatunk el, ahol $|N(k, l)| \geq 1 - \kappa$; itt κ előre adott küszöbszám.

További mértékek találhatóak pl. az [A-4], [A-10], [A-12], [A-14] stb. művekben.

A fenti és az ezekhez hasonló eljárások használata esetén az szükséges a biztonságos felismeréshez, hogy a képen található valós alakzat mérete és állása megegyezék a mintáéval. Ez a feltétel erős korlátozásokat szab az alkalmazásuknak, hacsak nem növeljük meg jelentősen a minták számát (ami mind a tárigény, mind a számítási idő nemkívánatos megnövekedésével jár).

A problémát megkerülhetjük, ha sikerül felbontanunk a mintákat olyan részmintákra, amelyeket méretüktől és állásuktól függetlenül is nagy biztonsággal tudunk azonosítani, és ezeket külön-külön illesztjük a képhez. Végül

a részminták egymáshoz való viszonyának ismeretében azonosítjuk a teljes mintát, a megtalált részobjektumok alapján.

4.3.2.2 Grafikus eljárások

A fentebb vázolt problémák elesnek, ha az objektumokat illesztjük a mintákhoz; mégpedig úgy, hogy előzőleg valamilyen "szabványos" helyzetbe forgatjuk és "normalizáljuk" a méretüket.

A feladatot az [1-48] rendszerben úgy oldottuk meg, hogy mind az objektumokat, mind a mintákat a *legkisebb tehetetlenségi nyomatéku* állásba forgattuk egy olyan koordináta-rendszerben, amelynek kezdőpontja az objektum súlypontja és tengelyei párhuzamosak a koordinátatengelyekkel [4-9]. Ezen kívül megkerestük azt a legkisebb befoglaló négyzetet, amelynek oldalhossza 2 hatványa, és középpontja egybeesik az objektum súlypontjával; ezt a mintákat befoglaló négyzet méretére hoztuk (nagyítással vagy kicsinyítéssel).

Az ilyen módon normalizált alakzatokat pontonként hasonlítjuk össze az ugyancsak normalizált mintákkal. Az egyezést $\%$ -ban fejezzük ki, a besorolásról küszöbök alapján döntünk.

A normalizáláshoz feltételezzük, hogy a szegmentált kép kétszintű (bináris), és az értékes képpontok "sulya" 1, a háttérpontoké 0.

Egy alakzat digitális nyomatékát az

$$M_{mn} = \sum_{k, l \in A} k^m l^n q(k, l) \quad (4-37)$$

képlet definiálja. Itt A a befoglaló négyzetet jelenti, és $q(k, l) \in \{0, 1\}$.

A folt területét az M_{00} nyomatóék adja; (4-37) alapján a súlypont koordinátáit a

$$k_0 = \frac{M_{10}}{M_{00}}, \quad l_0 = \frac{M_{01}}{M_{00}}$$

összefüggésekből számíthatjuk ki. A súlypont körüli elforgatáshoz szükségünk van a centrális nyomatóékokra; ezeket a következő képletből számíthatjuk ki:

$$I_{mn} = \sum_{k, l \in A} (k - k_0)^m (l - l_0)^n q(k, l). \quad (4-38)$$

A (4-38) alapján számított tehetetlenségi (=másodrendű centrális) nyomatóékokkal a forgatás szöge:

$$\vartheta_0 = \frac{1}{2} \arctg\left(\left|\frac{2I_{11}}{I_{20} - I_{02}}\right|\right). \quad (4-39)$$

Ha (4-39)-ben $I_{20} - I_{02}$ negatív, a $\vartheta = \vartheta_0 + \frac{\pi}{2}$ szöggel kell számolnunk; további korrekció szükséges, ha $I_{03} < 0$; ekkor a forgatási szög: $\vartheta = \vartheta_0 + \pi$.

Példaképpen osztályoztuk a 19/a képet, a VII/a képen látható minták szerinti ("kőrszerű"=bors, "ellipszerű"=kávészem, "cseppszerű"=almamag, és "buzogányszerű"=szegfűszeg) osztályokba. Kellő egyezés esetén besoroltuk az objektumokat a megfelelő osztályba, és az osztály mintájával azonos színűre színeztük át. Az eredményt a VII/b képen mutatjuk; látható, hogy a besorolás hibátlan. Általában természetesen nem ez a helyzet. Ilyenkor azokat az objektumokat, amelyek egyik mintával sem egyeznek meg a tűréshatáron belül, egy további "ismeretlen" osztályba soroljuk és rendszerint meg hagyjuk

eredeti színükben. Ezek osztályozására vagy interaktív módszerekkel vagy újabb osztályozási kritériumok bevezetésével kerülhet sor.

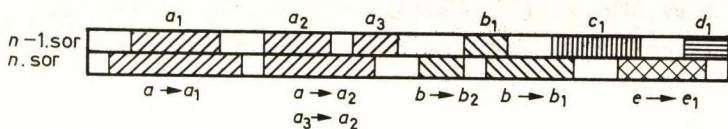
4.3.3 TERÜLET SZERINTI OSZTÁLYOZÁS

A foltok terület (nagyság) szerinti megoszlása az egyes osztályok fontos jellemzője. Erre vonatkozó statisztikák kiszámításához meg kell határoznunk az egyes foltok területét.

4.3.3.1 Foltszámlálás soronkénti növesztéssel

A 4.2.2.2 alpontban vázolt technika alapján *gyors foltszámláló* algoritmust dolgoztunk ki [4-15], kihasználva, hogy a kép a főtárban sorfolytonosan van tárolva. Elve a következő (lásd 4-12. ábra):

1. Meghatározzuk, hogy az aktuális sor képpontjai melyik osztályba tartoznak, illetve melyek a háttérpontok. Az egy osztályba tartozó összefüggő képpontsorozatokat *metszeteknek* nevezzük. (Az ábrán ezeket azonos vonalkázással jelöltük.) Tegyük fel, hogy az aktuális (n -edik) sor 5 foltot metsz át, az előző ($n - 1$ -edik) pedig 6-ot, s ezek rendre az a -, a -, b -, b - és e -, illetve az a -, a -, b -, c - és d -osztályba tartoznak.
2. Az aktuális sorban sorra vesszük a metszeteket. Aszerint, hogy az aktuális metszet a felette levőkhöz képest hogyan helyezkedik el, három esetet különböztetünk meg:



4-12. ábra. Sororkénti növesztés

- a (4)-szomszédai között *csak egy vele azonos* osztályu van: ez a folt kiegészül az aktuális metszettel ($a \rightarrow a_1, b \rightarrow b_1$);
- a (4)-szomszédai között *legalább két vele azonos* osztályu van: az eddig különböző foltok egyesülnek, s az új folt kiegészül az aktuális metszettel ($a \rightarrow a_2, a_3 \rightarrow a_2$),
- különben új folt kezdődik ($b \rightarrow b_2, e \rightarrow e_1$).

A feltérképezett foltokról lista készül, ez tartalmazza az azonosítási adatokat, az osztálykódot és a folt területét. Egyidejűleg a foltokat területük szerint *szóndákba* is soroljuk. Ehhez előre meg kell adni a zónahatárokat (z_i), számuk tetszés szerinti. A $z_0 = 0$ és $z_{\max} = 65535$ határok eleve adva vannak, így legalább egy zóna mindig létezik. A folt akkor kerül az i -edik zónába, ha területére teljesül az alábbi reláció:

$$z_{i-1} < T \leq z_i .$$

Az eredményekről összesítő, osztályonkénti és zónánkénti statisztikát lehet kérni, ez utóbbi a foltok terület szerinti felsorolásából áll. (Megjegyezzük, hogy az 1. zónát a statisztikákban nem vesszük figyelembe, ezért z_1 -et "szeméthatár"-nak is nevezzük.) A VII/b képre végrehajtott foltszámlálás eredményeit a 4-2., 4-3., illetve 4-4. táblázat tartalmazza.

4.2 táblázat

OSSZESITO STATISZTIKA

OSZTALYNEV	SZIN KOD	FOLTSZAM		KEPPONT	F O L T T E R U L E T			ATLAG	SZORAS	
		DB	%		OSSZ%	ERT%	MIN			MAX
BORS	L4	14	46.6	12134	10.9	26.0	446	2625	866	465
SZEGFUSZEG	S4	5	16.6	11768	10.6	25.2	1690	3365	2353	598
KAVE	P4	5	16.6	14586	13.1	31.2	2262	3650	2917	512
MAG	T3	6	20.2	8119	7.3	17.6	1163	1575	1353	132
FOLT OSSZESEN		30	100.0	46607	41.9	100.0				
KEP OSSZESEN				110592	100.0					

4.3 táblázat

OSZTALYONKENTI OSSZESITES

OSZTALYNEV	ZONAHATAR		DB	F O L T TERULET	
	ALSO<=T<FELSO				
BORS	400 - 600		4	2195	
	600 - 800		5	3699	
	800 - 1000		4	3615	
	2200 - 2800		1	2625	
	OSSZESEN		14		12134
SZEGFUSZEG	1400 - 1800		1	1690	
	1800 - 2200		2	3985	
	2200 - 2800		1	2728	
	2800 - 3500		1	3365	
	OSSZESEN		5		11768
KAVE	2200 - 2800		3	7587	
	2800 - 3500		1	3349	
	3500 - 5000		1	3650	
	OSSZESEN		5		14586
MAG	1000 - 1400		4	5103	
	1400 - 1800		2	3016	
	OSSZESEN		6		8119
	MINDOSSZESEN		30		46607

4.4 táblázat

ZONANKENTI ELOSZLÁS

1. ZONA	400<=T< 600				
BORS	OSZTALY:	446	565	591	593

2. ZONA	600<=T< 800				
BORS	OSZTALY:	681	743	747	750 778

3. ZONA	800<=T< 1000				
BORS	OSZTALY:	818	841	959	997

4. ZONA	1000<=T< 1400				
MAG	OSZTALY:	1168	1253	1295	1337

5. ZONA	1400<=T< 1800				
SZEGFUSZEG	OSZTALY:	1690			
MAG	OSZTALY:	1441	1575		

6. ZONA	1800<=T< 2200				
SZEGFUSZEG	OSZTALY:	1969	2016		

7. ZONA	2200<=T< 2800				
BORS	OSZTALY:	2625			
SZEGFUSZEG	OSZTALY:	2728			
KAVE	OSZTALY:	2262	2535	2790	

8. ZONA	2800<=T< 3500				
SZEGFUSZEG	OSZTALY:	3365			
KAVE	OSZTALY:	3349			

9. ZONA	3500<=T< 5000				
KAVE	OSZTALY:	3650			

4.3.3.2 Foltszámlálás szomszédságelemzéssel

Előfordulhat, hogy a további elemzéshez a foltok egymáshoz viszonyított helyzetét is ismernünk kell. Az alábbi algoritmus a területméréssel együtt a szomszédsági gráfot is előállítja [4-10]:

1. Jelöljük az egyes foltokat számkóddal (lásd 4-13/a ábra). Ezekhez a feldolgozás során különböző betűkódokat rendelünk, az összefüggőség és a szomszédság szerint. Legyen a háttérkód 0, és rendeljük ehhez az A betűt ($A = 0$). Legyen az aktuális kezdőpont a bal felső sarokpont.
2. Induljunk el az aktuális kezdőponttól sorfolytonosan, és cseréljük át a számkódokat az aktuális hozzárendelés szerint mindaddig, amíg jelöletlen új számkódot nem találunk, vagy a teljes alakzatot fel nem dolgoztuk.
3. Az előbbi esetben egy *sziget* bal felső sarokpontjához értünk, a továbbiakban ez lesz az aktuális kezdőpont. *Járjuk körül* a sziget határvonalát, és cseréljük át az aktuális osztályhoz nem tartozó, ezzel (4)-szomszédos pontok számkódját; egyszersmind jelöljük meg a sziget *bal oldali belső* határpontjait is. Az első körüljárás eredménye a 4-13/b ábrán látható.
4. Ezután *foltbejárás* következik, az aktuális kezdőpontból kiindulva. Előzőleg beállítjuk az új hozzárendelést (példánkban $B = 1$). A bejárás során át-cseréljük a folt pontjait, beleértve a bejelölt határpontokat, és egyuttal megjelöljük a folttal jobbról (4)-szomszédos pontokat is. Az első foltbejárás utáni állapotot a 4-13/c ábrán látjuk, a felderített szomszédsági gráffal együtt.

```

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 1 1 1 1 1 1 1 1 0 0 2 2 0 0 0
0 0 1 1 1 1 1 1 1 1 0 0 2 2 2 0 0
0 0 0 1 1 0 0 0 1 1 0 0 2 2 0 0 0
0 0 0 1 1 0 0 0 1 1 0 0 2 2 0 0 0
0 0 0 1 1 0 0 0 1 1 0 0 2 0 0 0 0
0 0 1 1 1 0 0 0 1 1 0 0 0 0 0 0 0
0 0 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0
0 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0
0 1 1 1 1 3 3 3 3 0 0 0 0 0 0 0 0
0 1 1 2 2 2 3 3 3 0 0 0 0 0 0 0 0
0 1 1 2 2 3 3 3 3 0 0 0 0 0 0 0 0
0 3 3 2 2 2 3 3 3 0 0 0 0 0 0 0 0
0 3 3 3 3 3 3 3 0 0 0 0 0 0 0 0 0
0 3 3 3 3 3 3 3 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

a)

```

A A A A A A A A A A A A A A A A A A
A A A A A A A A A A A A A A A A A A
A ① 1 1 1 1 1 1 1 1 A 0 2 2 2 0 0
0 A ① 1 1 1 1 1 1 1 1 A 0 2 2 2 0
0 0 A ① 1 0 0 1 1 1 1 A 0 2 2 2 0 0
0 0 A ① 1 0 0 0 1 1 1 A 0 2 0 0 0 0
0 A ① 1 1 0 0 1 1 1 1 A 0 0 0 0 0 0
0 0 A ① 1 1 1 1 1 1 1 A 0 0 0 0 0 0
A ① 1 1 1 1 1 1 1 1 A 0 0 0 0 0 0 0
A ① 1 1 1 3 3 3 3 3 A 0 0 0 0 0 0 0
A ① 1 2 2 2 3 3 3 A 0 0 0 0 0 0 0
A ① 1 2 2 3 3 3 3 A 0 0 0 0 0 0 0
A ③ 3 2 2 2 3 3 3 A 0 0 0 0 0 0 0
A ③ 3 3 3 3 3 3 A 0 0 0 0 0 0 0 0
A ③ 3 3 3 3 3 3 A 0 0 0 0 0 0 0 0
0 A A A A A A A A 0 0 0 0 0 0 0 0

```

b)

```

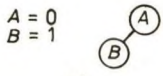
A A A A A A A A A A A A A A A A A A
A A A A A A A A A A A A A A A A A A
A B B B B B B B B B A 0 2 2 2 0 0
0 A B 1 1 1 1 1 1 B A 0 2 2 2 0 0
0 0 A B 1 ① 0 0 0 1 B A 0 2 2 2 0 0
0 0 A B 1 ② 0 0 0 1 B A 0 2 0 0 0 0
0 A B 1 1 ③ 0 0 1 1 B A 0 0 0 0 0 0
0 0 A B 1 1 1 1 1 B A 0 0 0 0 0 0 0
A B B 1 1 1 1 B A 0 0 0 0 0 0 0 0
A B 1 1 B ③ 3 3 3 3 A 0 0 0 0 0 0 0
A B 1 ② 2 2 3 3 3 3 A 0 0 0 0 0 0 0
A B B ② 2 3 3 3 3 3 A 0 0 0 0 0 0 0
A ③ 3 2 2 2 3 3 3 3 A 0 0 0 0 0 0 0
A ③ 3 3 3 3 3 3 3 A 0 0 0 0 0 0 0 0
A ③ 3 3 3 3 3 3 3 A 0 0 0 0 0 0 0 0
0 A A A A A A A A 0 0 0 0 0 0 0 0

```

```

A A A A A A A A A A A A A A A A A A
A A A A A A A A A A A A A A A A A A
A B B B B B B B B B A A F F F A A
A A B B B B B B B B B A A F F F A
A A A B B C C C B B B A A F F A A
A A A B B C C C B B B A A F A A A
A A B B B C C C B B B A A A A A A
A A A B B B B B B B B A A A A A A
A B B B B D D D D A A A A A A A A
A B B E E E D D D A A A A A A A A
A B B E E D D D D A A A A A A A A
A D D E E E D D D A A A A A A A A
A D D D D D D D A A A A A A A A A
A D D D D D D D A A A A A A A A A
A A A A A A A A A A A A A A A A A A

```



c)



d)

4-13. ábra. Foltszámlálás szomszédságelemzéssel:
a) kiindulási képrészlet,
b) eredmény az első szigetkörüljárás után,
c) eredmény az első foltbejárás után,
d) végeredmény

5. Most *visszatértünk* a sorfolytonos átcserélésre, az aktuális kezdőpontból indulva. A következő esetek lehetségesek:

a) A számkód nincs bejelölve:

- ha a bal oldali szomszédja vele azonos osztályu betűkód, erre átcseréljük; folytatás 2-től,
- ha a bal oldali szomszédja más osztály betűkódja, vagy a számkód még nem szerepelt hozzárendelésben, új sziget kezdődik; folytatás 3-tól.

b) A számkód be van jelölve, ekkor egy új folt bal felső sarokpontja. A továbbiakban ez lesz az aktuális kezdőpont; folytatás 4-től foltbejárással.

A végeredményt az eredő szomszédsgái gráffal és a hozzárendelésekkel a 4-14/d ábrán adtuk meg. Megjegyezzük, hogy a kapott foltok (4)-összefüggők, s a hozzárendelés egyértelmű.

4.4 ÉLKERESÉS

Ha a szegmentálás különbségi jellemzők alapján történik, *élkeresésről* beszélünk. A foltkeresésről mondtuk alapján nyilvánvaló, hogy az élek a foltok határvonalai. *Konturvonálnak* nevezzük az olyan folytonos és zárt élvonalakat, amelyek egy foltot, illetve egy lyukat teljesen körülhatárolnak. (A vonalak folytonosságának vizsgálatánál (8)-szomszédsgát veszünk figyelembe.)

Az élek *élpontokból* állnak. Az élpontokat ott keressük, ahol a sajátsgvektorok hirtelen, ugrásszerű változásokat mutatnak, más szóval a szomszédos képpontok sajátsgvektorának *s-távolsága elég nagy*. Bár elvileg nem

szükségszerű, az élkereső eljárások rendszerint *lokálisak*, s az esetek többségében a világosságkódok hirtelen változásait értékelik ki, ritkábban alapulnak statisztikai vagy texturaelemzésen. A lokális jellegből következik, hogy az élkereső eljárások érzékenyek a zajokra. A nehézségeket az okozza, hogy egyrészt a megtalált élpontok nem alkotnak folytonos vonalakat (vagyis nem (8)-összefüggők), másrészt olyan pontok is élpontnak minősülhetnek, amelyek a valóságban nem tartoznak egyetlen folt konturvonalához sem.

Elfogadható eredményeket csak a *kécmenetes* eljárások adnak. Az első menetet *élkitűzésnek* (edge detection) nevezzük: ekkor jelöljük ki az élpontokat. Ha van második menet, ezeket visszaminősítjük "*élgyanus*" pontokká, s az *élkorrekció* (edge correction) során alakítjuk ki a valódi éleket.

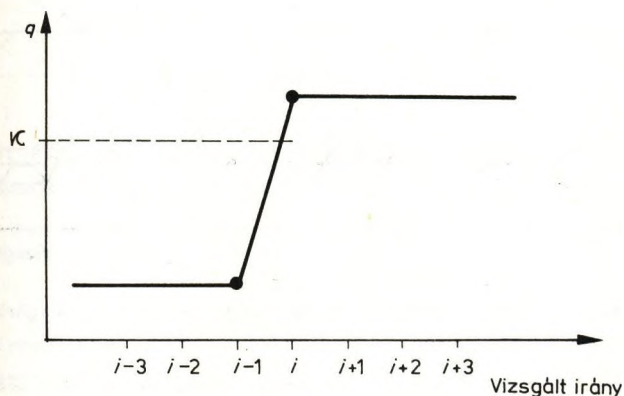
4.4.1 ÉLKITŰZÉS

A tulajdonképpeni szegmentálási eljárás az élkitűzés. Ekkor a képpontokat *két osztályba* soroljuk: az "értékes" pontokat nevezzük *élpontnak*, a többiek *háttérpontok*. Megjelenítéskor az éleket az élosztály saját színére színezzük át, a háttérpontokat vagy letöröljük a képről, vagy változatlanul meghagyjuk ("rarájzolás"). A további feldolgozáshoz természetesen nem kell a képet átszínezni, elegendő egy bináris élképet elkészíteni.

A feldolgozás szempontjából lépcső-, lejtő- és V-típusú éleket különböztetünk meg, az átmeneti tartomány szélessége, illetve az "ugrás" nagysága szerint. A változásokat különböző *irányokban* vizsgáljuk. *Átmeneti tartománynak* nevezzük egy konkrét élpont minősítése céljából egy adott irányban figyelembe vett két legtávolabbi képpont (8)-tá-

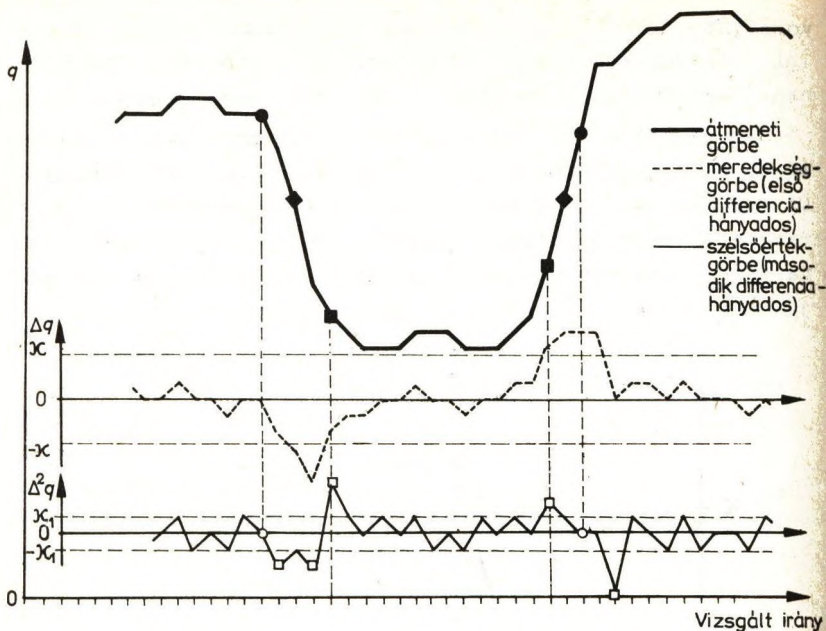
volságát. Az *ugrás nagyságát* ezen képpontok s-távolsága (pl. világosságkódjuk különbsége) határozza meg. Az "elég nagy" s-távolságot rendszerint egy küszöbvel definiáljuk.

Az "ideális" *lépcső* átmeneti tartománya 1 pont szélességű, és az ugrás nagyobb a küszöbnél (lásd 4-14. ábra). Ez esetben akár a "sötét"-oldali (az alacsonyabb szinthez tartozó), akár a "világos"-oldali (a magasabb szinthez tartozó) határpontot (esetleg mindkettőt is) kijelölhetjük élpontnak.



4-14. ábra. Ideális lépcsőátmenet

Az "ideális" *lejtőt* a 2-12. ábrán láthatjuk. A gyakorlatban az átmeneti görbe ezt közelítő lépcsősfüggvény, aminek következtében a meredekséggörbe is eltorzul (lásd 4-15. ábra; az ábrán a képpontokat a vizsgált irányban növekvően sorszámozzuk). Az átmeneti tartomány 5-11 képpont szélességű, amit a leképező eszköz nemlineáris működése és a diffúz fényvisszaverődésből eredő elmosódás okoz. Ezen kívül a meredekség változása is aszimmetrikus, ami a kvantálási hibák következménye. Az élpontokat ilyen esetben ott szokás kijelölni, ahol az átmeneti görbének

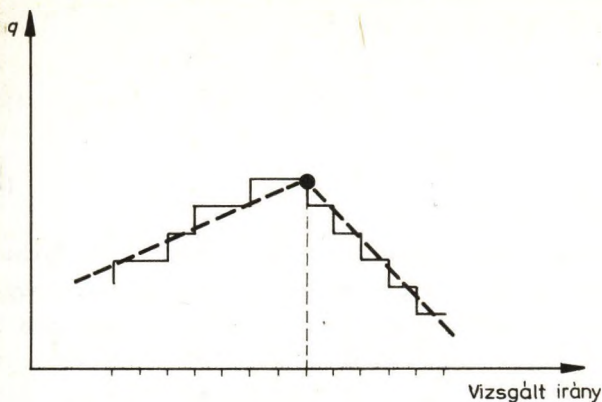


4-15. ábra. Digitális átmeneti, meredekség- és szélsőértékgörbe

"inflexiós pontja", vagyis az ábrán szaggatott vonallal rajzolt "meredekséggörbének" lokális szélsőértéke van.

A *V-típusú* élék felfedezése a legnehezebb, mivel sem az átmeneti tartományra, sem az ugrásra nem lehet egyértelmű becslést adni (lásd 4-16. ábra). Ez esetben a "meredekséggörbe" lépcsősfüggvény, s ennek ugrási helyéből lehet az él létezésére, illetve helyzetére következtetni.

Az élkereső eljárások nagy részében fontos szerepe van a képfelület gradiensének. Ezen azt értjük, hogy egy képpontot akkor minősítenek él(gyanus)-pontnak, ha - bizonyos környezetét vizsgálva - a képfüggvény gradiensének vagy valamilyen ebből lezármaztatott jellemzőnek az adott pontban szélsőértéke van.



4-16. ábra. Ideális V-típusú átmenet

Az élkijelölő módszereket a döntéshez figyelembe vett jellemzők szerint tekintjük át.

4.4.1.1 Gradiensmódszerek

Ismeretes, hogy egy függvény gradiensét bármely két, egymásra merőleges irányu parciális differenciálhányadosa alapján meg lehet határozni. A digitális "képfüggvény" esetében ezeket az (F4-21)-ben definiált (számítástechnikai okokból legtöbbször x és y irányu) differenciahányadosokból számítjuk ki.

A *gradiensmódszerek* a (k, l) pontot akkor minősítik élpontnak (vagy élgyanusnak), ha ott a gradiens (F4-23) szerinti normája *egy adott küszöbnél nagyobb*:

$$\|\nabla q(k, l)\| = + \sqrt{\Delta_x^2(q(k, l)) + \Delta_y^2(q(k, l))} > \kappa. \quad (4-40)$$

Az él iránya merőleges a gradiens irányára. A gyakorlatban e helyett a sokkal egyszerűbben kiértékelhető

$$\begin{aligned} \|\nabla q(k, l)\| &\approx |\Delta_x(q(k, l))| + |\Delta_y(q(k, l))| = \\ &= |q(k, l) - q(k-1, l)| + |q(k, l) - q(k, l-1)| > \kappa \end{aligned} \quad (4-41)$$

abszolútérték normát (lásd (F4-23/a)) szokás használni.

A (4-41) típusú közelítések legfőbb hibája, hogy aszimmetrikusak a vizsgált pontra nézve. A probléma kiküszöbölésére egy sereg konvolúciós szűrőt dolgoztak ki. Ezekkel a gradienst (illetve a szükséges differenciahányadosokat) a

$$\|\nabla q(k, l)\| \approx |\tilde{q}(k, l) \otimes G_x| + |\tilde{q}(k, l) \otimes G_y| \quad (4-42/a)$$

képlet alapján számítjuk ki; ahol $\tilde{q}(k, l)$ a (k, l) képkordinátájú pont $m \times n$ -es környezetét (lásd F4-3. ábra), a \otimes szimbólum az (F4-12) képlettel definiált konvolúciós összegzést jelenti.

Konvolúciós szűrők alkalmazása esetén a szűrő közép-pontjának megfelelő ponton áthaladó él irányát a

$$\vartheta \approx \arctg \frac{\tilde{q}(k, l) \times G_x}{\tilde{q}(k, l) \times G_y} \quad (4-42/b)$$

képletből becsülhetjük meg.

Az *optimális szűrőnek* ki kell elégítenie a következő feltételeket:

- *Izotrop* jellege megköveteli, hogy minden irányban azonos mértéket használjunk. Ezért a Δ_x , illetve a Δ_y differenciahányados kiszámítására szolgáló szűrők egymásnak a főátlóra vonatkozó tükörképei (lásd 4-17/a ábra).

$$\begin{array}{c}
 G_x \\
 \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad \begin{bmatrix} a & b & c \\ d & e & f \\ a & b & c \end{bmatrix} \quad \begin{bmatrix} a & b & -a \\ d & e & -d \\ a & b & -a \end{bmatrix} \quad \begin{bmatrix} a & b & -a \\ d & -2b & -d \\ a & d & -a \end{bmatrix} \quad \begin{bmatrix} a & 0 & -a \\ d & 0 & -d \\ a & 0 & -a \end{bmatrix} \\
 \\
 G_y \\
 \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} \quad \begin{bmatrix} a & d & a \\ b & e & b \\ c & f & c \end{bmatrix} \quad \begin{bmatrix} a & d & a \\ b & e & b \\ -a & -d & -a \end{bmatrix} \quad \begin{bmatrix} a & d & a \\ b & -2b & b \\ -a & -d & -a \end{bmatrix} \quad \begin{bmatrix} a & d & a \\ 0 & 0 & 0 \\ -a & -d & -a \end{bmatrix} \\
 \\
 \text{a)} \quad \quad \quad \text{b)} \quad \quad \quad \text{c)} \quad \quad \quad \text{d)} \quad \quad \quad \text{e)}
 \end{array}$$

4-17. ábra. Optimális gradiensszűrő kialakítása

- A környezeti szimmetriára vonatkozó követelmény miatt az elemeinek az y/x tengely irányában azonosnak kell lenniük (lásd 4-17/b ábra):

$$g = a, \quad h = b, \quad i = e.$$

- A vizsgált irányban akár a pozitív, akár a negatív változásokat egyformán kell érzékelnie, ezért a G_x/G_y szűrőnek az y/x tengelyre szimmetrikus elemei azonos értékűek és ellenkező előjelűek (lásd 4-17/c ábra):

$$c = -a, \quad f = -d.$$

- Homogén tartományokban a gradiens 0, ezért a szűrőelemek összegének 0-nak kell lennie, (szemben az élkiemelő és simító szűrőkkel; lásd 4-17/d ábra):

$$2b + e = 0 \rightarrow e = -2b.$$

Ha a 4-17/d ábrát összehasonlítjuk a differenciahányadosok kiszámítására szolgáló (F4-21) képlettel, szembevetően a hasonlóság, kivéve a b súlytényező hatását. Ezért kézenfekvőnek látszik a $b = 0$ választás. Hasonló következtetésre jutunk, ha a szűrő zajérzékenységének csökkentésére is törekszünk. Bebizonyítható [4-16], hogy a 4-17/d ábra szerinti konvolúciós szűrő alkalmazásával a zajspektrum σ^2 szórásnégyezete

$$\sigma'^2 = \sigma^2 |G_x^2| = \sigma^2 (4a^2 + 6b^2 + 2d^2) \quad (4-43)$$

lesz, ami nyilván akkor minimális, ha $b = 0$. Így végül is két szabad paraméterünk marad (lásd 4-17/e ábra).

$$\begin{array}{cccccc}
 G_x & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix} \\
 G_y & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} & & & & \\
 & \text{a)} & \text{b)} & \text{c)} & \text{d)} & \text{e)} & \text{f)}
 \end{array}$$

4-18. abra. Konvolúciós gradiensszűrők: a) Robert-, b) Söbel-, c) Prewitt-, d) Laplace1-, e) Laplace2-, f) Laplace3-szűrő

A gradiens közelítésére szolgáló egyik legrégebbi és egyszerűsége miatt elterjedt, bár a fentiek szerint nem optimális eszköz a Roberts-operátor [4-47]:

$$\|\nabla q(k, l)\| \approx |q(k, l) - q(k + 1, l + 1)| + |q(k + 1, l) - q(k, l + 1)|.$$

(4-44)

A megfelelő konvolúciós szűrőpárt a 4-18/a ábrán adtuk meg. Be lehet bizonyítani [4-21], hogy ez a szoban forgó 4 képpontban a képfüggvényt a legkisebb négyzetes hibával közelítő *sík normálisa*. A 2/a képre a Roberts-operátorral végzett élkitűzés eredménye a 25/a képen látható.

A gyakorlatban használt konvolúciós szűrők négyzet alakúak és 3×3 méretűek. Egyik közismert optimális szűrő *Söbeltől* származik [A-4; 271 old.] (lásd 4-18/b ábra). Bebizonyítható [4-29], hogy a $d = 2a$ választás következtében az operátor széles szögtartományban jól közelíti a gradienst (lásd 25/b kép).

Egy további optimális szűrőpár, a Prewitt operátor [4-46], a 4-18/c ábrán látható. Itt a $d = a$ választást a (4-43) szerinti szórás minimalizálása indokolja (lásd 25/c kép).

A számításigény csökkentésére kidolgozott autokorrelációs eljárás [4-26] lényege a következő:

1. Képezzünk egy $d(k, l)$ bináris képet az eredetiből a

$$d(k, l) = \begin{cases} 1, & \text{ha } |q(k, l) - q(k + \Delta k, l + \Delta l)| > \kappa_1; \\ 0, & \text{különben;} \end{cases}$$

definíció alapján. (Itt Δk és Δl önkényesen választott eltolási konstansok, κ_1 kísérleti uton meghatározott küszöb.)

2. Tekintsük élgyanusnak azokat a pontokat, amelyekben $d(k, l) = 1$ és 8 szomszédjuk közül legalább egyben 0, vagyis

$$\sum_{s=-1}^1 \sum_{t=-1}^1 d(k + s, l + t) < 9.$$

3. Az élgyanus pontokban számítsuk ki a gradienst a (4-42/a) képlet szerint, ahol G valamilyen optimális (pl. Söbel) szűrő. Élpontnak minősítjük azokat a pontokat; amelyekben a gradiens értéke nagyobb egy κ_2 küszöbnél.

Az eljárás nagyon gyors, de hibája, hogy a $(\Delta k, \Delta l)$ eltolási vektor irányába eső éleket rosszul detektálja.

Egyszerűsége miatt az élpontok kijelölésére igen közkedvelt az (F4-24), illetve az (F4-25) képlettel definiált Laplace-operátort realizáló konvolúciós szűrő (lásd 4-18/d és 4-18/e ábra). Egy harmadik változatot a 4-18/f ábrán adunk meg [A-13, 482. old.] szerint. Mint látjuk, ezek nem teljesítik az optimalitási követelményeket, és szigorúan véve nem is a gradienst közelítik, hanem a 4-15. ábrán egy dimenzióban felrajzolt *második differenciához* szélsőérték-helyeit jelölik ki. Az ábráról jól látható, hogy csak azokat a szélsőérték-helyeket szabad figyelembe venni, amelyek meghaladnak egy κ_1 küszöbértéket (ezeket \square jelöli), különben nagyon sok hamis élpontot kapnánk. Figyeljük meg továbbá, hogy minden átmeneti tartományhoz 2 élpont tartoznék, ha nem lennének tekintettel a polaritásra. Ezért a végleges élkijelöléshez csak az egyik fajta szélső értéket tartjuk meg. (Pl. az ábrán \blacksquare -tel jelölt pontok lesznek élpontok.)

A Laplace-operátor legfőbb előnye, hogy egyetlen szűrési menetben ad eredményt; semmit sem mond viszont az él irányáról. A 4-18/d, illetve 4-18/f ábra szerinti szűrővel kapott eredményeket a 25/d, illetve a 25/e képen mutatjuk.

A Laplace-operátor általánosítható (Marr-Hildreth-operátor, [4-36]), ha a szűrőelemeket az

$$a \left[1 - b \frac{m^2 + n^2}{\sigma^2} \exp\left(-\frac{m^2 + n^2}{2\sigma^2}\right) \right]$$

magfüggvény segítségével számítjuk ki. A képletben m , illetve n a szűrő x , illetve y irányu mérete, σ^2 pedig a Gauss-görbe szabadon választható szórásnégyzete. A b állandót úgy kell megválasztani, hogy a szűrőelemek összege 0 legyen, a értékét pedig úgy, hogy egész értékű szűrőelemeket kapjunk. A 4-19. ábrán 2 különböző σ értékhez tartozó szűrőt adunk meg [4-22] alapján, az $m = n = 11$ esetre. (Az elemek jellegzetes eloszlása miatt ezt "burkalap" szűrőnek is nevezik.)

A szűrő alkalmazása esetén az élpontokat ott kell kijelölni, ahol a második differenciához negatív meredekségű nullaátmeneteli vannak, feltéve, hogy az esés értéke nagyobb egy κ_2 küszöbnél. (A 4-15. ábrán a $\kappa_2 = 3$ esetnek megfelelő helyeket σ , a megfelelő élpontokat \bullet jelzi.)

A szűrést elvégeztük a $\sigma = 5$ "burkalap" szűrővel is; az eredmény a 25/f képen látható.

4.4.1.2 Iránykereső módszerek

Az iránykereső módszerek különböző irányokban vizsgálják az átmeneti görbe meredekségét, és ott jelölnék ki élpontokat, ahol a meredekséggörbének (egy vagy több irányban) maximuma van.

Az eljárások egyik csoportját *szűrőillesztésnek* nevezzük. Ehhez 8 konvolúciós szűrőre van szükség, amelyekkel a vizsgált pont körüli 8 első differenciához számítjuk ki. A (k, l) pont akkor élgyanus, ha teljesül a

$$\max_j (q(k, l) \otimes M_j) > \kappa \quad (j = 0, 1, \dots, 7) \quad (4-45/a)$$

feltétel, ahol κ előre adott küszöb. Az él irányát ekkora

$$\begin{bmatrix}
 0 & 0 & 0 & -1 & -1 & -2 & -1 & -1 & 0 & 0 & 0 \\
 0 & 0 & -2 & -4 & -8 & -9 & -8 & -4 & -2 & 0 & 0 \\
 0 & -2 & -7 & -15 & -22 & -23 & -22 & -15 & -7 & -2 & 0 \\
 -1 & -4 & -15 & -24 & -14 & -1 & -14 & -24 & -15 & -4 & -1 \\
 -1 & -8 & -22 & -14 & 52 & 103 & 52 & -14 & -22 & -8 & -1 \\
 -2 & -9 & -23 & -1 & 103 & 178 & 103 & -1 & -23 & -9 & -2 \\
 -1 & -8 & -22 & -14 & 52 & 103 & 52 & -14 & -22 & -8 & -1 \\
 -1 & -4 & -15 & -24 & -14 & -1 & -14 & -24 & -15 & -4 & -1 \\
 0 & -2 & -7 & -15 & -22 & -23 & -22 & -15 & -7 & -2 & 0 \\
 0 & 0 & -2 & -4 & -8 & -9 & -8 & -4 & -2 & 0 & 0 \\
 0 & 0 & 0 & -1 & -1 & -2 & -1 & -1 & 0 & 0 & 0
 \end{bmatrix}
 \quad \sigma = 1,4$$

$$\begin{bmatrix}
 -24 & -21 & -17 & -13 & -10 & -10 & -10 & -13 & -17 & -21 & -24 \\
 -21 & -16 & -10 & -4 & 0 & 2 & 0 & -4 & -10 & -16 & -21 \\
 -17 & -10 & -1 & 6 & 11 & 13 & 11 & 6 & -1 & -10 & -17 \\
 -13 & -4 & 6 & 15 & 22 & 24 & 22 & 15 & 6 & -4 & -13 \\
 -10 & 0 & 11 & 22 & 29 & 31 & 29 & 22 & 11 & 0 & -10 \\
 -10 & 2 & 13 & 24 & 31 & 34 & 31 & 24 & 13 & 2 & -10 \\
 -10 & 0 & 11 & 22 & 29 & 31 & 29 & 22 & 11 & 0 & -10 \\
 -13 & -4 & 6 & 15 & 22 & 24 & 22 & 15 & 6 & -4 & -13 \\
 -17 & -10 & -1 & 6 & 11 & 13 & 11 & 6 & -1 & -10 & -17 \\
 -21 & -16 & -10 & -4 & 0 & 2 & 0 & -4 & -10 & -16 & -21 \\
 -24 & -21 & -17 & -13 & -10 & -10 & -10 & -13 & -17 & -21 & -24
 \end{bmatrix}
 \quad \sigma = 5$$

4-19. ábra. Marr-Hildredth ("bur kalap") szűrők

$$\vartheta \approx (j \pm 0,5) \frac{\pi}{4} \quad (4-45/b)$$

Összefüggés adja.

A 4-20/a és 4-20/b ábrán Robinson, illetve Kirsch nyomán két szűrősorozatot adunk meg ([4-48], [4-28]); az előbbit "iránytű" (compass) szűrőnek is nevezik. A 2/a képre

Robinson-
("iránytű")
szűrők

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}
 \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -1 \\ 1 & -1 & -1 \end{bmatrix}
 \begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & -1 \end{bmatrix}
 \begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & 1 \end{bmatrix}
 \begin{bmatrix} -1 & -1 & -1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}
 \begin{bmatrix} -1 & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}
 \begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}
 \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$$

a)

Kirsch-
szűrők

$$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}
 \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}
 \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}
 \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}
 \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}
 \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}
 \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}
 \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$$

M_0 M_1 M_2 M_3 M_4 M_5 M_6 M_7

b)

4-20. ábra. Iránykereső szűrősorozatok

ezekkel meghatározott élképek a 26/a, illetve 26/b képen láthatók.

Hasonló sorozatokat lehet kialakítani a 4-18/b és a 4-18/c ábrán bemutatott Söbel-, illetve Prewitt-operátorból.

Az elvet rendkívül érdekes módon általánosították, amit "élbázisvektorok" módszerének nevezhetünk [4-17]:

1. Tekintsük a vizsgált képpont $n \times n$ -es környezetében levő világosságkódokat az n^2 -dimenziós S sajáttság-térben egy helyvektor összetevőinek, amelyet egy $n \times n$ -es mátrixszal ábrázolunk. Pl. $n = 3$ esetén:

$$\underline{Q} = \begin{bmatrix} q(k-1, l-1) & q(k, l-1) & q(k-1, l+1) \\ q(k, l-1) & q(k, l) & q(k, l+1) \\ q(k+1, -1) & q(k+1, l) & q(k+1, l+1) \end{bmatrix} \equiv \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

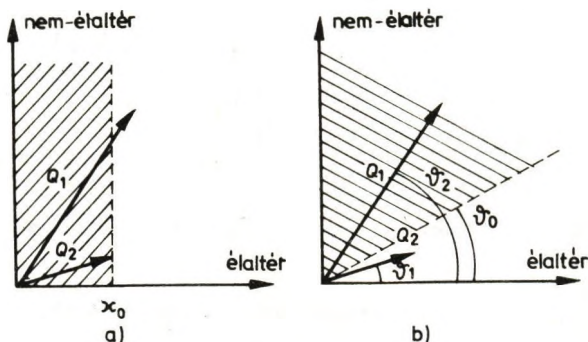
2. Értelmezzük továbbá két vektor skalárszorzatát az S térben a konvolúciós szorzatösszeggel.
3. Bontsuk fel az S sajátságteret "élaltérre" és "nem-élaltérre". Az előbbit feszítse ki e db ortogonális "élbázisvektor": $\{\underline{E}_1, \underline{E}_2, \dots, \underline{E}_e\}$, az utóbbit $n^2 - e$ db "nem-élbázisvektor": $\{\underline{E}_{e+1}, \underline{E}_{e+2}, \dots, \underline{E}_{n^2}\}$.
4. Határozzuk meg a \underline{Q} vektor vetületét az élaltérre, illetve hosszát:

$$|\underline{Q}_e|^2 = \sum_{j=1}^e (\underline{Q} \otimes \underline{E}_j)^2, \quad |\underline{Q}|^2 = \sum_{j=1}^{n^2} (\underline{Q} \otimes \underline{E}_j)^2 = (\underline{Q} \otimes \underline{Q}).$$

5. Az élaltérbe eső vetületet jellemezhetjük a $|\underline{Q}_e|$ "élhosszal" és a

$$\vartheta = \arccos \left(\frac{|\underline{Q}_e|}{|\underline{Q}|} \right) \quad (0 \leq \vartheta \leq \pi)$$

"élszöggel". Nyilvánvaló, hogy minél kisebb ez a szög, annál jobban illeszkedik \underline{Q} az éalattérre és fordítva. A döntést egy ϑ_0 küszöb alapján hozzuk.



4-21. ábra. Élkitűzés elve: a) küszöb, b) "élszög" alapján

Az elvet a 4-21. ábrán szemléltetjük. Legyen két sajátvektorunk, \underline{Q}_1 és \underline{Q}_2 . A mintaillesztési módszer alkalmazása esetén a küszöböt egy állandó "élhossz" (κ_0), a mi esetünkben egy állandó "élszög" (ϑ_0) jelenti. Az előbbi esetben a \underline{Q}_1 vektorral jellemzett pontot élpontnak minősítjük, a \underline{Q}_2 -vel jellemzettet nem (4-22/a ábra). Az élbázisvektorok módszerével éppen fordított döntést hozunk, mivel \underline{Q}_1 -nek nagy a vetülete a nem-éalattérre, \underline{Q}_2 -nek viszont az éalattérre (4-21/b ábra).

A 4-22. ábrán egy lehetséges ortogonális bázisvektorsorozatot ("Frei-szűrő") adunk meg $n = 3$ esetre, [4-17] alapján. Ezek felhasználásával egy adott altérbe tartozó képpontokat a

$$\sum_{j=m}^{m+1} (\underline{Q} \otimes \underline{B}_j)^2 |\underline{Q}|^{-2} < \vartheta_0 \quad (4-46)$$

élpontaltér			élvonalaltér			
$\begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & \sqrt{2} \\ 1 & 0 & -1 \\ -\sqrt{2} & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix}$			
$m = 1$	$m = 3$	$m = 5$	$m = 7$			
$\begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -\sqrt{2} \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 2 & -1 & 2 \\ -1 & -4 & -1 \\ 2 & -1 & 2 \end{bmatrix}$			$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \frac{1}{9}$
$m = 9$						
izotrop átlag	"fodor"	vonal	Laplace	átlag		

4-22. ábra. Az S-tér egy lehetséges ortogonális bázisvektorrendszere

összefüggés alapján jelölhetjük ki. A 26/c képen az $m = 1$ esetnek megfelelő élképet látjuk. Ha ezt összehasonlítjuk a Robinson-, illetve a Kirsch-szűrővel nyert élképekkel (26/a, illetve 26/b kép), meg lehet állapítani, hogy finomabb élrészleteket is felfed, mint a másik módszerek. Ezen kívül az "erős" élek vékonyabbak, és a sötétebb háttérű területeken is jó a határfok. Hátránya a nagyobb zajérzékenység, amit előzetes zajszűréssel lehet kompenzálni.

Más elvű, de gyors és hatékony a *főirányszűrők* módszere [4-32]. Eszerint végigjárjuk a képet a 4 főirányban ($x, y, p = 45^\circ$ és $n = -45^\circ$), és megvizsgáljuk az aktuális pont 5×5 -ös környezetében a meredekséggörbét. A főirányszűrőket - a szokásostól eltérően - szimmetrikus differenciák definiálják a

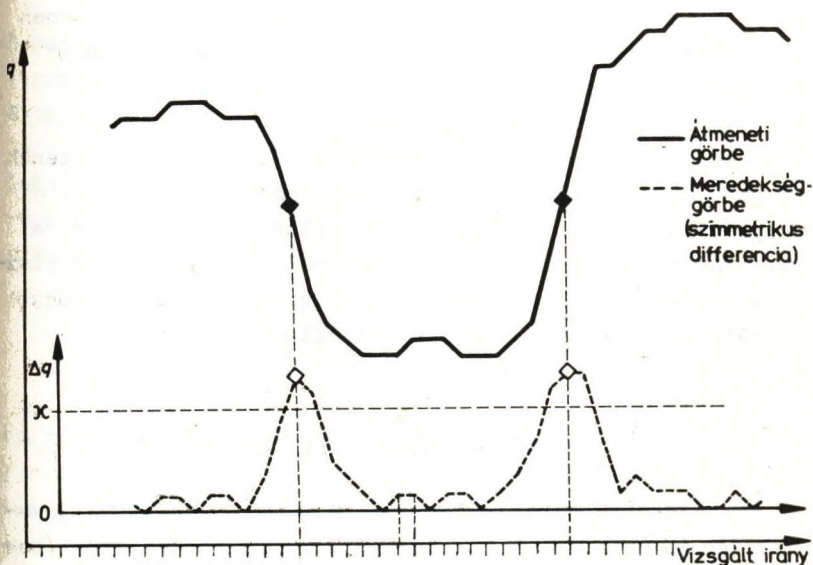
$$\Delta_x(k, l) = |q(k-1, l) - q(k+1, l)|; \quad (4-47/a)$$

$$\Delta_y(k, l) = |q(k, l-1) - q(k, l+1)|; \quad (4-47/b)$$

$$\Delta_p(k, l) = |q(k + 1, l - 1) - q(k - 1, l + 1)|; \quad (4-47/c)$$

$$\Delta_n(k, l) = |q(k - 1, l - 1) - q(k + 1, l + 1)| \quad (4-47/d)$$

képletek szerint. A 4-23. ábrán az egyik főirányszűrő hatását mutatjuk; itt az átmeneti függvény azonos a 4-15. ábrán megadottal.



4-23. ábra. Élkitűzés elve a főirányszűrők segítségével

- Egy képpontot akkor minősítünk élgyanusnak, ha ott
- a differenciagörbének *legalább két* irányban maximuma van, és
 - a maximumok értéke *nagyobb* egy κ küszöbnél.

(Ez elvileg az átmeneti görbe egy elég meredek szakaszának inflexiós pontját jelenti.)

Mivel a differenciagörbe is digitális jellegű, a maximumot egy kicsé tágabban értelmezzük a szokásosnál [4-2]. A 4-23. ábrán az elfogadott maximumot \diamond , a megtalált élpontokat \blacklozenge jelöli; ezt a jelölést - az összehasonlítás kedvéért - a 4-15. ábrára is átvezettük. Ezzel a módszerrel nyert élképet láthatunk a 4-26/d képen.

Az eddigieknél jobb élfolytonosságot és kisebb zajérzékenységet eredményez a fazettamodell (lásd 4.2.2.1 alpont) továbbfejlesztésén alapuló módszer [4-22]. Ebben egy fazetta felett a képfüggvényt harmadfoku, diszkrét ortogonális polinomokkal közelítik; ezek segítségével állítják elő a kívánt irány szerinti differenciahányadosokat. Egy képpontot az eljárás során akkor minősítenek élpontnak, ha ott a *gradiens irányu* második differenciahányadosnak *negatív* meredekségű nullaátmenete van, és a változás értéke *nagyobb* egy küszöbnél. A küszöböt statisztikai becsléssel, az élek előfordulásának (helytől függő) feltételes valószínűsége alapján állapítják meg.

4.4.1.3 Vonalkereső módszerek

Az eddigiek során olyan módszereket vizsgáltunk, amelyek az éleket pontokból rakják össze. A továbbiakban bemutatunk néhány módszert, amely az éleket jól közelítő egyenesszakaszok megkeresésére irányul, s ezekből fűzi össze a végleges élképet. Ezek átmenetet képeznek a lokális és a regionális eljárások között, mivel az aktuális képpont vizsgált környezete általában elég nagy (7×7, illetve 9×9 képpont).

A *Hükel-féle* "klasszikus" vonalkereső eljárás egy *ötparaméteres, ideális élfüggvényből* (lépcsőből) indul ki [4-25]:

$$e(x, y, \alpha, \beta, \rho, \gamma, \delta) = \begin{cases} \gamma, & \text{ha } \alpha x + \beta y \leq \rho; \\ \gamma + \delta, & \text{ha } \alpha x + \beta y > \rho. \end{cases}$$

Jelölje a képfüggvényt $f(x, y)$, akkor a feladat az 5 paraméter meghatározása úgy, hogy a

$$h(\alpha, \beta, \rho, \gamma, \delta) = \int_A [f(x, y) - e(x, y, \alpha, \beta, \rho, \gamma, \delta)]^2 dx dy \quad (4-48)$$

hibafüggvény *minimális* legyen. Itt A a képsík vizsgált tartománya, amelyben az $\alpha x + \beta y = \rho$ egyenes felett $e(x, y)$ értéke $\gamma + \delta$, alatta pedig γ .

A paraméterek kiszámítása céljából az A tartomány felett mindkét függvényt felbontják egy *ortonormált bázisfüggvényrendszer* szerint. Ezzel a (4-48) kifejezést vizsgálva lehet vezetni a

$$h(\alpha, \beta, \rho, \gamma, \delta) = \sum_{i=0}^r [a_i - b_i(\alpha, \beta, \rho, \gamma, \delta)]^2 \quad (4-49)$$

végtelen összeg minimalizálására, ahol a_i , illetve b_i az $f(x, y)$, illetve az $e(x, y)$ függvény függvénytörvényének együtthatói. (Elvileg $r = \infty$, a gyakorlatban csak a sor néhány első tagját számítják ki.)

A feladat megoldásához a polárkoordinátás *Fourier-bázisfüggvényeket* választották, ami azzal a számítástechnikai hátránnyal jár, hogy az A tartományt *kör alakúnak* kell venni. Az *algoritmust* nagy vonalakban így lehet megfogalmazni (*Hückel-operátor*):

1. Fedjük le a képsíkot hézagmentesen (közel) kör alakú (átfedő) ablakokkal.
2. A képfüggvényt bontsuk fel minden ablak fölött 9 bázisfüggvény szerint.

3. Határozzuk meg minden ablakban a legjobban illeszkedő élfüggvényt (4-49) alapján.
4. Ha az illeszkedés elég jó, a δ paraméter elég nagy, és az éldarab nem túl hosszú, fogadjuk el azt az adott ablak felett a valódi él közelítésének.

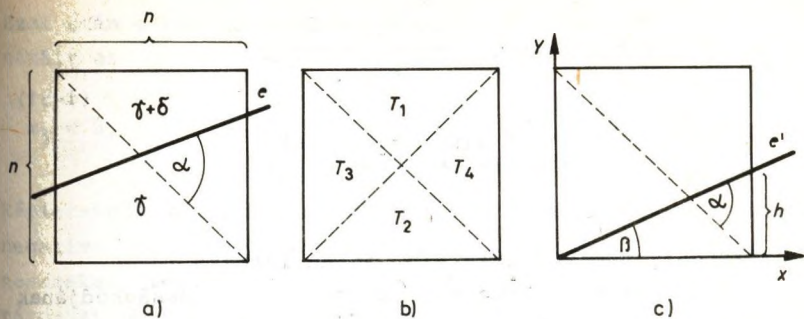
Az eljárás meglehetősen *számításigényes*, és további gondot okoz a megtalált éldarabok összefűzése. Ezen kívül bebizonyították [4-52], hogy a módszer bizonyos ritkán előforduló esetekben *hibás döntéshez* vezethet. Nevezetesen előfordulhat, hogy a (4-49) minimalizálásával komplex paramétereket kapunk, amelyek nem értelmezhetők; továbbá lehetséges az is, hogy a megtalált él kívül esik a vizsgált tartományon. Bár ezeket a problémákat ki lehet védeni a konkrét programozás során, a módszer továbbfejlesztése nem látszik indokoltnak.

Egy másik eljárás, a *Mérő-Vassy-operátor*, azon a felismerésen alapul, hogy ha egy ablakban létezik határvonal, akkor ennek iránya *egyértelműen meghatározható* a világosságkódoknak az ablakon belüli *eloszlásából* [4-37].

Vizsgáljunk $n \times n$ méretű négyzet alakú ablakokat. Az ablak középpontja a szokásos módon illeszkedjék az aktuális (k, l) képkoordinátájú képpontra; az egyszerűbb írásmód kedvéért azonban helyezzük a helyi koordináta-rendszer kezdőpontját az ablak bal alsó sarokpontjába. Tegyük fel, hogy az ablakon áthalad egy e él, amely felett a képpontok világosságkódja $\gamma + \delta$ (illetve γ), alatta pedig γ (illetve $\gamma + \delta$) (lásd 4-24/a ábra). Ekkor az e egyenesnek a főátlóval bezárt α szögére felírhatjuk:

$$\operatorname{tg} \alpha = \frac{\tilde{q}(k, l) \otimes A}{\tilde{q}(k, l) \otimes B}. \quad (4-50)$$

A képletben A , illetve B a 4-25. ábrán megadott két



4-24. ábra. Élvonaldarabok kijelölése a világosságkód-
eloszlás alapján

eloszlásmaszk; $\tilde{q}(k, l)$ az ablakba eső képpontok halmaza;

⊗ pedig a (F4-12)-ben definiált konvolúciós összegzést
jelenti.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 1 \end{bmatrix}$$

4-25. ábra. Eloszlásmaszkok

(A bizonyítást lásd [4-37]-ben. Megjegyezzük, hogy sokkal jobb eredményeket kaptunk, ha a főátlóra, illetve a mellékátlóra eső képpontokat $+1/2$, illetve $-1/2$ súllyal vettük figyelembe az átló feletti, illetve alatti tartományok kiértékelésénél.)

Ha az ablakba eső képrészlet nem tekinthető kétszintűnek, használjuk a

$$\gamma + \delta = \text{ent}\left(\frac{4(\max_j(\tilde{q}_j) + 2n)}{n^2}\right), \quad (4-51)$$

$$\gamma = \text{ent}\left(\frac{4(\min_j(\tilde{q}_j) + 2n)}{n^2}\right)$$

közelítést, ahol \tilde{q}_j a 4-24/b ábrán látható T_j ($j = 1, 2, 3, 4$) területbe eső képpontok világosságkódjának összege. (Itt is célszerű az összegzésnél a határvonalakra eső képpontok világosságkódjának csak a felével (a középpont esetében negyedével) számolni.)

A (4-50)-nel megadott irányu és a helyi origón áthaladó egyenesre (4-24/c ábra):

$$\text{tg } \beta = \text{tg}(\alpha - 45^\circ) = \frac{h}{n},$$

amiből elemi átalakításokkal következik, hogy

$$h = n \frac{\text{tg } \alpha - 1}{1 + \text{tg } \alpha}. \quad (4-52)$$

Számítsuk még ki az

$$m = \frac{\sum_{s=1}^n \sum_{t=1}^n q(s, t) - \gamma n^2}{n\delta} = \frac{c}{n} \quad (4-53)$$

mennyiséget, ahol c a $\gamma + \delta$ világosságkódú képpontok száma az ablakban. (Szemléletesen $n \rightarrow m$ azt jelenti, hogy ha az e egyenes vízszintes volna, milyen magasságban metszené az ablakot.)

Ezek után a keresett éldarab végpontjainak elvi koordinátáit az

$$x_1 = 0, y_1 = m - \frac{h}{2}; \quad x_2 = n, y_2 = m + \frac{h}{2} \quad (4-54)$$

képletekből számíthatjuk ki. (Ha valamelyik y -koordináta negatív, vagy n -nél nagyobb volna, az ablakba eső egyenesszakasz végpontjainak koordinátáit egyszerűen meghatározhatjuk lineáris interpolációval.)

Végezetül az éldarabnak az ablakon belüli pontos helyét a következő döntések alapján jelöljük ki:

$$\begin{aligned} \operatorname{tg} \alpha \geq 0 & \begin{cases} \tilde{q}_1 \geq \tilde{q}_2 \rightarrow \text{változatlan;} \\ \tilde{q}_1 < \tilde{q}_2 \rightarrow \text{tükrözés a középpontra;} \end{cases} \\ \operatorname{tg} \alpha < 0 & \begin{cases} \tilde{q}_3 \geq \tilde{q}_4 \rightarrow \text{tükrözés a főátlóra;} \\ \tilde{q}_3 < \tilde{q}_4 \rightarrow \text{tükrözés a mellékátlóra.} \end{cases} \end{aligned} \quad (4-55)$$

A Mérő-Vassy-operátor sokkal gyorsabb és valódi élek esetében legalább olyan pontos, mint a Hückel-operátor. Hibája, hogy mindig *egy* éldarabot talál: akkor is, ha valódi él nem halad át az ablakon, illetve ha az áthaladó élvonalnak töréspontja vagy több élvonal metszéspontja esik az ablakba. Az első esetet egyszerű kijavítani úgy, hogy élszakaszt csak a $\delta > \kappa$ esetben jelölünk ki, ahol κ előre adott küszöb. A második és a harmadik esetet csak élkorrekciónal lehet kiigazítani.

Az eddig tárgyalt eljárások *nem* alkalmasak a V -típusú élek felderítésére. Erre a 4.2.2.1 alpontban ismertetett *fazettamodell* alapján kínálkozik lehetőség [4-21]. Mint láttuk, a modell lényege az, hogy a képsíkot hézagmentesen

illeszkedő, diszjunkt, sokszög (ez esetben négyzet) alakú területekre, (fazettákra) bontják, és az egyes fazetták felett a képfüggvényt a legkisebb négyzetes hibával illeszkedő ferde sikkal ("palatetővel") helyettesítik. (Ideális esetben minden képpont egy fazetta középpontja.) Az egyes fazetták feletti közelítő sík egyenletét a (4-33), gradiensét a (4-34) összefüggések definiálják. Ezek után két szomszédos fazetta határvonalát akkor tekintjük közelítő éldarabnak, ha gradiensük egy küszöbnél nagyobb mértékben különbözik egymástól. (Természetesen mind a 4 szomszédot megvizsgáljuk.) Mivel a képek mindig zajosak, a küszöbértéket [4-21]-ben szignifikancianalízissel állapítják meg, amihez F-statisztikát konstruálnak.

Megjegyezzük még, hogy éldarabokat az előző alpontban leírt szűrőillesztési eljárással is kereshetünk a (4-45) képletek alapján. Erre a célra használhatók pl. a 4-26. ábrán megadott szűrők.

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

4-26. ábra. Élvonal-kijelölő szűrősorozat

Hasonló eredményhez jutunk, ha a (4-46) kifejezést $m = 5$ mellett értékeljük ki, vagyis meghatározzuk a képsajátságvektornak a "vonaltérbe" eső vetületét.

4.4.1.4 Élkijelölés texturaelemzéssel

A "texturaélek" a texturában található hirtelen (ugrásszerű) változások mértani helyei. Az alábbiakban két, a texturális jellemzők elemzésén alapuló élkijelölő eljárás-

rást ismertetünk. (A texturaelemzés vázlatos áttekintését lásd az 5.4 pontban.)

A textura vizsgálatára gyakorlatban a *kétdimenziós*, képpont-, lokális minta-, illetve tartományalapú modellek terjedtek el [4-19]. Ezek segítségével főleg orvosi és geológiai felvételek elemezhetők. Más célokra a texturát *térbeli* felületelemekként értelmezik, amelyeket "egyéni" visszaverő képességükkel és irányítottágukkal definiálnak.

A *képpontalapú* modelleket időbeli vagy véletlenszerű mintavételezéssel előállított képek elemzésére használják; a természetes képekre inkább a *lokális minta-* és a *tartományalapú* modellek alkalmazhatók. Az utóbbiak a kép mozaikként kezelik: minden (általában konvex) *mozaikcellához* tartozik valamilyen sajátság(vektor) és előfordulási valószínűség.

Élkitűzésre egy *linedris tartományalapú* modellre épülő eljárást ismertetünk [4-11], amelyben csak a cellák sorirányu szomszédainak világosságkódját vizsgálják. (A matematikai egyszerűségeen kívül ezt még az is indokolja, hogy a képpontokat rendszerint sorfolytonosan lehet elérni.) Az élkijelölő operátor a (sorirányban) szimmetrikusan szomszédos cellák átlagértékének különbségét méri az

$$e_m(i) = \frac{1}{m} \sum_{s=1}^m [q(i-s) - q(i+s)]$$

definíciós képlet alapján. Az eljárás három lépésből áll:

1. Minden tartományra meghatározzák az $e_m(i)$ értékét; (a $2m + 1$ cellaméret a hibás éldetektálás valószínűségét befolyásolja; a számítások szerint $5 \leq m \leq 16$ választás az optimális).
2. Törölnek minden cellát, amelyre $|e_m(i)| < \kappa_1$.

3. Törölnek minden cellát, amelyik nála nagyobb él-értékű cella környezetében van, vagyis

$$|e_m(i)| < |e_m(i+j)|;$$

ahol $|j| < \kappa_2$. (Ezt az utóbbi lépést *álmáximum el-
nuomásnak* (nonmaxima suppression) nevezik.)

Az eljárás hatékonysága az m paraméter és a κ_1, κ_2 küszöb megfelelő megválasztásától függ.

Egy másik eljárásban a textura első- és másodrendű *statisztikai jellemzőit* használják az élek megkeresésére [4-43]. A hamis éldetektálás csökkentése érdekében a képen először mediánszűrést végeznek, majd minden képpontban meghatározzák a gradienst a Prewitt-szűrővel (lásd 4-18/c ábra), amelyet - a 4-20. ábra mintájára - minden pont körül "körbeforgatnak" a maximumot adó irány megkeresése céljából.

Ezután "*antiparalel élpárok*" keresnek. Egy élpont antiparalel párja a gradiensének irányában, adott távolságon belül található, az adott iránnyal $j \cdot 45^\circ$ ($j = 3, 4, 5$) szöget bezáró másik élpont. Minden élpárra kiszámítják a következő, *elsőrendű* statisztikai sajátságokat:

- távolság ($1 \leq d \leq 10$);
- az összekötő egyenes mentén elhelyezkedő képpontok átlagos világosságkódja (μ);
- ugyanezek szórása (σ);
- a két gradiens iránya közti különbség abszolút értéke ($\delta \in \{0^\circ, 45^\circ\}$);
- a két pont világosságkódja különbségének abszolút értéke (Δq).

(Ezeknél a számításoknál figyelembe veszik, hogy az élpárok a "világos" vagy a "sötét" oldalakra illeszkednek-e.) Ezekon kívül még *másodrendű* statisztikai sajátságokat is

kiszámítanak. Minden élpár két végpontjában olyan másodrendű felülettel közelítik a képfüggvényt, amelyik a legkisebb négyzetes hibával illeszkedik a képpontok 5×5-ös környezetére. A közelítő függvény parciális differenciálhányadosai segítségével ismét meghatározzák a gradiens nagyságát és irányát, valamint a görbületet a

$$|\nabla f| = \sqrt{f_x'^2 + f_y'^2}; \quad \operatorname{tg} \vartheta = \frac{f_y'}{f_x'}$$

$$\frac{1}{\bar{Y}} = \frac{f_{xx}'' f_y'^2 + f_{yy}'' f_x'^2 - 2 f_{xy}'' f_x' f_y'}{(f_x' + f_y')^{3/2}}$$

képletekből. Ennek alapján álmaximum-elyomást végeznek a gradiens irányára merőlegesen 2-2 képpont távolságig. Az éleket azok a megmaradó antiparalel élpárok alkotják, amelyeknek a fentiek szerint számított sajáttségvektorai elég kis távolságra vannak egymástól.

Megjegyezzük, hogy a cikkben ismertetett eredmények szerint a másodrendű statisztikák figyelembevétele nem hozott észrevehető javulást az élképben.

4.4.2 ÉL KORREKCIÓ

A tapasztalat azt mutatja, hogy nem lehet vizuálisan jónak elfogadható élképet kijelölni pusztán a képsajátságok elemzésével. Kielégítő eredményeket csak az élponatok, illetve az éldarabok egymáshoz viszonyított helyzetének vizsgálatával, az élkijelölési eljárások eredményeinek korrigálásával lehet elérni.

Az *élkorrekció* célja folytonos és "kellően sima" élek, illetve (zárt) konturvonalak előállítás. (Az utób-

biak felelnek meg a képen található foltok határvonalainak.) A célt a szakadások megszüntetésével (vagyis az egymást kiegészítő éldarabok összekötésével), illetve a felesleges (hamis) éldarabok törlésével lehet elérni.

4.4.2.1 Analitikus módszerek

Az analitikus módszerek alapelve az, hogy az éleket a megtalált *élpontokra illeszkedő digitális görbék* segítségével alakítják ki. A gyakorlatban - a számítások egyszerűsítése érdekében - a közelítés legtöbbször szakaszosan *linedris*, vagyis görbék helyett *tört vonalakkal* történik.

Mivel ezek a módszerek nem képesek a hamis élpontok felfedezésére, az élkijelölés során a küszöb(ke)t olyan magasra kell választani, hogy ilyenek lehetőleg ne keletkezzenek.

A *polinomos közelítés* alapja az az algebrai tétel, hogy mindig létezik olyan $(r + 1)$ -ed fokú polinom, amely r számú pontra illeszkedik.

Az együtthatók meghatározása azonban r növekedésével egyre bonyolultabbá válik; ráadásul a magasabb fokszámu görbék az illeszkedési pontok között rendszerint "roszszul" viselkednek.

A nehézségek elkerülésére a leggyakrabban a "*legkisebb négyzetes hibájú*" (LKNH) módszert alkalmazzák. Ennek lényege, hogy a polinom m fokszáma kisebb, mint a pontok r darabszáma, s így az együtthatók tulhatározottá válnak. Mivel most nem lehet elérni, hogy a görbe minden adott pontra pontosan illeszkedjék, ehelyett meghatározzák az egyes pontok távolságát a közelítő görbétől, s ezek *négyzetösszegét minimalizálják*. Legyen a közelítendő pontok halmaza $\{k_j, l_j\}$ ($j = 1, 2, \dots, r$), akkor a feladatot az

$$\varepsilon = \sum_{j=1}^r [(a_0 + a_1 k_j + a_2 k_j^2 + \dots + a_m k_j^m) - l_j]^2 = \text{minimum} \quad (4-56)$$

Összefüggéssel definiálhatjuk. Legyen a továbbiakban $n = m+1$ és vezessük be a kényelmesebb mátrixjelöléseket:

$$\underline{K} = \begin{bmatrix} 1 & k_1 & k_1^2 & \dots & k_1^m \\ 1 & k_2 & k_2^2 & \dots & k_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & k_r & k_r^2 & \dots & k_r^m \end{bmatrix}; \quad \underline{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}; \quad \underline{l} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_r \end{bmatrix} \quad (4-57)$$

Ezekkel és az (F3-8) euklidészi normafüggvénnyel (4-56)-ot a

$$\|\underline{K} \underline{a} - \underline{l}\|^2 = \text{minimum} \quad (4-58)$$

alakban írhatjuk. A problémát az okozza, hogy a

$$\underline{K} \underline{a} = \underline{l} \quad (4-59)$$

egyenletrendszert közvetlenül nem tudjuk megoldani, mivel - a túlhatározottság miatt - \underline{K} $r \times n$ -es téglalapmátrix ($r > n$). Szorozzuk azonban (4-59) mindkét oldalát balról \underline{K} transzponáltjával:

$$\underline{K}^T \underline{K} \underline{a} = \underline{K}^T \underline{l}. \quad (4-60)$$

Ekkor $\underline{K}^T \underline{K}$ már $n \times n$ -es négyzetes mátrix, s ha - mint általában - nem szinguláris, akkor invertálható. Ezzel a (4-58) minimumfeladat megoldását visszavezettük az \underline{a} , együtthatókra nézve elsőfoku, $m+1$ ismeretlenes (4-60) egyenletrendszer megoldására. A megoldás pl. az

$$\underline{a} = (\underline{K}^T \underline{K})^{-1} \underline{K}^T \underline{z} \quad (4-61)$$

alakban írhatjuk fel; de más módszerekkel, pl. Gauss-Seidel-iterációval is meghatározhatjuk.

Nézzük meg kissé részletesebben a gyakorlatban legfontosabb *linedíris* esetet. Ekkor

$$\underline{K} = \begin{bmatrix} 1 & k_1 \\ 1 & k_2 \\ \vdots & \vdots \\ 1 & k_r \end{bmatrix} \quad \underline{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}; \quad \underline{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{bmatrix} \quad (4-57/a)$$

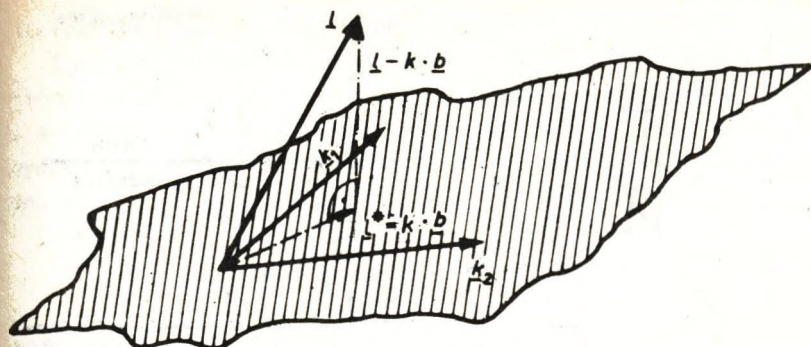
és (4-60)-at kis számolással a

$$\begin{bmatrix} \sum_{i=1}^r 1 & \sum_{i=1}^r k_i \\ \sum_{i=1}^r k_i & \sum_{i=1}^r k_i^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^r z_i \\ \sum_{i=1}^r k_i z_i \end{bmatrix} \quad (4-60/a)$$

alakra hozhatjuk, ami a_0, a_1 -re nézve kétismeretlenes elsőfoku egyenletrendszer. Könnyen belátható [A-4; 328.old.], hogy (4-60/a) megoldásai teljesítik a (4-56) minimumfeltételt.

A LKNH-módszernek szemléletes geometriai tartalma is van. Az ábrázolhatóság kedvéért álljon a \underline{K} mátrix két (k_1, k_2) 3D oszlopvektorból, amelyek a 4-27. ábrán látható síkor feszítik ki. Mindazok az \underline{z}^* vektorok, amelyek előállíthatók a bázisvektorok lineáris kombinációjaként

$$\sum_{j=1}^2 b_j \underline{k}_j = \underline{z}^*, \quad \text{illetve} \quad \underline{K} \underline{b} = \underline{z}^* \quad (4-60/b)$$



4-27. ábra. A LNKH közelítés szemleltetése

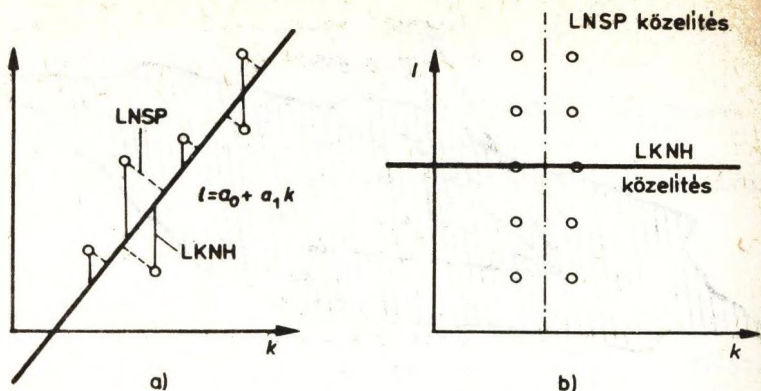
alakban írhatók. Itt $\underline{b}^T = (b_1, b_2)$ alkalmasan választott súlyvektor, és az \underline{l}^* vektor a $\{k_1, k_2\}$ síkban van. Ha azonban az \underline{l} kívül esik a síkra, a legjobb közelítését akkor kapjuk, ha az $\underline{l} - \underline{l}^*$ különbségvektor merőleges a síkra, illetve az azt kifeszítő \underline{k}_1 és \underline{k}_2 vektorra. Ezt a (4-60/b) felhasználásával a

$$\underline{K}^T (\underline{l} - \underline{K} \underline{b}) = 0$$

mátrixegyenlet fejezi ki, ami (4-60)-nal azonos. Vagyis a legjobb közelítést adó \underline{l}^* vektor négyzetes eltérése \underline{l} -től minimális.

A módszer súlyos elvi hibája, hogy a távolság mértékegysége függ a koordináta-rendszer megválasztásától. A 4-28/a ábrán látható pontok esetében pl. a legjobban illeszkedő egyenes meghatározásához a távolságokat a LKNH-módszer esetében az l tengellyel párhuzamosan mértük. Szélsőséges esetben (lásd 4-28/b ábra) így a lehetséges legrosszabb közelítést kapjuk.

Nyilvánvaló, hogy megbízhatóbb a közelítés, ha a távolságokat a közelítő görbére merőlegesen mérjük (lásd



4-28. ábra. A LKMH és a LNSP közelítés hibája:
 a) a távolságmérés elve, b) a legrosszabb eset

4-28/a ábra szaggatott vonalakat), ezt a *legnagyobb sajátvektorral párhuzamos* (LNSP) illesztésnek nevezzük. Nem nehéz belátni (lásd [A-4] 333. old.), hogy ez a vonal át-
 megy a ponthalmaz súlypontján. A közelítéshez meg kell
 határoznunk a ponthalmaz normalizált *szórásmatrixát*,
 amelyet az $\underline{M} = (\underline{k}^T, \underline{l}^T)$ koordinátamatrixból az

$$\underline{\Omega} = \underline{M}^T \underline{M} = \begin{bmatrix} \sum_{j=1}^r (k_j - k_0)^2 & \sum_{j=1}^r (k_j - k_0)(l_j - l_0) \\ \sum_{j=1}^r (k_j - k_0)(l_j - l_0) & \sum_{j=1}^r (l_j - l_0)^2 \end{bmatrix} \quad (4-62)$$

összefüggéssel számíthatunk ki, ahol $k_0 = \frac{1}{r} \sum_{j=1}^r k_j$, illetve
 $l_0 = \frac{1}{r} \sum_{j=1}^r l_j$ a súlypont koordinátái. A legjobb LNSP-köze-
 litő egyenes irányát $\underline{\Omega}$ legnagyobb sajátvektorának iránya
 jelöli ki. (Ilyen pl. a 4-28/b ábrán a pontvonallal raj-
 zolt egyenes. Megjegyezzük, hogy a koordinátatengelyek

lcserélése esetén ez egybeesik a LKNH közelítéssel.)

Egy másik egyenesillesztési eljárás a *Hough-transzformáció*n alapul [4-12]. A transzformáció elvét a 4-29/a s 4-29/b ábrán mutatjuk be. Eszerint az (x, y) sík bármely egyenesének megfelel egy pont a (ρ, ϑ) transzformációs síkon, az

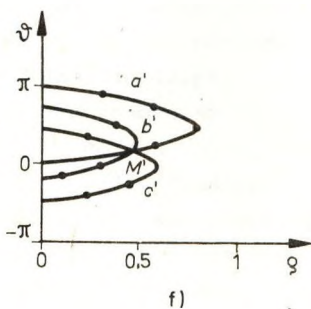
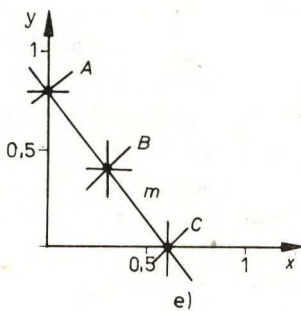
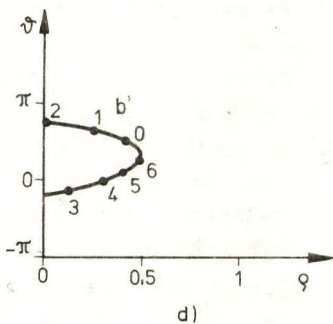
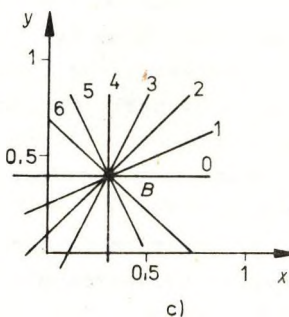
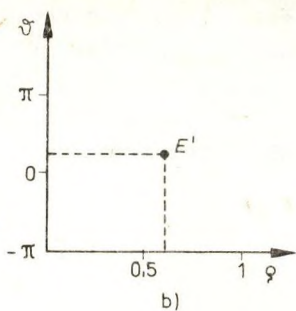
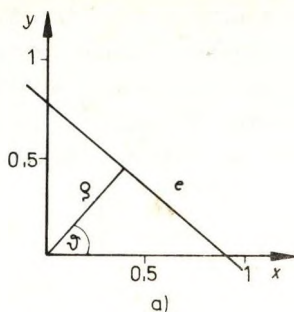
$$x \sin \vartheta + y \cos \vartheta = \rho$$

összefüggés szerint $(0 \leq \vartheta < 2\pi)$. Tekintsük most a B ponton átmenő sugársort (4-29/c ábra), ezek képe egy görbe (4-30/d ábra). Határozzuk meg további két (A és C) pont épét: ha ezek egy egyenesbe esnek, a transzformációs örbéik egy pontban fogják metszeni egymást (4-29/e és 4-29/f ábra). Ez utóbbiak M metszéspontjának megfelelő m egyenes mindhárom pontra illeszkedik.

Ha ezzel a módszerrel akarjuk meghatározni az adott pontra legjobban illeszkedő, egyenes szakaszokból álló törtvonalat, $r^2/2$ egyenletet kellene felírunk a pontokon áthaladó összes lehetséges egyenes transzformálásához. Ráadásul a zajhatások erősen leronthatják az eredményt. Mindkét nehézség megszűnik, ha a (ρ, ϑ) síkot kvantáljuk, vagyis elemi téglalaprácselhálóval fedjük le. Ezek után azonosnak tekintjük azokat az egyeneseket, amelyeknek képe egy cellába esik. A transzformálás során növeljük a megfelelő cellaszámláló értékét, és csak azokat az egyeneseket állítjuk vissza, amelyeknek megfelelő cella számlálójának értéke egy küszöbnél nagyobb.

Végül megemlítünk még egy viszonylag egyszerű törtvonalas közelítési módszert. Az alapelvét a 4-30. ábrán szemléltetjük (lásd [A-4]; p. 338. old.):

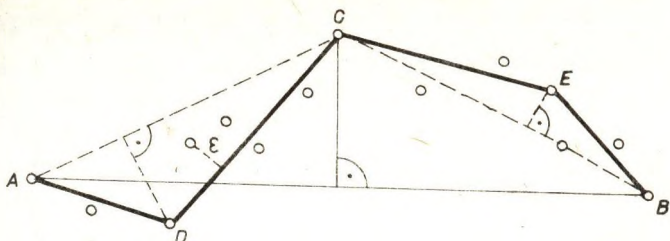
1. Válasszuk ki a ponthalmaz két legszélső elemét (az ábrán az A és a B pont), és kössük össze egy



4-29. ábra. A Hough-transzformáció szemléltetése:

a)-b) egyenes leképzése, c)-d) pont leképzése

e)-f) egy egyenesbe eső pontok megkeresése



4-30. ábra. Tört vonalas közelítés végpontillesztéssel

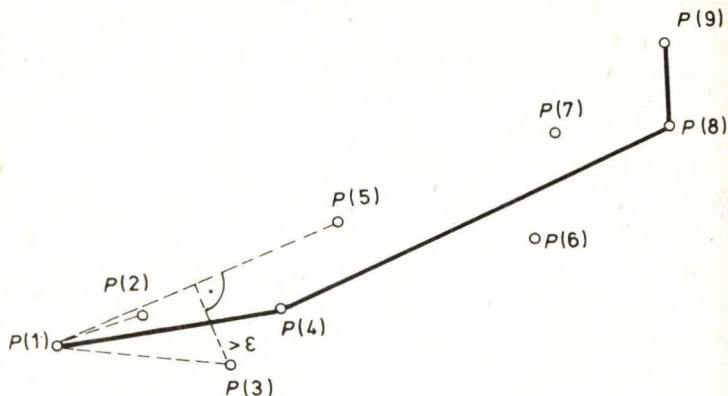
egyenessel. Ha ettől az egyenestől az összes pont a hibahatárnál kisebb távolságra van, az eljárás véget ért; különben folytatás 2-től.

2. Válasszuk ki az egyenestől legtávolabbi pontot (C). Töröljük az AB egyenest, és rajzoljuk meg az ACB tört vonalat. Ha az összes többi pont ezektől mért legkisebb távolsága a hibahatáron belül van, az eljárás véget ért; különben folytatás 3-tól.
3. Folytassuk az eljárást külön-külön mindegyik vonalszakaszra mindaddig, míg az összes pontot meg nem közelítettük a hibahatáron belül. A 4-30. ábrán a végeredményt az $ADCEB$ törtvonal adja.

A módszer súlyos hibája, hogy a kiválasztott töréspontok közelében levő pontokat esetleg rosszul közelíti, és hogy egyetlen "kilógó" hibapont teljesen félrevezető eredményt adhat. Bár az utóbbiak hatását csökkentheti egy előzetes sűrítő szűrés, célszerűen csak egy összetettebb eljárás kezdeti értékeinek meghatározására érdemes használni.

Gyors és kielégítő eredményt kaphatunk a *minimális hosszúságú poligon* (MHP) kereső módszerrel [4-53]. Legyen a pontok száma $r \geq 3$; jelölje az aktuális pont sorszámát j , az utolsó elfogadott szakasz határpontját s és a meg-

talált töréspontokat $T(i)$. Az eljárás a következő lépésekből áll (lásd 4-31. ábra):



4-31. ábra. Minimális hosszúságú poligon rajzolása

1. Kezdetben legyen $i = 1$, $T(1) = P(1)$ és $j = s = 2$.
2. Növeljük meg j értékét: $j = j + 1$. Ha $j > r$, folytatás 5-től; különben 3-tól.
3. Határozzuk meg az összes $P(n)$ ($i < n < j$) pont távolságát a $\overline{T(i)P(j)}$ szakasztól. Ha mindegyik kisebb a hibakorlátnál, legyen $s = j$ és ismétlés 2-től; különben folytatás 4-től.
4. Növeljük meg i értékét: $i = i + 1$, és legyen $T(i) = P(s)$; folytatás 2-től. (Az ábra szerint ez első ízben $j = 5$ esetén következik be, ekkor a $P(3)$ távolsága a $\overline{T(1)T(5)}$ szakasztól $> \epsilon$. Így a $T(2) = P(4)$ lesz.)
5. Ha $T(i) \neq P(j-1)$, növeljük meg i értékét: $i = i + 1$, és legyen $T(i) = P(j-1)$; ezzel, vagy különben az eljárás véget ért.

Az MHP poligont a 4-31. ábrán a $P(1)-P(4)-P(8)-P(9)$ töréspont jelöli ki.

Az eljárás a *Hausdorff-euklidészi* távolságdefiníció értelmében minimális hosszúságu tört vonallal közelíti a pont-halmaz pontjait. Ezt a távolságot a fenti jelölésekkel a

$$\rho_{HE} = \max_i [\max_j (\min(\rho_E(P(j)-T(i))), \max_j (\min(\rho_E(P(j)-T(i)))))] \quad (4-63)$$

képlet definiálja, ahol ρ_E az euklidészi távolságot jelent [4-53].

Az eljárás hátránya, hogy a távolságokat minden aktuális szakasztól újra kell számolni, s ez nagyon időigényes. Jelentősen javított változata a *rektifikált minimális hosszúságu poligon* (RMHP) kereső eljárás [4-54].

4.4.2.2 Vonalvékonyítás

Az élkorrekciónak viszonylag egyszerű és elterjedten alkalmazott másik módja a következő:

- A döntési küszöböt alacsonyra választják. Ezáltal sok hamis élpontot jelölnek ki, de várható lesz, hogy az összefüggő élek nem szakadnak el.
- Az élképre *vonallvékonyító* eljárást alkalmaznak, miáltal a széles "élszalagok" 1 képpont vastagságu élvonalakká zsugorodnak össze, a folytonosság megtartása mellett.

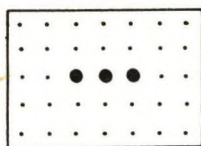
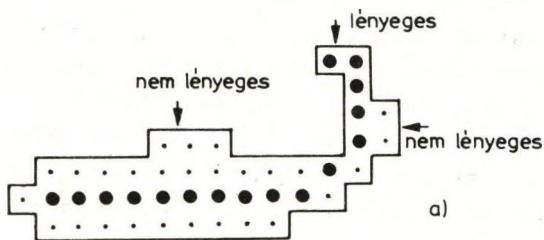
A módszer hiányossága, hogy a hamis "élnyulványokat" nem tudja megkülönböztetni a valódi élektől, így ezek megmaradnak a végső élképben.

A vonalvékonyítás során mindig *bináris* (értékes pont: 1, háttérpont: 0) képből indulunk ki. A célkitűzés az, hogy megkeressük az értékes képpontokból álló, kiterjedt alakzatok *vázpontjait*, amelyek teljesítik a következő feltételeket:

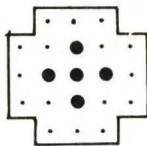
- a lehető legtávolabb vannak a határpontoktól,
- megtartják az alakzat topológiai tulajdonságait (vagyis a (4)-összefüggő tartományok (8)-összefüggő vázvonala zsugorodnak);
- számuk a lehető legkisebb (vagyis a vázvonal a lehető legsímább, azaz töréspontjainak száma minimális).

A módszer fő problémái:

- a digitális vonalnak - szemben a geometriaival - véges (1 képpont) vastagsága van, ezért a vonalat alkotó képpontokat nem mindig lehet egyértelműen kijelölni (vö. F3.2.2 alpont);
- a képhibák befolyásolják a konvexitást; a "nem lényeges" kitüremkedéseket nem kell figyelembe venni (4-32/a ábra);
- ha az alakzat nem eléggé "elnyúlt" (lásd alább), a vázvonal értelmezése nehézséget okoz (lásd 4-32/b és 4-32/c ábra).



b)



c)

4.32. ábra. Vázpontok kitűzése

Az irodalomban "klasszikusnak" nevezett vonalvékonyító eljárás a *vázpontok meghatározásán* alapul [A-12]. Elve a következő (az alábbiakban egy képpont nyolc szomszédját az iránykódok szerint α_i -vel jelöljük; lásd F3.1.2 alpont):

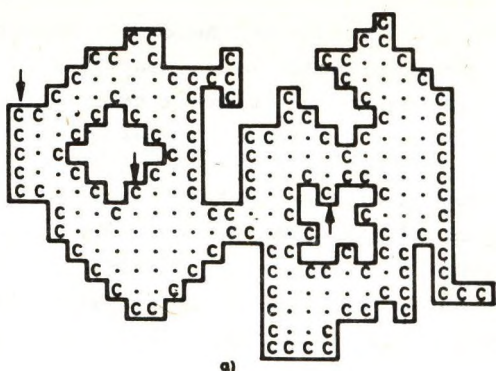
1. Keressük meg az alakzat konturpontjait:

- a) Induljunk ki egy olyan pontból, amelynek α_4 szomszédja nem tartozik az alakzathoz (lásd a 4-33/a ábrán a nyíllal megjelölt pontokat). Az aktuális konturpontot jelöljük C -vel. Legyen $s = 6$.
- b) Ha C -nek α_{s-1} szomszédja nem eleme az alakzatnak, folytatás c)-től; különben legyen ez az új aktuális konturpont, és legyen $s = s - 2 \pmod{8}$; folytatás b)-től.
- c) Ha C -nek α_s szomszédja nem eleme az alakzatnak, folytatás d)-től; különben legyen ez az új aktuális konturpont; folytatás b)-től.
- d) Ha C -nek α_{s+1} szomszédja nem eleme az alakzatnak, folytatás e)-től; különben legyen ez az új aktuális konturpont; folytatás b)-től.
- e) Legyen $s = s + 2 \pmod{8}$ és folytatás b)-től mindaddig, amíg vissza nem értünk a kiindulási pontba.

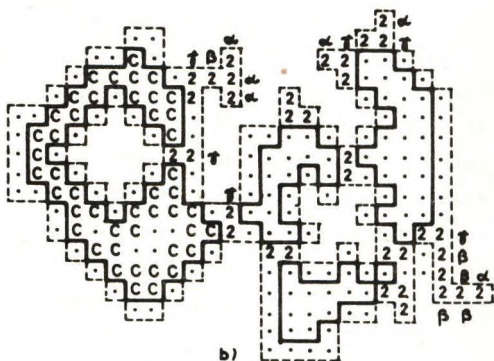
Többszörösen összefüggő alakzatok esetében az eljárást minden "lyukra" is le kell futtatni. Az első menet után konturpontnak minősülő pontokat a 4-33/a ábrán C -vel jelöltük.

2. Jelöljük ki a konturpontok közül a vázpontokat az alábbi feltételek szerint.

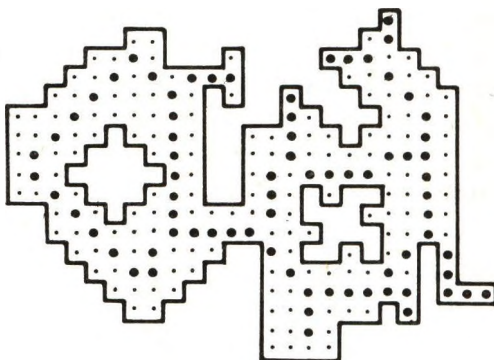
- α) minden értékes (8)-szomszédja konturpont; vagy
- β) legalább kétszer áthaladtunk rajta; vagy
- γ) legalább egy olyan konturponttal (4)-szomszédos, amelyik a haladás irányát tekintve a konturvonalon nem szomszédja.



a)



b)



c)

4-33. ábra. A vonalvékonyítás szemléltetése

A 4-33/b ábrán 2-vel jelöltük a megtalált vázpon-
tokat. Ezek között példaképpen megjelöltünk néhányat aszerint, hogy melyik fentebbi feltétel tel-
jesült rá.

3. Hagyjuk el a megtalált C -jelű konturpontokat, így
új alakzat keletkezik. Ha ennek van még jelöletlen
pontja, ismételjük meg a konturkijelölő eljárást.
Ha nincs, a 2 jelű pontokból - megfelelő törlés
után - megmaradó pontok az alakzat vázát adják.
(Lásd a 4-33/c ábrán a \bullet -tal megjelölt pontokat.)

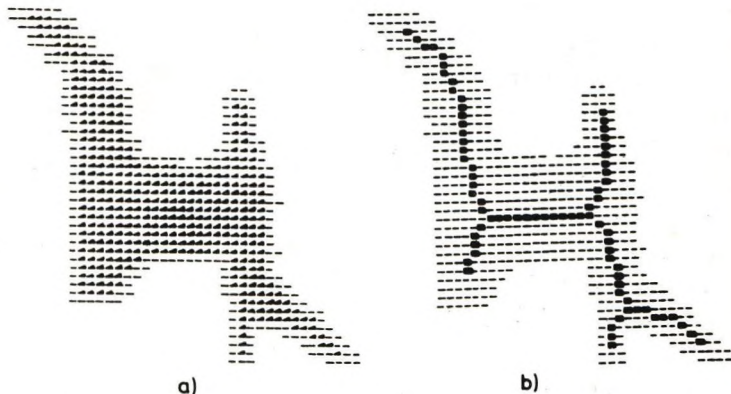
A törlés során ügyelni kell az *összefüggőség* megtartá-
sára. Az erre vonatkozó vizsgálat az eredeti algoritmus-
ban nem volt korrekt, ezért módosítanunk kellett [4-45].
A módosított eljárást lefuttattuk a 26/c képen látható
kétszintes élképre. Az 1. iteráció eredményét a 27/a
végeredményt a 27/b képen mutatjuk.

Egy másik eljárás a "*Lényeges kitüremkedések*" megke-
resésén és megőrzésén alapul [4-1]. Vezessük be a követ-
kező jelöléseket:

- A az aktuális alakzathoz tartozó-, \bar{A} a háttérpontok
halmaza;
- $C \subset A$ azon pontok halmaza, amelyeknek legalább egy
(8)-szomszédjuk $\in \bar{A}$ (=konturvonal; figyeljük meg,
hogy C (4)-összefüggő!);
- nevezzük "*mag*"-nak a $K = A \setminus C$ halmazt.

A kitüremkedések felismeréséhez pontosítani kellett az
elnyultság fogalmát. Eszerint *elnyultnak* (= kitüremkedés-
nek) tekintik a C halmaz azon részhalmazait, amelyek nem
szomszédosak K -val. *Lényegesnek* minősülnek azok a *kitü-
remkedések*, amelyek (8)-távolsága K -tól nagyobb egy adott
küszöbnél. (Ez a definíció zajos képeknél előnyös, és a
küszöb módosításával igazítható a kép jellegéhez.) Az el-
járás fő lépései a következők:

1. Jelöljük ki A konturpontjait, majd a K halmazt.
Ha K nem üres, folytatás 2-től; különben 4-től.
2. Keressük meg a lényeges kitüremkedéseket.
3. Töröljük A -ból azokat a pontokat, amelyek nem elemei K -nak, és nem tartoznak egyetlen lényeges kitüremkedéshez sem; továbbá nem szükségesek az összefüggőség fenntartásához. Az így kapott A' alakzatra ismétlés 1-től.
4. Az eddig megtartott pontok halmazára indítsunk vonalvékonyító eljárást, amely 1 képpont vastagságú, (8)-összefüggő vonalakat eredményez.



4.34. ábra. Vonásvékonyítás a "lényeges kitüremkedések" megtartásával: a) eredeti kép, b) végeredmény

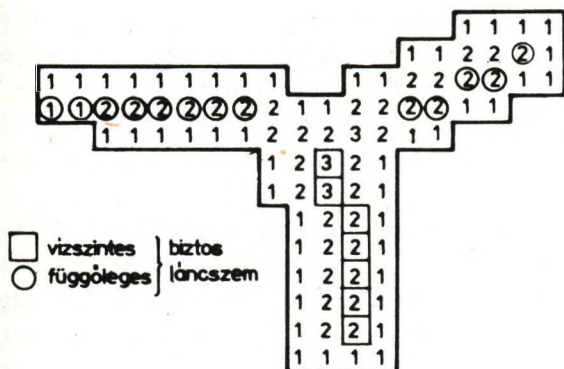
A 4-34. ábrán egy mintaalakzaton mutatjuk be az eljárás eredményét [4-1] alapján.

Végül ismertetünk még egy eljárást, amely *egyenes vonalakkal határolt* alakzatok felismerésére alkalmas [4-45]. Az algoritmus a következő lépésekből áll:

1. Meghatározzuk az alakzat minden pontjának legkisebb (4)-távolságát a háttértől.
2. Soronként, majd oszloponként végigvizsgáljuk az alakzatot, és "biztos láncszemek" (BL) jelöljük ki azokat a pontokat, (egy sorban, illetve oszlopban egyet), amelyeknek t távolságkódjára teljesül a

$$t = \min\left(\text{ent}\left(\frac{h_s + 1}{2}\right), \text{ent}\left(\frac{h_o + 1}{2}\right)\right),$$

ahol h_s , illetve h_o a pontot tartalmazó sorban illetve oszlopban az alakzat belsejébe eső képpontok száma (lásd 4-35. ábra).



4-35. ábra. Egyenesvonalu vonalvékonyítás szemléltetése

3. Megkeressük a BL-sorozatok kezdő- és végpontjait, figyelembe véve a lehetséges töréspontokat.
4. Összefűzzük azokat a BL-sorozatokat; amelyeknek iránytényezője és végpontjaik távolsága elég közel esik egymáshoz.

5. Szükség esetén megnyújtjuk a BL-szakaszokat az alakzat határáig; illetve töröljük a túl rövid szakaszokat.
6. Végül meghatározzuk az így kapott egyésszakaszok hordozó egyeneseinek metszéspontjait, s ha ezek elég közel esnek a szakaszvégpontokhoz, megnyújtjuk a szakaszokat.

Az eredményeket a 28. képen szemléltetjük. Ezen egy alakzat kameraképe (a), a két szintre vágott kép (b) és a vonalvékonyítás végeredménye (c) látható.

4.4.2.3 Relaxációs módszerek

Relaxációs módszereknek nevezzük a fokozatos közeli-tés elvén alapuló, a környezethez alkalmazkodó iteratív módszereket, amelyeket széles körben alkalmaznak a közelítő számításokban. Ez - a képfeldolgozás területén - kissé részletesebben azt jelenti, hogy egy iterációs lépésben egy képpont sajátságvektorát aszerint változtatjuk meg, hogy mennyire van összhangban (kompatibilis), illetve ellentmondásban környezetének sajátságvektoraival. Az iterációk során az "erős" elemek mintegy magukba szívják a kompatibilis környezet sajátságait, illetve fordítva: a "gyengék" sajátságai szétoszlanak a velük kompatibilis környezetben. Az összhangot, illetve hiányát statisztikus módszerekkel mérjük.

Az elvet precízen is megfogalmazzuk [4-50]:

Jelölje $B = \{b_j\}$ a vizsgált objektumok (pl. képpontok, élpontok vagy bármilyen más alakzatok), $C = \{c_i\}$ pedig a lehetséges osztályok halmazát ($j = 1, 2, \dots, p$; $i = 1, 2, \dots, h$); és legyen $D = \{b_d\}$ az egyes objektumok figyelembe vett környezete ($d = 1, 2, \dots, r$).

Jelöljük $P_{ji}(\tau)$ -val annak becsült valószínűségét, hogy a τ -adik iterációs lépés végén b_j a c_i osztályba tartozik; nyilván $\sum_{i=1}^h P_{ji}(\tau) = 1$. Végül vezessük be a kompatibilitás jellemzésére szolgáló, korrelációs típusu relaxációs paramétert:

$$|R(b_j, c_i, b_d, c_m)| \leq 1.$$

Ez azt fejezi ki, hogy mennyire van összhangban a $b_j \in c_i$ ($= b_j$ -t a c_i osztályba soroltuk) a $b_d \in c_m$ ($= b_d$ a c_m osztályba tartozik) relációval. (Az $R = 1$ eset a környezet "teljes egyetértését", az $R = -1$ "maximális ellenkezését" jelenti. Az $R \approx 0$ esetekben a besorolás közömbös.)

A környezet "véleményét" az összes lehetséges besorolással kapcsolatban a

$$w_{ji}(\tau) = \sum_{m=1}^h s_m \sum_{d=1}^r P_{dm}(\tau) R(b_j, c_i, b_d, c_m) \quad (4-64)$$

kifejezés adja. (Ekkor hallgatólagosan feltételeztük, hogy a környezet egyes elemeinek "véleménye" egymástól független, ami teljesen sohasem igaz; ez az egyszerűsített modell azonban a gyakorlatban sok esetben kielégítő.) Az s_m súlytényezőkre azért van szükség, hogy az egyes szomszédok "egyéni szavazatából" (amit a második összeg reprezentál) képzett eredmény is a $[-1, 1]$ intervallumba essék. Az új besorolási valószínűséget a

$$P_{ji}(\tau + 1) = \frac{P_{ji}(\tau) (1 + w_{ji}(\tau))}{\sum_{i=1}^h (1 + w_{ji}(\tau))} \quad (4-65)$$

képlettel definiálhatjuk.

A számítást - elvben - egyidejűleg kell elvégezni a B halmaz minden elemére. Megfelelő számú iterációs lépés után az eljárás általában stabilitást mutat: a P_{ji} valószínűségek egy adott osztályra nézve nagyok, egyszersmind a többire nézve kicsik lesznek. Ilymódon az osztályozást a kívánt pontossággal tudjuk elvégezni.

Először nézzünk egy olyan eljárást, amely a megtalált élek "megerősítését" a "nem-élek" elnyomásával egyidejűleg végzi [4-51]. Ehhez határozzuk meg minden képpontban a gradiens értékét és irányszögét. Jelöljük ezeket a (k, l) képpontkoordinátájú pontban $\nabla q(k, l)$ -lel, illetve $\theta(\tau, k, l)$ -lel, és számítsuk még ki az él valószínűségét is a

$$P(\tau, k, l) = \frac{\nabla q(k, l)}{\max_{u, v}(\nabla q(u, v))}$$

kifejezésből, ahol τ az iteráció sorszáma; kezdetben $\tau = 0$. (Nyilvánvalóan $0 \leq P(\tau, k, l) \leq 1$; az eljárásban a maximumot az egész képre vonatkozóan értelmezték, de ez nem szükség-szerű.) Ezután minden képpontban 4 kölcsönhatási tényezőt számolunk ki, amelyek az él (e)-, illetve nem-él (n) pontok egymásra gyakorolt hatását adják meg:

$$R_{ee}(\tau, k, l, u, v) = 2^{-\rho} \cos(\alpha - \gamma) \cos(\beta - \gamma);$$

$$R_{en}(\tau, k, l, u, v) = \min(0, -2^{-\rho} \cos(2\alpha - 2\gamma));$$

(4-66)

$$R_{ne}(\tau, k, l, u, v) = 2^{-(\rho+1)} (1 - \cos(2\beta - 2\gamma));$$

$$R_{nn}(\tau, k, l, u, v) = 2^{-\rho}.$$

A (4-66) képletekben $\alpha = \vartheta(\tau, k, l) + 90^\circ$, illetve $\beta = \vartheta(\tau, u, v) + 90^\circ$, vagyis az élek irányszöge az aktuális képpontban, illetve a relaxációs környezetének egy (u, v) képkoordinátájú pontjában; γ a két pontot összekötő egyenes irányszöge; ρ pedig a két pont (8)-távolsága.

Kis számolással belátható, hogy $R_{ee} = 0$ az "oldalsó" pontokhoz tartozó ($\gamma = \alpha \pm 90^\circ$) élek esetében. Az egy egyenesbe eső (kollineáris) élek ($\gamma = \pm\alpha$) viszont a távolságuktól függő mértékben erősítik vagy gyengítik egymást aszerint, hogy azonos ($\beta = \alpha$), illetve ellenkező értelműek ($\beta = \alpha + 180^\circ$). Láthatjuk továbbá, hogy az éleket gyengítik a kollineáris nem-élek; a nem-éleket erősítik az oldalsó élek; és végül a közeli nem-élek is erősítik egymást.

Ezek után a relaxációs "élmegerősítési" algoritmus a következő:

1. Számítsuk ki minden képpontban az új élvalószínűséget az

$$E \equiv E(\tau, k, l) = \sum_{u=k-s}^{k+s} \sum_{v=l-t}^{l+t} [\lambda_1 P(\tau, u, v) R_{ee}(\tau, k, l, u, v) + \lambda_2 (1 - P(\tau, u, v)) R_{en}(\tau, k, l, u, v)]; \quad (4-67/a)$$

$$N \equiv N(\tau, k, l) = \sum_{u=k-s}^{k+s} \sum_{v=l-t}^{l+t} [\lambda_3 P(\tau, u, v) R_{ne}(\tau, k, l, u, v) + \lambda_4 (1 - P(\tau, u, v)) R_{nn}(\tau, k, l, u, v)] \quad (4-67/b)$$

él-, illetve nemél-valószínűségnövekmények alapján a

$$P(\tau + 1, k, l) = \left[1 + \frac{1 - P(\tau, k, l)}{P(\tau, k, l)} \frac{N + |E| + |N|}{E + |E| + |N|} \right]^{-1} \quad (4-68)$$

képletből.

2. Határozzuk meg az új éllirányszöveget az

$$A \equiv A(\tau, k, l) = \lambda_5 P(\tau, k, l) \cos(\vartheta(\tau, k, l)) + \\ + \sum_{u=k-s}^{k+s} \sum_{v=l-t}^{l+t} P(\tau, u, v) R_{ee}(\tau, k, l, u, v) \cos(\vartheta(\tau, k, l));$$

(4-69/a)

illetve a

$$B \equiv B(\tau, k, l) = \lambda_5 P(\tau, k, l) \sin(\vartheta(\tau, k, l)) + \\ + \sum_{u=k-s}^{k+s} \sum_{v=l-t}^{l+t} P(\tau, u, v) R_{ee}(\tau, k, l, u, v) \sin(\vartheta(\tau, k, l))$$

(4-69/b)

összefüggések alapján a

$$\vartheta(\tau + 1, k, l) = \arctg\left(\frac{B}{A}\right) \quad (4-70)$$

képletből. A (4-67) és a (4-69) képletekben a relaxációs környezet a (k, l) pont körüli $(2s + 1) \times (2t + 1)$ méretű téglalap, de az összegzésből kizárjuk az $u = k$ és az $l = v$ esetet. Érvényes továbbá a

$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$$

összefüggés. Az új irányszöveget annál nagyobb mértékben befolyásolják a környező pontok, minél kisebb λ_5 értéke.

3. A (4-70)-ből kapott új szöggel újraszámítjuk az éll valószínűségeket a (4-66)-(4-68) képletek

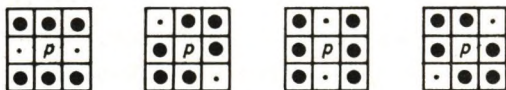
alapján, majd meghatározzuk az új élszöget stb. Az eljárást $\tau = \tau + 1$ -gyel addig ismételjük, míg a $P(\tau + 1, k, l) - P(\tau, k, l)$ különbség elég kicsi nem lesz.

4. Élpontoknak tekintjük azokat a pontokat, amelyeknek élvalószínűsége nagyobb egy megadott küszöbnél.

Érezhető, hogy az eljárás - mint általában a relaxációs eljárások - nagyon számításigényes.

Ebből a szempontból kedvezőbb az alábbi hárommenetes (Eberlein-) algoritmus, amely simítási és vonalvékonyítási lépésekből áll [4-13]:

- Határozzuk meg minden pontban a gradiens nagyságát és az élrányszöget (= a gradiens iránya $+90^\circ$).
- Végezzünk ismételt simítást a gradienstérben szükség szerint (esetleg egyszer sem) a hamis élpontok eltüntetése céljából. A simításban a pontok 3×3 -as környezetének 6 pontja vesz részt. Az él irányszöge szerint lehetséges 4 esetet a 4-36. ábrán mutatjuk; itt a simításban részt vevő szomszédokat • jelöli. A simítás két elven történhet:
 - az aktuális pont gradiensét megnöveljük a megfelelő 6 szomszéd gradiensének súlyozott átlagával (infúzió; pl. úgy, hogy az új érték felét adja a szomszédok átlaga);
 - a 6 szomszédjának gradiensét megnöveljük az aktuális pont gradiensének egy hányadával (diffúzió; pl. $1/6$ -od részével).



4-36. ábra. Infúziós és diffúziós környezet

Mivel az eljárás *nem konvergens*, a két módszer más eredményhez vezet. A megvizsgált esetekben az infúziós módszer bizonyult jobbnak, szemben a simítás nélküli esettel, illetve a diffúziós simítással.

3. Végezzünk *ismételt vonalvékonyítást* avégett, hogy az esetleges széles éltartományok egy pont vastagságu vonalakra csökkenjenek, illetve a "gyenge" élek kiemelkedjenek a környezetből. A cél elérésére az *abszorpciós* módszert használták. Ennek lényege, hogy egy élpont "elnyel" a gradiensének irányába eső két szomszédja közül egyiknek vagy mindkettőnek gradienséből egy hányadot (célszerűen a felét), ha ennek iránya az övével párhuzamos.

Az eljárás az egzakt módszernél sokkal gyorsabb, mivel a gradiensek egyszer megtalált irányát nem változtatja meg. Ez egyben gyengéje is, amit fokoz, hogy a vékonyítás lényegében lineáris algoritmuson alapul.

4.4.2.4 Heurisztikus módszerek

Tekintve, hogy egyrészt ez idő szerint nincs olyan elmélet, amely támpontot adna, hogy milyen képfeldolgozási feladatokra milyen módszereket kell vagy lehet használni, másrészt az egzaktságra törekvő módszerek számításigénye sokszor nem elviselhető, indokolt a *heurisztikus módszerek* alkalmazása is. Ezekben az élképet - tapasztalati alapon - úgy igyekeznek korrigálni, hogy az eredmény "kellemes látványt" nyújtson, mérsékelt futásidők mellett. (Az eddigieken - szakadások megszüntetésén, "szőrök" letörlésén - tulmenően ez még a felesleges töréspontok eltávolítását is jelenti.)

A *vonalvezési eljárás* [4-38] az irányított éldarabok megkeresésén és fokozatos összekapcsolásán alapul.

Az eljárásban az élpontok kijelöléséhez 6 speciális szűrőt használnak (lásd 4-37. ábra) s ezek alapján minden élponthoz egy élírányt is meghatároznak. Az élkorrekcio 4 menetben történik:

$$\begin{bmatrix} -100 & -100 & 0 & 100 & 100 \\ -100 & -100 & 0 & 100 & 100 \\ -100 & -100 & 0 & 100 & 100 \\ -100 & -100 & 0 & 100 & 100 \\ -100 & -100 & 0 & 100 & 100 \end{bmatrix} \quad \begin{bmatrix} -100 & 32 & 100 & 100 & 100 \\ -100 & -78 & 92 & 100 & 100 \\ -100 & -100 & 0 & 100 & 100 \\ -100 & -100 & -92 & 78 & 100 \\ -100 & -100 & -100 & -32 & 100 \end{bmatrix}$$

$$\begin{bmatrix} 100 & 100 & 100 & 100 & 100 \\ -32 & 78 & 100 & 100 & 100 \\ -100 & -92 & 0 & 92 & 100 \\ -100 & -100 & -100 & -78 & 32 \\ -100 & -100 & -100 & -100 & -100 \end{bmatrix} \quad \begin{bmatrix} 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 100 & 100 \\ 0 & 0 & 0 & 0 & 0 \\ -100 & -100 & -100 & -100 & -100 \\ -100 & -100 & -100 & -100 & -100 \end{bmatrix}$$

$$\begin{bmatrix} 100 & 100 & 100 & 100 & 100 \\ 100 & 100 & 100 & 78 & -32 \\ 100 & 92 & 0 & -92 & -100 \\ 32 & -78 & -100 & -100 & -100 \\ -100 & -100 & -100 & -100 & -100 \end{bmatrix} \quad \begin{bmatrix} 100 & 100 & 100 & 32 & -100 \\ 100 & 100 & 92 & -78 & -100 \\ 100 & 100 & 0 & -100 & -100 \\ 100 & 78 & -92 & -100 & -100 \\ 100 & -32 & -100 & -100 & -100 \end{bmatrix}$$

4-37. ábra. Éldetektáló szűrők 6 fűirányra

1. Először *vonalvékonyítást* végeznek. Ez után egy élpontnak nem lehet élpont szomszédja az élírányra merőlegesen, élírányban viszont legfeljebb 3 lehet (lásd 4-38/a ábra; itt és a további ábrákon az élpontokat •, az elemzés során figyelembe vett pozíciókat x jelöli.)

2. Ezután *összefűzik* azokat a szomszédos élpontokat, amelyeknek élíránya $\pm 30^\circ$ -on belül megegyezik. (A 30° , illetve 60° élírányu pontokból a megfelelő (8)-szomszédra lépnek.) Három változat lehetséges:
- Ha az aktuális élpontnak csak 1 élszomszédja van, ehhez hozzáfűzik;
 - Ha 2 élszomszéd van, a 4-38/b ábra szerint 3 féle elrendezésben állhatnak (nem szomszédos, szomszédos különböző irányu és szomszédos egyirányu utódok);
 - Ha 3 élszomszéd van, a 4-38/c ábra szerinti 2 eset lehetséges.

A két utóbbi változatban az összefűzést a szaggatott vonalak jelzik.

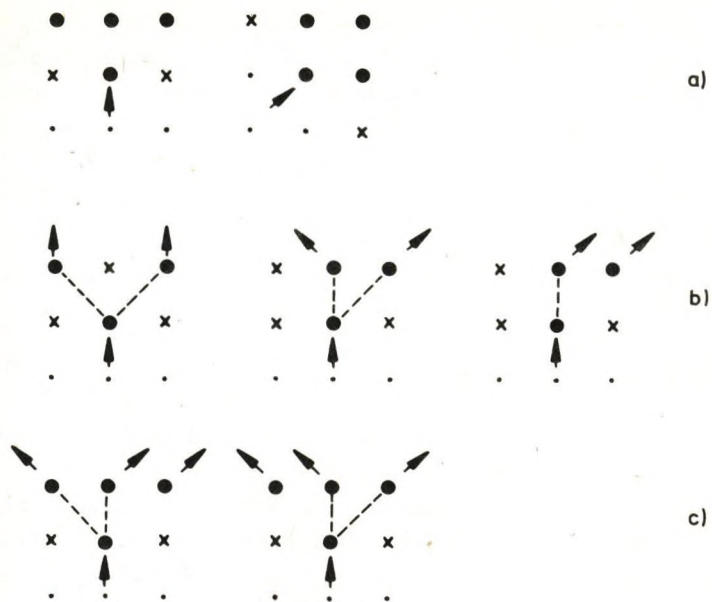
3. *Vonalkövetéssel* meghatározzák az élvonalakat.

Először olyan élpontokból indulnak ki, amelyeknek nincs elődjük, és egy vonalba gyűjtik a velük összefűzött pontokat (figyelembe véve az elágazásokat is). Ezután zárt élvonalakat keresnek, amikor is szükség szerint kitöltik az 1 képpontnyi hézagokat.

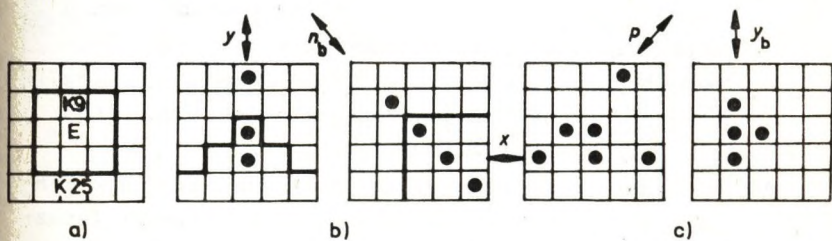
4. Végül a kapott élpontsorozatokat *szakaszonként lineáris* polinomokkal helyettesítik a 4.4.2.1 alponthban vázolt tört vonalas közelítéssel.

Az eljárás tapasztalati értékelése során gyorsabbnak és jobbnak bizonyult a módosított Hückel-féle [4-25] eljárásnál.

Végül ismertetünk még egy saját algoritmust, amely az élpontok 5×5 -ös környezetében (K25; lásd 4-39/a ábra) található *irányok kiegészítésére* törekszik ([4-3], [4-4]). Az eljárás alapvető fogalma az *irány*, amit legalább 3, egy egyenesbe eső (kollineáris) élpont határoz meg. *Főirányról* beszélünk, ha az iránynak az aktuális élpont (E) is eleme.



4-38. ábra. Élösszefüzési stratégia szemléltetése

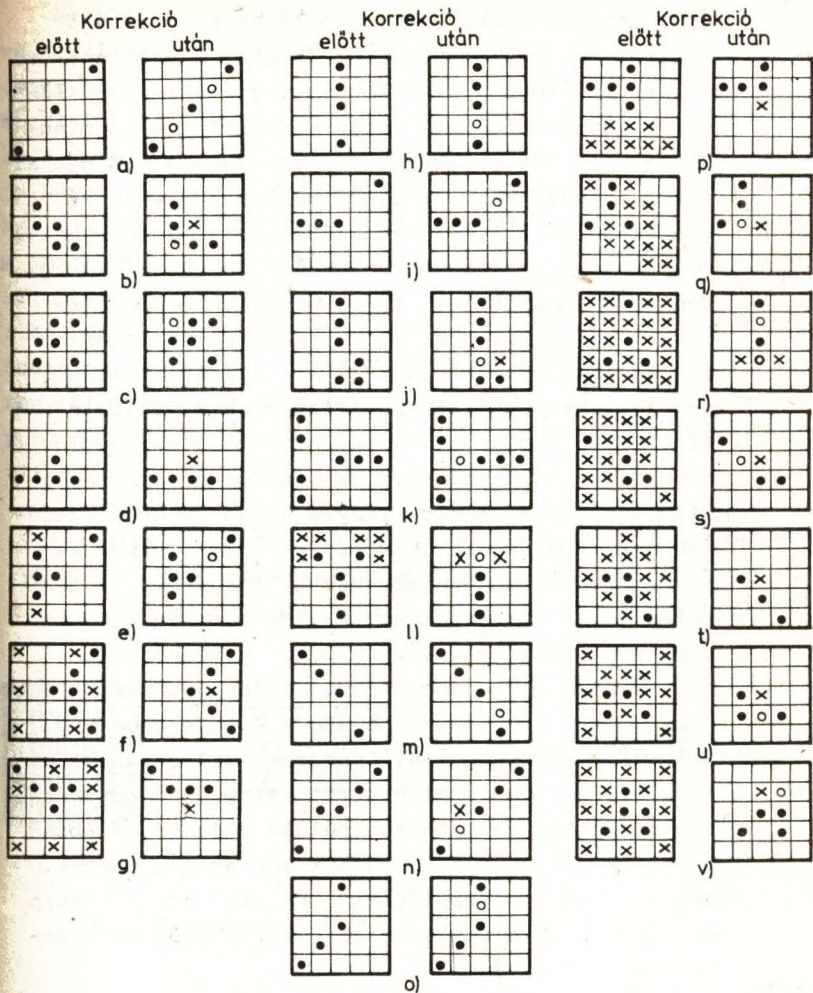


4-39. ábra. Elkorrekciós környezet és a lehetséges irányok

(A lehetséges négy főirányt x , y , $p = +45^\circ$ és $n = -45^\circ$ szimbólummal jelöljük, lásd pl. 4-39/b ábra; a többi irányt mellékiránynak nevezzük, lásd pl. 4-39/c ábra.) Mindkét esetben megkülönböztetjük a belső irányokat: ezeknek 3 pontja E 3×3 -as környezetébe (K9) esik.

Az algoritmusnak 5 fő ága van, ezek - az elemzés sorrendjében - a következők:

1. *Külső főirány kiegészítésre* akkor kerül sor, ha E belső környezete (K9) "üres", vagyis nem tartalmaz más élpontot. Ha ilyenkor létezik főirány, kiegészítjük. (Lásd 4-40/a ábra; itt a P -főirányt egészítettük ki. A "betoldott" élpontokat itt és a továbbiakban o ; a korrekció során megvizsgált *lényeges* pozíciókat - ha azok nem élgyanus pontok -, illetve a korrekció után "kitörölt" élpontokat is \times jelzi.) Ha egyetlen főirány sem létezik, E -t töröljük az élpontok halmazából. A további esetekben K9-ben legalább még egy élgyanus pont van.
2. *Belső sarokpont vizsgálatot* végzünk, ha létezik belső főirány, és ez ferde (P , illetve N). Attól függően, hogy az eloszlásból egy sarokba befutó ferde élre lehet-e következtetni vagy sem, E -t meg hagyjuk, illetve töröljük. A "sarok" esetleg hiányzó csucsát mindenképpen kiegészítjük (lásd pl. 4-40/b és 4-40/c ábra).
3. Ha nincs belső főirány, a *belső mellékirányok korrekciójára* kerül sor, feltéve, hogy ilyenek léteznek. Ha a mellékirány külső is, E -t töröljük (4-40/d ábra). Különben elsődlegesen a még esetleg létező további irányok kiegészítésére törekszünk (pl. 4-40/e ábra). Ha ilyenek nincsenek, a vonalvezetés "simitását" végezzük el (pl. 4-40/f és 4-40/g ábra) a felesleges élpontok törlésével.



4-40. ábra. Jellegzetes élponteloszlások

4. Ha semmilyen belső irány nem létezik, most azt vizsgáljuk meg, hogy *van-e főirány*, s ha igen, az algoritmus legbonyolultabb ága, a *szomszédságelemzés* következik. Ennek során minden főirányt mindkét értelemben bejárunk, és megvizsgáljuk E -nek az adott irányra szimmetrikus, 90° -os környezetét (=szomszédságát). (A szomszédsághoz E -vel együtt 9 pont tartozik, lásd pl. a 4-39/b ábrán a vastag vonallal bekeretezett tartományt.) Az elemzésre csak akkor kerül sor, ha E -nek az aktuális értelemben első szomszédja (P_0) *nem* élpont. (Pl. a 4-39/b ábra szerinti y főirányt csak "felfelé" értelemben fogjuk elemezni, az n főirányt pedig egyáltalán nem.) Mivel a szóban forgó tartomány szimmetrikus az adott irányra nézve, mindkét felében *azonos* vizsgálatokat végzünk. Elsősorban az irány kiegészítésére törekszünk (vagyis P_0 -t lehetőleg élponttá változtatjuk). Ha azonban a többi szomszéd eloszlásából arra lehet következtetni, hogy E töréspont, akkor az új irányt erősítjük meg. x és y irányban - a lehetséges 4 értelemben 932, a ferde irányokban 168, összesen 1100 olyan eloszlás van, amikor az algoritmus valamilyen korrekciót végez. Néhány jellegzetes példát a 4-40/h-o ábrán mutatunk. Külön megvizsgáljuk még azt az esetet, amikor a szomszédság "üres". Ha ilyenkor a 4-40/p ábra szerinti eloszlás van, E -t töröljük mint felesleges "nyulványt".
5. Ha semmilyen irány sem létezik, végül az algoritmus másik igen bonyolult ágára térünk. Ekkor teljesen *átvizsgáljuk E belső környezetét* (K_9 -et), s minden lehetséges esetet kiértékelünk. Attól függően, hogy 1, 2, 3, illetve 4 élszomszéd létezik, még mintegy további 250 eloszlás esetében végzünk korrekciót (lásd pl. 4-40/q-v ábra).

Megemlítünk még egy programozási fogást, amely lehetővé teszi az irányok gyors felismerését, s így jelentősen csökkenti a futási időt. A K25 pontjait a 4-41/a ábra szerint beszámozzuk, majd leképezzük 3 byte-ra. Ezek rendre a P (primer), S (szekunder) és T (tercier) szomszédokat jelképezik (lásd 4-41/b ábra). Az egyes bitek értéke akkor 1, ha a megfelelő sorszámú pont élpont. Ekkor pl. az *y* főirányt a 4-41/c ábra szerinti maszkkal választhatjuk ki. Ebben az irányban korrekciót akkor végzünk, ha P-re alkalmazva *legfeljebb* egy, S-re alkalmazva *legalább* egy, összesen *legalább két* koincidenciát kapunk. Figyeljük meg, hogy a maszk jobbra, illetve balra léptetésével az *E* pont megfelelő irányu S, illetve T szomszédait tudjuk vizsgálni.

Az élpontokat a 4.4.1.2 alpontban ismertetett "főirányszűrők" módszerével jelöljük ki. A 29/a képen megadjuk a 26/d képre végrehajtott iránykiegészítő korrekció eredményét; egy további példa a 29/b és c) képen látható.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

a)

12	7	8	9	14	19	18	17	P
11	1	3	5	15	25	23	21	S
6	2	4	10	20	24	22	16	T

b)

0	0	1	0	0	0	1	0
---	---	---	---	---	---	---	---

c)

4-41. ábra. Az élkorrekciós környezet leképzése

5. KÉPOSZTÁLYOZÁS

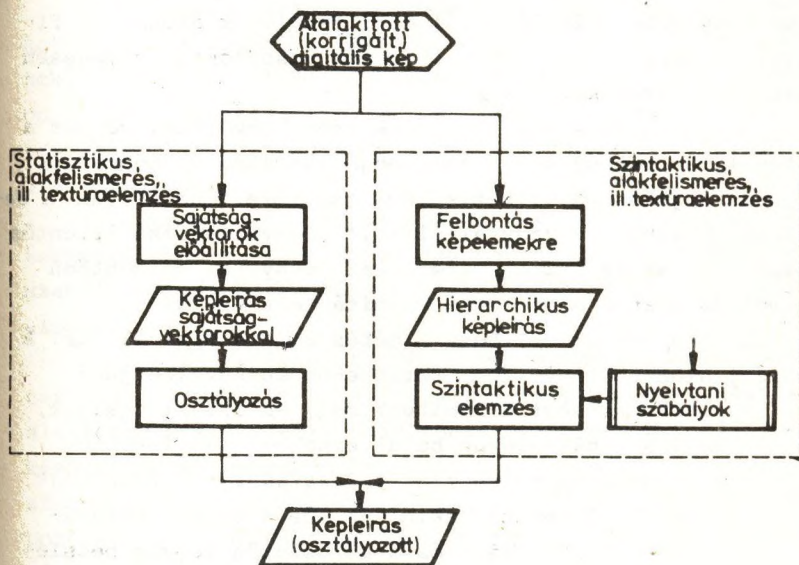
Az elemzési szinten végzett képosztályozások két nagy területe az alakfelismerés (pattern recognition) és a texturaelemzés (texture analysis; lásd 1-3. ábra). Ide tartoznak azok a módszerek, amelyek a kép lényeges jellemzőinek, a *sajátságok*nak felismerésére és kinyerésére, végeredményben pedig - a mért sajátóságok alapján - a képrészletek osztályozására és ezzel a kép valamilyen *leírására* irányulnak (kép → leírás leképezések).

Alakfelismerésről beszélünk, ha a sajátóságok a kép makroszerkezetét, az *objektumok*at jellemzik; ehhez a képet előzőleg rendszerint szegmentálni kell. (Megjegyezzük, hogy az "alakfelismerés" elnevezést nem tartjuk szerencsésnek, de elterjedtsége miatt mégis ezt használjuk. Ugyanis olyan eljárásokat és módszereket takar, amelyeket egyrészt egy sereg más területen is lehet alkalmazni, másrészt az esetek nagy részében semmilyen kapcsolatban sincsenek a hétköznapi életben megszokott, geometriai alakfogalommal.)

A *texturaelemzés* során a kép mikroszerkezetét vizsgálják. Ehhez definiálni kell a *hasonlóság* fogalmát, ami igen gyakran (de nem szükségszerűen) azon a feltételezésen alapul, hogy a *texturaelemek* elkülöníthetők. A feldolgozás az esetek túlnyomó többségében *homogén* texturákra terjed ki, amelyekben minden elempár hasonló.

5.1 ELVI ÁTTEKINTÉS

Módszereit tekintve mind az alakfelismerésnek, mind a texturaelemzésnek két, lényegesen eltérő ága van (lásd 5-1. ábra).



5-1. ábra. A képosztályozási folyamat modellje

A *statisztikus* (vagy döntéseméleti) ágon valószínűségi számítási és matematikai statisztikai módszerekkel dolgoznak. Az ily módon előállított jellemzők az objektumok, illetve a texturaelemek eloszlását írják le, de elhanyagolják a síkbeli (illetve térbeli) összefüggéseket. Előnyük a hatékonyságuk és viszonylagos egyszerűségük, ami gyakran kisebb számítási igényt is jelent.

A *szintaktikus* (vagy strukturális) módszerek ezzel szemben az objektumok, illetve texturaelemek síkbeli (térbeli)

viszonyaira koncentrálnak. Az alakfelismerésben elsősorban matematikai nyelvészeti módszereket adaptáltak. Szabályos texturák esetében ezek a módszerek hatékonyabbak a statisztikusaknál, de véletlenszerűek leírására kevésbé alkalmasak.

Az 5-1. ábrából látható, hogy mindkét terület mindkét ága lényegében két fázisból áll; felhívjuk azonban a figyelmet arra, hogy a modell szintű hasonlóság lényegesen eltérő módszereket takar.

Az *első fázis* a sajáttságvektorok kinyerése, vagyis a figyelembe veendő sajáttságok meghatározása és megmérése minden objektumra, illetve texturaelemre vonatkozóan. A sajáttságok alkalmas vagy alkalmatlan megválasztása jelentős hatással van az osztályozás végeredményére, de döntően befolyásolhatja a megoldás számításigényét is.

Emlékeztetünk rá, hogy a mérés mindig hibával jár. A mérési hibákat lényegében kétféleképpen kezelhetjük:

- *tűrőhatárokat* állapítunk meg, és azt vizsgáljuk, hogy a hibák ezeken belül esnek-e; illetve
- a mérendő mennyiségeket *valószínűségi változóknak* (lásd F2.2. pont) tekintjük, és a mérési eredményeket az eloszlás- és a sűrűségfüggvényük becslésére használjuk. (Megjegyezzük, hogy a valószínűségi mező általában nem ismeretes, de erre nincs is szükség.)

Egy színösszetevő sajáttságvektorait az első esetben a

$$\underline{z}^T = [z^{(1)}, z^{(2)}, \dots, z^{(n)}], \quad (5-1)$$

a másodikban a

$$\underline{\zeta}^T = [\zeta^{(1)}, \zeta^{(2)}, \dots, \zeta^{(n)}] \quad (5-2)$$

szimbóllummal jelöljük, ahol a $z^{(n)}$, illetve a $\zeta^{(n)}$ valós szám

az r -edik sajáttság mérőszáma. Említettük, hogy a saját-
ságvektorokat az n -dimenziós sajátstágtér helyvektorai-
ként kezeljük. Ez azt jelenti, hogy minden sajáttságvektor
egy pontot határoz meg \mathbb{R}^n -ben; egy képre meghatározott
összes sajáttságvektor pedig együttesen \mathbb{R}^n egy részhalma-
zát jelöli ki.

Az alakfelismerés során használt *geometriai jellemző*-
ket a kép jellegzetes alakzatainak (foltoknak, illetve élek-
nek) vizsgálatával nyerjük, ezekből származtatjuk le az
objektumok sajátságait. Ezek az alakzatok kerületével,
területével és lineáris méreteivel függnek össze (lásd
5.2.1.1 alpont).

A texturaelemzéshez szükséges *texturális jellemzők* a
texturaelemek 3 alapvető tulajdonságát: a finomságot, a
szabályosságot és az anizotrópiát írják le (lásd 5.4.1
alpont).

A *második fázisban* történik meg az objektumok, illetve
texturák besorolása az előre adott, vagy menet közben ki-
alakított osztályokba, majd ennek alapján a képleírás el-
készítése. A besorolás kritériuma a szóban forgó saját-
ságvektor hasonlósága valamelyik osztályéhoz, illetve külön-
bözősége más osztályokéitól. (Ezzel hallgatólagosan fel-
tételtettük, hogy az egyes osztályok sajáttságvektorai
"kellő mértékben" különböznek egymástól!)

A számítógépes eljárások segítségével nagy elemszámu
halmazok, illetve bonyolult strukturák belső szerkezetét is
fel tudjuk tárni, miáltal gyakran juthatunk addig nem
ismert összefüggésekhez.

A továbbiakban részletesen csak a statisztikus alak-
felismerési ágat tárgyaljuk (lásd 5.2 pont); a szintak-
tikus alakfelismerést éppen csak érintjük az 5.3 pontban,
a texturaelemzési területet pedig az 5.4 pontban tekint-
jük át röviden.

5.2 STATISZTIKUS ALAKFELISMERÉS

Mint már említettük, a statisztikus alakfelismerési ágon az osztályba sorolást valószínűségyszámítási és matematikai statisztikai módszerekkel végezzük. A tárgyalás során 3 alapesetet különböztetünk meg:

1. Arra az esetre, amikor az osztályok statisztikusan *teljesen meghatározottak*, a Bayes-féle döntési módszereket ismertetjük. Ehhez az osztályok feltételes sűrűségfüggvényeit (lásd 5.2.2 alpont) vagy az egyéb statisztikai jellemzőkből leszarmaztatott - és rendszerint könnyebben kezelhető - diszkriminánsfüggvényeket (lásd 5.2.3 alpont) használjuk.
2. A *részlegesen meghatározott* esetben az osztályokat csak egyes mintaelemek, a *tanítók* révén ismerjük, és ezért nem biztos, hogy minden paraméterük egyértelműen adott; sőt a tanítók halmazából álló *tananyagban* átfedések és ellentmondások is lehetnek. Az osztályozásra a távolságmérő módszereket (lásd 5.2.4 alpont) és a potenciálfüggvény-módszert (lásd 5.2.5 alpont) ismertetjük.
3. Ha az osztályokról *semmit sem tudunk*, sőt legtöbb-ször még a számukat sem ismerjük, a döntéseket *klasszterezéssel* hozhatjuk. Itt bizonyos értelemben az osztályozás megfordításáról van szó: adott elemekhez keressük a megfelelő osztályokat. Néhány ilyen módszert az 5.2.6 alpontban mutatunk be.

A képelemek (képpontok, objektumok, texturaelemek) sohasem függetlenek egymástól, ezért a jó osztályozáshoz a környezeti információkat is figyelembe kellene venni; ettől azonban az eljárások lényegesen bonyolultabbakká válnának. Néhány ilyen problémát említünk meg az 5.2.7 alpontban.

5.2.1 AZ OSZTÁLYOZÁS ELŐKÉSZÍTÉSE

Az előkészítő fázis célja az objektumokat kellő pontossággal leíró, de - a számítások csökkentése érdekében - minimális méretű sajátságvektorok előállítására. Ez rendszerint két lépésben történik: először megmérjük a kiválasztott sajátságokat, majd minimalizáljuk a kapott sajátságvektorok dimenziószámát.

5.2.1.1 A sajátságvektorok fontosabb típusai

Az alakfelismerési osztályozásokhoz leggyakrabban a képen megtalált foltok és élek geometriai paramétereiből származtatjuk le a sajátságvektorok összetevőit; ezek *lokális* tulajdonságokat mérnek. Ilyenek:

- ivhossz vagy kerület (L), terület (A),
- alaktényező (A/L),
- legnagyobb/legkisebb átmérő (D, d),
- elnyújtottság (D/d),
- konvexitás (a foltot befoglaló legkisebb területű konvex poligon területének aránya a foltéhoz) stb.

A legfontosabb *globális* jellemző a világosságkódok eloszlása; ezért az osztályozásban fontos szerep jut a különböző szempontok szerint készített *hisztogram*oknak.

A digitális képfeldolgozás egyik fontos alkalmazási területe a műholdfelvételek kiértékelése [5-19]. A felvételek két alapelv szerint készülhetnek (lásd 2.5 pont).

A *multispektrális kép* egyidejűleg több frekvenciasávban, ugyanarról a területről készített felvételekből áll. Ekkor egyszerűen a különböző frekvenciákhoz tartozó világosságkódokat szokás a sajátságvektorok összetevőinek tekinteni. Az ilyen típusu képek közül talán legismerteb-

bek a LANDSAT erőforráskutató műholdakkal készített 4, illetve ujabban a LANDSAT 4 RM esetén 7-sávós felvételek; valamint a METEOSAT meteorológiai műholdak 3-sávós képei. (Felhasználásukról pl. az [5-6], [5-7], [5-11] tanulmányban található leírás.)

Multitemporális képen ugyanazon területről különböző időpontokban készült felvételek sorozatát értjük. Bár ezek a képek geometriailag azonosak, a környezeti feltételek és az objektumok megváltozása miatt általában jelentősen eltérnek egymástól. Ilyen esetekben szokás a mért jellemzőkből olyan újabbakat származtatni, amelyek alkalmasabbak az objektumok leírására. (Pl. két vagy több kép átlaga, maximuma, minimuma, szórása stb. vagy egyéb statisztikai jellemzők.)

A felsoroltakon kívül sok más jellemzőt is lehet használni; egyes feladatoknál az ismertektől, gyakorlatiaktól esetleg teljesen eltérő sajátosságok választása bizonyulhat megfelelőnek.

5.2.1.2 Dimenziócsökkentés

Minél több - jól kiválasztott és független - sajátosságot mérünk, annál biztosabban ismerhetők fel a kép elemei, annál jobban le tudjuk írni a képet. Másrésztől a sajátosságvektor dimenziószámával, azaz a mérések számával általában erősen nő a számításigény: mind az osztályozási eljárások hatékonysága, mind pedig az igénybe veendő (számítógépes) erőforrások mennyisége jelentősen függ tőle.

Ezért ésszerűnek látszik, hogy a képi információkból csak az osztályozás szempontjából legfontosabbakat tartjuk meg. Az erre irányuló *dimenziócsökkentő módszereknek* eszerint fontos szerepük lehet az alakfelismerési folya-

matban. Mégsem alkalmazhatók korlátlanul: részben mert ezáltal csökken a döntés pontossága, részben mert a dimenziócsökkentés számitásigényét is figyelembe kell venni.

A dimenziócsökkentő módszerek két csoportját szokás megkülönböztetni:

- A *lényegkiemelési* (feature selection) csoportba azok az eljárások tartoznak, melyekben az új saját-ságvektort a régeből bizonyos komponensek - mérések - elhagyásával kapjuk.
- *Lényegtömörítésnek* (feature extraction) nevezzük mindazokat az eljárásokat, melyekben a régi vektorból az újat valamilyen transzformációval állítjuk elő. Nyilvánvaló, hogy a lényegtömörítő eljárások egyuttal lényegkiemelésnek is tekinthetők.

A dimenziócsökkentés operátora egy

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (m < n) \quad (5-3)$$

leképezés; a feladat általában ennek az operátornak a meghatározása. Legtöbbször *lineáris* operátorokra szorítunk; a leképezést ekkor mátrixszorzással valósíthatjuk meg.

A lényegtömörítő operátor által előállított új saját-ságvektor komponensei rendszerint - de nem feltétlenül - fiktív saját-ságok lesznek. Egyszerű példaként ismertetjük a *Knuth-Thomas transzformációt*, amelyet LANDSAT-felvételek esetén szoktak alkalmazni. Mint ismeretes, a régebbi LANDSAT műholdak 4-sávos felvételeket készítenek. Elméleti megfontolások - és gyakorlati tapasztalatok - alapján megállapították, hogy a felvételek valódi (belső) dimenziószáma kettő; a megfelelő lineáris transzformáció mátrixa pedig:

$$\underline{A} = \begin{bmatrix} 0,406 & 0,600 & 0,645 & 0,234 \\ -0,386 & -0,530 & 0,535 & 0,532 \end{bmatrix}.$$

Az így kapott sajátvektor első komponense lényegében a világosságnak, a második a zöldtartalomnak felel meg.

Karhunen-Loève-transzformáció. A lényegtömörítő eljárások általában valamilyen értelemben optimális operátort használnak, hogy ezáltal a legfontosabb információt emeljék ki. A kiemelés, a transzformáció "jóságát" gyakran valamilyen függvénnyel szokás mérni, a cél ennek optimalizálása.

Tételezzük fel, hogy a sajátvektorok az \mathbb{R}^n tér egy véges (r elemű) részhalmazát alkotják. Keressünk (5-3) szerint olyan *lineáris* transzformációt, amelyik optimális abban az értelemben, hogy *átlagos négyzetes hibája minimális* a mért halmazon, s egyuttal az eredményvektor átlagos hossza a lehető *legnagyobb*. Ezt a transzformációt *Karhunen-Loève transzformációnak* nevezzük.

Jelöljük a továbbiakban (5-2) szerint a sajátvektorokat. Legyen $s(\underline{x})$ az n -dimenziós sűrűségfüggvény. Az egyszerűség kedvéért tegyük fel, hogy a várható érték $\underline{\mu} = M(\underline{x}) = 0$. (Ez nyilván nem lényeges korlátozás, hiszen egyszerű kivonással elérhető.) Jelölje továbbá $\underline{\Omega} = M(\underline{x} \underline{x}^T)$ a *kovariancia-* vagy *szórásmatrixot*. Mivel ez mindenkor pozitív definit, az

$$\underline{\Omega} \underline{u} = \lambda \underline{u} \quad (\underline{u} \in \mathbb{R}^n, \lambda \in \mathbb{R}) \quad (5-4)$$

sajátérték-problémának van n lineárisan független megoldása: $\underline{u}_1, \dots, \underline{u}_n$; ezeket az $\underline{\Omega}$ *sajátvektorainak* nevezzük. Az algebrából ismeretes, hogy létezik olyan sajátvektorokból álló rendszer is, amely \mathbb{R}^n -ben *ortonormált bázist* alkot (lásd F3.2.1.1 alpont) és amelyekhez tartozó λ_j *sajátértékek* nemnegatívak. Rendezzük egy ilyen megoldás sajátvektorait

ugy, hogy $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ legyen. Ezek után bármely sajáttségvektor a

$$\underline{\zeta} = \sum_{j=1}^n \alpha_j \underline{u}_j \quad (5-5)$$

alakba írható, s a komponensek átlagértékére (5-4) alapján az

$$M(\alpha_j^2) = M(\underline{u}_j^T \underline{\zeta} \underline{\zeta}^T \underline{u}_j) = \underline{u}_j^T M(\underline{\zeta} \underline{\zeta}^T) \underline{u}_j = \underline{u}_j^T \underline{\Omega} \underline{u}_j = \lambda_j \quad (5-6)$$

eredményt kapjuk.

Visszatérve az eredeti problémánkra: bebizonyítható (lásd pl. [5-16], [A-16]), hogy ekkor a Karhunen-Loève-transzformáció mátrixa

$$\underline{\underline{K}}^T = (\underline{u}_1, \underline{u}_2, \dots, \underline{u}_m)$$

alakú; vagyis $\underline{\underline{K}}$ az $\underline{\Omega}$ első m sajátvektorából, mint sorokból képzett $m \times n$ méretű mátrix. (Vegyük észre, hogy $\underline{\underline{K}} \underline{u}_j = \underline{e}_j$ ($= \mathbb{R}^n$ j -edik egységvektora), ha $j \leq m$; és 0, ha $j > m$.) Továbbá bármely

$$\underline{v} = \underline{\underline{K}} \underline{\zeta} = \sum_{j=1}^n \alpha_j \underline{\underline{K}} \underline{u}_j = \sum_{j=1}^m \alpha_j \underline{e}_j \quad (5-7)$$

transzformált sajáttségvektor legfeljebb m ($< n$) dimenziós, és a leképezés átlagos hibája

$$M(\|\underline{\zeta} - \underline{v}\|^2) = M(\|\underline{\zeta} - \underline{\underline{K}} \underline{\zeta}\|^2) = \text{minimum}$$

az adott sajátvektor-halmazon. (A $\|\cdot\|$ szimbólum itt és a továbbiakban egy vektor euklidészi normáját jelöli; lásd

F3.2.1.1 alpont, (F3-8) képlet).

Végül (5-7) és (5-6) felhasználásával

$$\begin{aligned} M(\|\underline{v}\|^2) &= M(\underline{v}^T \underline{v}) = M\left(\sum_{j=1}^m \alpha_j \underline{e}_j\right) \left(\sum_{j=1}^m \alpha_j \underline{e}_j\right)^T = \\ &= M\left(\sum_{j=1}^m \alpha_j^2\right) = \sum_{j=1}^m \lambda_j; \end{aligned}$$

vagyis az eredményvektor átlagos normanégyzete az *első m sajátérték összegével* egyezik meg.

A Karhunen-Loève-transzformáció további érdekes és hasznos tulajdonsága, hogy az eredményvektor komponensei korrelálatlanok lesznek.

A transzformáció végrehajtásánál bármilyen m értéknél gyakorlatilag ugyanazt a mátrixot kell használnunk: ezt az ortonormált sajátvektorok által alkotott mátrix első m sora adja. Ha m értékét előre nem ismerjük, az információ megőrzéséhez célszerű olyan kritériumot választani, mely a lehető legnagyobb eredményvektorhoz vezet; pl.:

$$\left(\sum_{j=1}^m \lambda_j\right) \left(\sum_{j=1}^n \lambda_j\right)^{-1} \approx 1.$$

A 30. képen a mintaképen (lásd 13. kép) végrehajtott, kétsávos eredményt adó Karhunen-Loève-transzformációt mutatunk be. (A feldolgozást az [1-31]-ben leírt MIP képfeldolgozó rendszerrel végeztük.)

Lényegtömörítés célfüggvény alapján. A lényegtömörítő eljárásokban keresett - valamilyen szempontból - legfontosabb információk "fontosság"-át alkalmasan választott célfüggvénnyel is lehet mérni; az optimális transzformáció megkeresése ilyen esetben a *célfüggvény optimalizálásának problémájához* vezet. A Karhunen-Loève-

transzformáció esetén pl. ilyen célfüggvény lehet a (minimalizálendő) várható átlagos hiba vagy a (maximalizálendő) várható átlagos hossz.

A leggyakrabban használatos három célfüggvény a következő:

1. *Entrópia*. Egy n dimenziós, folytonos eloszlású valószínűségi változó entrópiáját a

$$H(\mathbf{s}) = - \int_{\mathbb{R}^n} \mathbf{s}(\underline{\zeta}) \log_2(\mathbf{s}(\underline{\zeta})) \, d\underline{\zeta} \quad (5-8/a)$$

összefüggés definiálja (ahol $\mathbf{s}(\underline{\zeta})$ az n -dimenziós sűrűségfüggvény). Diszkrét esetben

$$H = - \sum_{i=1}^{\infty} P(\underline{\zeta}_i) \log_2(P(\underline{\zeta}_i)) \quad (5-8/b)$$

a megfelelő definíció. A feladat az (5-3) szerint olyan $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m < n$) lineáris transzformáció megkeresése, amit a sűrűségfüggvényre alkalmazva a $H(\mathbf{s}_T)$ entrópia maximális lesz (\mathbf{s}_T a transzformált sűrűségfüggvény). Igazolható, hogy *normális eloszlás* esetén, ha a várható érték 0, a megoldást a Karhunen-Loève-transzformáció szolgáltatja.

2. *Divergencia*. Jelöljük a lehetséges d_1, d_2, \dots, d_h osztály *feltételes* sűrűségfüggvényét az egyszerűség kedvéért s_1, s_2, \dots, s_h -val (pontosabban lásd 5.2.2.1 al-pont). A sűrűségfüggvények divergenciáját az

$$E(s_1, \dots, s_h) = \frac{1}{2} \sum_{j=1}^h \sum_{i=1}^h \int_{\mathbb{R}^n} [s_j(\underline{\zeta}) - s_i(\underline{\zeta})] \log_2\left(\frac{s_j(\underline{\zeta})}{s_i(\underline{\zeta})}\right) \, d\underline{\zeta} \quad (5-9/a)$$

kifejezés definiálja. Ez a gyakorlatban fontos kétosztályos esetben az

$$E(s_1, s_2) = \int_{\mathbb{R}^n} [s_2(\underline{\zeta}) - s_1(\underline{\zeta})] \log_2\left(\frac{s_2(\underline{\zeta})}{s_1(\underline{\zeta})}\right) \, d\underline{\zeta} \quad (5-9/b)$$

alakra egyszerűsödik. A divergencia az egyes osztályok közötti eltérést méri: minél nagyobb, annál jobban különböznek az osztályok, annál kisebb az osztályozás hibája. Egy tetszőleges $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m < n$) lineáris transzformáció esetén a transzformáltak $E(s_{1T}, s_{2T})$ divergenciája kisebb lehet, mint az eredeti osztályoké. A legjobb transzformációhoz az $E(s_{1T}, s_{2T})$ maximálásával jutunk.

3. *Bhattacharyya-együttható*. Két osztály esetén szokás a *Bhattacharyya-együtthatót* a

$$\beta(s_1, s_2) = \int_{\mathbb{R}^n} \sqrt{s_1(\underline{\zeta}) \cdot s_2(\underline{\zeta})} d\underline{\zeta} \quad (5-10/a)$$

összefüggéssel definiálni. (A β mérőszámot gyakran Hellinger-integrálnak is nevezik.) Az (5-10/a)-ból a

$$\rho_B(s_1, s_2) = -\log_2(\beta(s_1, s_2)) \quad (5-10/b)$$

segítségével definiáljuk a *Bhattacharyya-távolságot*. A lényegtömörítés ezuttal a $\rho_B(s_1, s_2)$ maximalizálása útján hajtható végre. Két megjegyzés:

- a $\beta(s_1, s_2)$ segítségével meg lehet becsülni a Bayes-módszer (lásd 5.2.2 alpont) hibáját: $R_0^* \leq \frac{1}{2} \beta(s_1, s_2)$;
- a Bhattacharyya-távolság *nem metrika*, mivel nem elégíti ki a háromszög-egyenlőtlenséget.

Megjegyezzük, hogy a fentiektől lényegesen eltérő lényegtömörítő (dimenziócsökkentő) eljárások is léteznek. Részletesebb ismertetés nélkül utalunk a *Foley-Sammon*-[5-10], illetve a *Fukunaga-Koontz*-[5-14] eljárásra, amelyek az osztályok szeparáltságát is figyelembe veszik. Az utóbbi ugyancsak a Karhunen-Loève-transzformáción alapszik.)

5.2.1.3 A döntési feladat megfogalmazása

Megadjuk a döntési feladat pontos megfogalmazását, amelyet a továbbiakban is használni fogunk. Az itt bevezetett jelölések a továbbiakban ugyancsak érvényesek lesznek anélkül, hogy jelentésükre ismételten kitérnénk.

Feltesszük egy $D = \{d_1, d_2, \dots, d_h\}$ döntéshalmaz létezését, amely definiálja a mért objektumok lehetséges osztályait.

Az osztályozási feladat egy

$$\delta: \mathbb{R}^n \rightarrow D; \underline{z} \mapsto d_i \quad (5-11)$$

leképezés előállítására; vagyis a sajátástér egy eleméhez egy megfelelően kiválasztott D -beli elemet kell rendelnünk. Magyarán szólva ez azt jelenti, hogy a δ döntéssel valamelyik (d_i nevű) lehetséges osztályba soroljuk a \underline{z} sajátástvektorral jellemzett objektumot. Ezzel tulajdonképpen az R_i diszjunkt régiókra bontottuk az \mathbb{R}^n teret, amelyek pontosan megfelelnek a lehetséges osztályoknak:

$$\mathbb{R}^n = \bigcup_{i=1}^h R_i \quad (R_i \cap R_j = \emptyset, \quad \forall i \neq j); \quad (5-12/a)$$

$$R_i = \{\underline{z} \in \mathbb{R}^n \mid \delta(\underline{z}) = d_i\} \quad (i = 1, \dots, h). \quad (5-12/b)$$

A sajátástér (a mérések eredménye, azaz az objektumok leírása) lehetne sokkal általánosabb is: nem kellene feltétlenül valós értékekre szorítkoznunk. Bizonyos esetekben másfajta "sajátástértékek" a szokásos gyakorlathoz közelebb álló leírást eredményeznének, egyszersmind jelentősen bonyolultabbá tennék a vizsgálatokat. Az \mathbb{R}^n választás a gyakorlatban általában elégségesnek bizonyul. További általánosítási lehetőség a D döntéshalmaz mó-

*dosít*ása. Nem feltétlenül szükséges *véges* elemű döntéshalmazzal dolgoznunk: választhatunk végtelen számosságú is. Mi a továbbiakban is megmaradunk a véges döntéshalmaz mellett. A döntésfüggvényről a korábban mondottak szerint feltételezzük, hogy a teljes M^n téren értelmezve van. Természetesen előfordulhat, hogy ez nem áll fenn; ekkor célszerű a döntéshalmazt kiegészíteni egy d_0 döntéssel, melynek jelentése "nincs döntés":

$$\bar{D} = D \cup \{d_0\}.$$

A kiterjesztett döntésfüggvényt most a következőképpen definiálhatjuk:

$$\bar{\delta}(\underline{z}) = \begin{cases} d_0, & \text{ha nem volt döntés;} \\ \delta(\underline{z}), & \text{különben.} \end{cases}$$

Az előforduló feladatok egy jelentős részében bizonyos objektumok besorolása valamely osztályba *csak erőszakoltan* képzelhető el, valójában *nincs olyan* osztály, melyhez szervesen hozzátartoznának. Ilyenkor gyakorlati megfontolások alapján célszerű olyan döntést hozni, hogy az objektum "nem besorolható", azaz "nincs döntés". Ezzel tulajdonképpen azt tesszük, hogy az eredeti döntésfüggvény értelmezési tartományát - valamilyen gyakorlati-technikai kritérium alapján - beszűkitjük egy "értelmes" döntéstartományra, majd az így kapott szűkebb döntésfüggvényt terjesztjük ki az előbbi módon.

A döntésfüggvényről általában feltesszük, hogy *egyértelmű*, azaz egy objektumhoz egy osztályt rendel. Valójában ez a kikötés nem teljesen indokolt, hiszen az osztályok nem feltétlenül jelentenek kizáró eseteket. Ez a probléma azonban egyszerűen kiküszöbölhető: a D döntéshalmaz helyett az összes részhalmazainak ($2^{|D|}$ számú) halmazát tekintve döntéshalmaznak, a döntésfüggvény egyér-

telművé válik. A gyakorlati megoldásokat is pontosan ezen az úton valósítják meg: a lehetséges sajátságokból csoportokat alkotnak.

A döntésfüggvénynek számos, sokszor ellentmondó követelményt kell kielégítenie. Ezek közül legfontosabb a *pontosági követelmény*, vagyis hogy a mérések alapján lehetőleg minden objektumot a valóságosnak megfelelő osztályba soroljunk. A döntések pontatlansága alapvetően a következő okokra vezethető vissza:

- a mérési eredmények valószínűségi változók, ennél fogva a mérés eredménye azonos osztályba tartozó objektumokra is eltérő lehet;
- a mért sajátságok nem feltétlenül hordozzák az osztályok elkülönítéséhez szükséges összes információt; vagyis különböző osztályokba tartozó objektumokra is mérhetünk azonos sajátságvektorokat.

Másik fontos követelmény, hogy a döntést realizáló algoritmusnak lehetőleg könnyen programozhatónak és a programnak "olcsón" futtathatónak kell lennie.

5.2.2 BAYES-DÖNTÉSELMÉLET

5.2.2.1 Bayes-típusú döntések

Az alakfelismerési feladat megoldása során természetes kiváncsóság olyan döntésfüggvények választása, amelyek hibája - a rossz döntések száma - a lehető legkisebb. A korábban bevezetett jelöléseket megtartva feltételezzük az alábbiakat:

- ismert a $D = \{d_1, d_2, \dots, d_h\}$ döntéshalmaz, mely az osztályokat definiálja;
- ismertek az egyes osztályok $p_i = P(d_i)$ ($i = 1, 2, \dots, h$) "a priori" valószínűségei;

- ismertek az egyes osztályok $s(\underline{\zeta}|d_i)$ feltételes sűrűségfüggvényei, melyek a $\underline{\zeta}$ valószínűségi változó eloszlását határozzák meg arra az esetre, ha tudjuk, hogy az i -edik osztályhoz tartozó objektumról van szó;
- adottak az $L_{ij} \geq 0$ ($i, j = 1, 2, \dots, h$) veszteségi sulyok; ezek tulajdonképpen a téves besorolások miatt bekövetkező veszteségeket fejezik ki: az L_{ij} suly a "büntetés" azért, hogy az i -edik osztályba tartozó objektumot a sajáttságvektora alapján a j -edik osztályba soroltuk; legegyszerűbb esetben az

$$L_{ij} = \begin{cases} 1, & \text{ha } i \neq j; \\ 0, & \text{ha } i = j \end{cases} \quad (5-13)$$

sulyokat használjuk ("bináris veszteségmátrix"). Jelölje $P(d_i|\underline{\zeta})$ annak "a posteriori" valószínűségét, hogy egy konkrét mérési eredmény (sajáttságvektor) éppen a d_i osztályba tartozó objektumot határoz meg. Ekkor az

$$R(d_i|\underline{\zeta}) = \sum_{j=1}^h L_{ij} P(d_j|\underline{\zeta}) \quad (5-14)$$

összefüggés definiálja a feltételes döntési kockázatot, amelyet azzal vállalunk, hogy a $\underline{\zeta}$ sajáttságvektorral jellemzett objektumot a d_i osztályba soroljuk. Az átlagos veszteséget a δ döntésfüggvény esetén pedig az

$$R_\delta = \int_{\mathbb{R}^n} \sum_{j=1}^h R(\delta(\underline{\zeta})|\underline{\zeta}) s(\underline{\zeta}) d\underline{\zeta} \quad (5-15)$$

kifejezés alapján számíthatjuk ki; ahol a teljes valószínűség tétele (F2-5) alapján

$$s(\underline{\zeta}) = \sum_{i=1}^h s(\underline{\zeta}|d_i) P(d_i).$$

Bayes-döntésről akkor beszélünk, ha az átlagos veszteség minimális, vagyis $R_B = \min(R_\delta)$. Ezt a minimális veszteséget nevezzük Bayes-vesztésnek. Tekinthetjük ezt az adatokban lévő belső bizonytalanság mértékének is: semmilyen döntéssel sem lehet ennél kisebb átlagos veszteséget elérni.

A veszteségfüggvény minimalizálásához elégséges a feltételes döntési kockázatot minimalizálni. Ezek alapján $\delta_B(\underline{\zeta}) = d_i$, vagyis a $\underline{\zeta}$ sajáttságvektor alapján az objektumot akkor soroljuk minimális veszteséggel a d_i osztályba, ha

$$R(d_i | \underline{\zeta}) \leq R(d_m | \underline{\zeta}) \quad (m = 1, 2, \dots, h).$$

Helyettesítsük be (5-14)-et, ekkor a minimális veszteségű döntés feltétele:

$$\sum_{j=1}^h L_{ij} P(d_j | \underline{\zeta}) \leq \sum_{j=1}^h L_{mj} P(d_j | \underline{\zeta}) \quad (m = 1, 2, \dots, h). \quad (5-16)$$

(Ha az egyenlőség több osztályra is fennáll, tetszős szerint választhatjuk bármelyiket.) Alkalmazzuk (5-16)-ra az (F2-6)-ban levezetett Bayes-tételt, ekkor a

$$\frac{\sum_{j=1}^h L_{ij} s(\underline{\zeta} | d_j) P(d_j)}{\sum_{j=1}^h s(\underline{\zeta} | d_j) P(d_j)} \leq \frac{\sum_{j=1}^h L_{mj} s(\underline{\zeta} | d_j) P(d_j)}{\sum_{j=1}^h s(\underline{\zeta} | d_j) P(d_j)}$$

illetve - mivel a nevezők egy adott mérés esetén egyenlők - végül is a

$$\sum_{j=1}^h L_{ij} s(\underline{\zeta} | d_j) P(d_j) \leq \sum_{j=1}^h L_{mj} s(\underline{\zeta} | d_j) P(d_j) \quad (5-17)$$

($m = 1, 2, \dots, h$) feltételhez jutunk. Ez a *bináris veszteségmátrixu* speciális esetben a következőre egyszerűsödik:

$$s(\underline{\zeta}|d_i)P(d_i) \geq s(\underline{\zeta}|d_m)P(d_m) \quad (m = 1, 2, \dots, h) \quad (5-18)$$

Ha definiáljuk a

$$g_i(\underline{\zeta}) = \sum_{j=1}^h L_{ij} s(\underline{\zeta}|d_j)P(d_j) \quad (i = 1, 2, \dots, h) \quad (5-19)$$

diszkriminánsfüggvényeket (részletesebben lásd 5.2.3.1 alpont), akkor a Bayes-döntést ezek segítségével hozhatjuk meg. Speciálisan a bináris veszteségmátrixu esetben

$$g_i(\underline{\zeta}) = s(\underline{\zeta}|d_i)P(d_i) \quad (5-20)$$

alaku függvényeket használhatunk. Ekkor a várható átlagos veszteség megegyezik a várható *hibaarány*al, vagyis (5-15) és (5-20) felhasználásával

$$R_B = 1 - \int_{\mathbb{R}^n} \max_{1 \leq i \leq h} (P(d_i|\underline{\zeta})) \sum_{j=1}^h s(\underline{\zeta}|d_j)P(d_j) d\underline{\zeta}. \quad (5-21)$$

Célszerű a döntés vizsgálatát a $h = 2$ esetre külön is elvégezni. A döntési szabályt (5-17) szerint azonnal felírhatjuk:

$$\delta_B(\underline{\zeta}) = \begin{cases} d_1, & \text{ha } L_{11}s(\underline{\zeta}|d_1)P(d_1) + L_{12}s(\underline{\zeta}|d_2)P(d_2) \leq \\ & \leq L_{21}s(\underline{\zeta}|d_1)P(d_1) + L_{22}s(\underline{\zeta}|d_2)P(d_2); \\ d_2, & \text{különben;} \end{cases}$$

azaz

$$\delta_B(\underline{\zeta}) = \begin{cases} d_1, & \text{ha } \frac{L_{12} - L_{22}}{L_{21} - L_{11}} \frac{P(d_2)}{P(d_1)} \leq \frac{s(\underline{\zeta}|d_1)}{s(\underline{\zeta}|d_2)}; \\ d_2, & \text{különben.} \end{cases} \quad (5-22)$$

A bináris veszteségmátrixu speciális esetben ez egyszerűen a

$$\delta_B(\underline{\zeta}) = \begin{cases} d_1, & \text{ha } \frac{P(d_2)}{P(d_1)} \leq \frac{s(\underline{\zeta}|d_1)}{s(\underline{\zeta}|d_2)}; \\ d_2, & \text{különben.} \end{cases} \quad (5-23)$$

szabályra redukálódik. Ekkor a várható hibaarány:

$$R_B = \frac{1}{2} - \frac{1}{2} \int_{\mathbb{R}^n} |P(d_2|\underline{\zeta}) - P(d_1|\underline{\zeta})| s(\underline{\zeta}) d\underline{\zeta}. \quad (5-24)$$

5.2.2.2 A normális eloszlás esete

A továbbiakban - a gyakorlatban legtöbbször feltételezett - *n*-dimenziós normális eloszlás esetével foglalkozunk. Ekkor az *i*-edik osztály feltételes sűrűségfüggvénye

$$s(\underline{\zeta}|d_i) = \frac{1}{\sqrt{(2\pi)^n |\underline{\Omega}_i|}} \exp\left[-\frac{1}{2}(\underline{\zeta} - \underline{\mu}_i)^T \underline{\Omega}_i^{-1} (\underline{\zeta} - \underline{\mu}_i)\right]$$

ahol $\underline{\mu}_i$ az *i*-edik osztály várhatóérték-vektora, $\underline{\Omega}_i$ pedig a szórásmátrixa. Bináris veszteségmátrixot feltételezve a Bayes-döntést az (5-20) szerint kiszámított

$$g_i(\underline{\zeta}) = s(\underline{\zeta}|d_i)P(d_i) = \frac{P(d_i)}{\sqrt{(2\pi)^n |\underline{\Omega}_i|}} \exp\left[-\frac{1}{2}(\underline{\zeta} - \underline{\mu}_i)^T \underline{\Omega}_i^{-1} (\underline{\zeta} - \underline{\mu}_i)\right]$$

($i = 1, 2, \dots, h$) diszkriminánsfüggvények segítségével hozhatnók meg, célszerűbb azonban helyettük a logaritmusukból származtatott újabb függvényekkel dolgozni. Vezessük be a

$$\tilde{g}_i(\underline{\zeta}) = b_i - (\underline{\zeta} - \underline{\mu}_i)^T \underline{\Omega}_i^{-1} (\underline{\zeta} - \underline{\mu}_i) \quad (5-25)$$

diszkriminánsfüggvényeket, ahol $b_i = \ln(P(d_i)) - \ln(|\underline{\Omega}_i|)$. A gyakorlati megvalósítás során a b_i konstansok előzetesen kiszámíthatók, a továbbiakban csak a $(\underline{\zeta} - \underline{\mu}_i)^T \underline{\Omega}_i^{-1} (\underline{\zeta} - \underline{\mu}_i)$ kifejezést kell kiértékelni. A fentiekből az is kiderül, hogy a számításigény az osztályok számával lineárisan, a dimenziószámmal négyzetesen nő.

Megvizsgálunk még további három speciális esetet:

1. Ha az egyes osztályok szórásmatrixa azonos (jelöljük most ezt $\underline{\Omega}$ -val), akkor (5-25) helyett a

$$\tilde{g}_i(\underline{\zeta}) = b_i^* + \underline{w}_i \underline{\zeta} - \underline{\zeta}^T \underline{\Omega}^{-1} \underline{\zeta}$$

diszkriminánsfüggvénnyel számolhatunk, ahol

$$b_i^* = b_i - \underline{\mu}_i^T \underline{\Omega}^{-1} \underline{\mu}_i \quad \text{és} \quad \underline{w}_i = 2 \underline{\mu}_i^T \underline{\Omega}^{-1}.$$

Mivel a $\underline{\zeta}^T \underline{\Omega}^{-1} \underline{\zeta}$ kifejezés mindegyik függvényben szerepel, a döntést nem befolyásolja: elhagyható.

2. Ha az előbbi mellett feltesszük, hogy a sajátságvektorok komponensei függetlenek és szórásuk minden koordinátában azonos, azaz $\underline{\Omega}_i = \sigma^2 \underline{E}$, (\underline{E} az n -edrendű egységmatrix), akkor az előbbi b_i és \underline{w}_i konstansok egyszerűen számolhatók. Az így meghatározott *diszkriminánsfelületek* merőlegesek lesznek az osztályok várhatóérték-vektorait összekötő szakaszokra.
3. Másfajta egyszerűsítési lehetőség adódik, ha az a priori valószínűségeket nem ismerjük, ezért azonosnak tételezzük fel. Ekkor a döntési szabály a következő lesz: egy objektumot akkor sorolunk az i -edik osztályba, ha sajátságvektorának előfordulási valószínűsége abban a legnagyobb.

A Bayes-féle döntéseméletnek igen nagy szakirodalma van; további tanulmányozásra javasolhatjuk pl. az [A-6], [A-8], [A-16], [5-16], [5-20] stb. műveket.

A 31. képen a 13. kép (mintakép) Bayes-féle osztályozására mutatunk példát, normális eloszlást feltételezve. (A feldolgozás az [1-31]-ben leírt MIP rendszerrel készült.)

5.2.2.3 Paraméterbecslés

Az osztályozás végrehajtásához általában szükséges döntésfüggvényeket csak az egyes osztályok eloszlásainak, az eloszlások bizonyos paramétereinek (pl. várhatóértékvektor (átlag), szórásmatrix stb.) ismeretében tudjuk definiálni. Az eddigiekben feltételeztük, hogy ezek pontos értéke rendelkezésünkre áll, a gyakorlatban azonban ez rendszerint nincs így. A döntések pontosságát viszont - és ez a következőkre is vonatkozik, nem csak a Bayes-féle döntésekre - befolyásolja ez a körülmény, amelyet eddig nem vettünk figyelembe.

A pontosan eleve nem ismert paramétereket valamilyen *tananyag* alapján kell becsléssel meghatároznunk (lásd pl. [5-16], [F-2] stb.). A becslési módszerekről - bár nem tartoznak szorosan a témánkhoz - feltétlenül legalább említést kell tennünk.

Normális eloszlású valószínűségi változók várhatóértékére és szórásmatrixára a következő becslések használatosak:

Legyen $Z = \{\xi_1, \dots, \xi_N\}$ normális eloszlású és független valószínűségi változók egy halmaza. Ekkor a

$$\hat{\mu} = \frac{1}{N} \sum_{j=1}^N \xi_j;$$

$$\hat{\underline{\Omega}} = \frac{1}{N} \sum_{j=1}^N (\underline{\zeta}_j - \hat{\underline{\mu}})(\underline{\zeta}_j - \hat{\underline{\mu}})^T;$$

$$\tilde{\underline{\Omega}} = \frac{N}{N-1} \hat{\underline{\Omega}}$$

definíciókkal adott $\hat{\underline{\mu}}$, illetve $\hat{\underline{\Omega}}$ a várhatóérték-vektorok, illetve a szórásmátrixuk *legnagyobb valószínűségű* (maximum likelihood) becslése; míg $\tilde{\underline{\Omega}}$ a *korrigált empirikus szórás-mátrix*. Igazolható, hogy N növekedésével a kapott $\hat{\underline{\mu}}$ és $\tilde{\underline{\Omega}}$ egyre pontosabban megközelíti az elméleti várhatóértéket és szórás-mátrixot.

Más módszerrel kell a becslést végeznünk, ha az eloszlás típusáról nincs semmiféle információnk. Most csak az *empirikus sűrűségfüggvény* meghatározásával foglalkozunk.

Az első esetben \mathbb{R}^n -t felosztjuk h régióra: R_1, \dots, R_h . Legyen az R_i régió térfogata V_i , és legyen $c_i = |\{\underline{\zeta}_j \in Z \mid \underline{\zeta}_j \in R_i\}|$. Ekkor a közelítő sűrűségfüggvény

$$\hat{s}(\underline{\zeta}) = \frac{c_i}{hV_i} \quad (i = 1, 2, \dots, h);$$

minden $\underline{\zeta}_j \in R_i$ esetén.

Másik lehetőség a *Rosenblatt-Parzen-típusú empirikus sűrűségfüggvények* használata. Legyen $s: \mathbb{R}^n \rightarrow \mathbb{R}$ olyan szimmetrikus sűrűségfüggvény, amely kétszer korlátosan differenciálható, és teljesíti a következő feltételeket:

$$\int_{\mathbb{R}^n} s(\underline{\zeta}) \, d\underline{\zeta} = 1;$$

$$\int_{\mathbb{R}^n} \underline{\zeta} s(\underline{\zeta}) \, d\underline{\zeta} = 0;$$

$$\int_{\mathbb{R}^n} \|\underline{\zeta}\|^2 s(\underline{\zeta}) \, d\underline{\zeta} < +\infty.$$

Legyen továbbá $a_j \gg 0$, $\lim_{j \rightarrow \infty} a_j = 0$ tetszőleges számsorozat. Ekkor az

$$\hat{s}(\underline{\zeta}) = \frac{1}{Na_N^n} \sum_{j=1}^N s\left(\frac{\zeta - \zeta_j}{a_N}\right)$$

segítségével definiált empirikus sűrűségfüggvény ad használható becslést. A fenti képletekben $s(\underline{\zeta})$ -ként gyakran a normális eloszlás sűrűségfüggvényét szokták használni. Megjegyezzük, hogy a megfelelő a_j számsorozat konkrét meghatározása komoly problémát okozhat.

5.2.3 LINEÁRIS DISZKRIMINÁNSFÜGGVÉNYEK

5.2.3.1 Döntés diszkriminánsfüggvények alapján

Diszkriminánsfüggvények szerinti döntésről akkor beszélünk, ha valamilyen

$$g_i: \mathbb{R}^n \rightarrow \mathbb{R} \quad (i = 1, 2, \dots, h) \quad (5-26)$$

valós értékű függvények felhasználásával a

$$\delta(\underline{z}) = d_i$$

döntést akkor és csak akkor hozzuk (másszóval: a \underline{z} saját-ságvektorral jellemzett objektumot akkor és csak akkor soroljuk az i -edik osztályba), ha

$$g_i(\underline{z}) \geq g_m(\underline{z}) \quad (m = 1, 2, \dots, h). \quad (5-27)$$

Nyilvánvalóan igazak a következők:

— a diszkriminánsfüggvények egyértelműen meghatározzák a döntésfüggvényt;

- tetszőleges, monoton növekedő $f: \mathbb{R} \rightarrow \mathbb{R}$ függvény esetén a $h_i = f \cdot g_i$ függvényrendszer is ugyanazt a döntésfüggvényt határozza meg.

Az (5-26) típusú diszkriminánsfüggvényekkel tulajdonképpen h számú diszjunkt régióra bontjuk az \mathbb{R}^n döntési teret, s a sajátságvektorokat - amelyek n -dimenziós pontokat reprezentálnak - az (5-27) típusú döntéssel valamelyik régióba soroljuk. Az egyes régiókat a

$$g_i(\underline{z}) = g_m(\underline{z}) \quad (i \neq m; i, m = 1, 2, \dots, h) \quad (5-28)$$

egyenletekkel meghatározott n -dimenziós hiperfelületek határolják.

A továbbiakban a lineáris esetre szorítkozunk. Ekkor az (5-26) szerinti függvények felírhatók

$$g_i(\underline{z}) = \underline{a}_i^T \underline{z} + b_i \quad (i = 1, 2, \dots, h) \quad (5-26/a)$$

alakban, ahol $\underline{a}_i \in \mathbb{R}^n$ adott súlyvektorok, $b_i \in \mathbb{R}$ pedig adott valós küszöbsúlyok. Egy lineáris diszkriminánsfüggvény-rendszer esetén a létrejövő döntési régiókat *hipersíkok* határolják, amelyeket (5-28) szerint és (5-26/a) alapján az

$$(\underline{a}_i - \underline{a}_m)^T \underline{z} + (b_i - b_m) = 0 \quad (i \neq m; i, m = 1, 2, \dots, h)$$

egyenletek definiálnak. Ennélfogva az $(\underline{a}_i - \underline{a}_m)$ különbségvektor *merőleges* a döntési felületre; továbbá egy tetszőleges \underline{z} pontnak a döntési felülettől mért (előjeles) távolsága

$$\rho_{ij}(\underline{z}) = \frac{g_i(\underline{z}) - g_m(\underline{z})}{\|\underline{a}_i - \underline{a}_m\|}.$$

Kétosztályos esetben a probléma lényegesen egyszerűbb; ilyenkor a döntéshez egyetlen lineáris függvényt szokás használni; amely

$$D(\underline{z}) = g_1(\underline{z}) - g_2(\underline{z})$$

alaku, a döntésfüggvény pedig

$$\delta(\underline{z}) = \begin{cases} d_1, & \text{ha } D(\underline{z}) > 0; \\ d_2, & \text{ha } D(\underline{z}) < 0. \end{cases}$$

5.2.3.2 Lineáris szeparálás

Tegyük fel, hogy meg vannak adva a $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_t$ mintavektorok, amelyeknek ismerjük az osztályozását. Egyelőre tételezzük fel, hogy csupán két osztály létezik. Azt mondjuk, hogy a minták lineárisan szeparálható rendszert alkotnak, ha van olyan g lineáris diszkriminánsfüggvény, melyre

$$\begin{aligned} g(\underline{w}_j) > 0, & \text{ ha } \delta(\underline{w}_j) = d_1 \\ g(\underline{w}_j) < 0, & \text{ ha } \delta(\underline{w}_j) = d_2 \end{aligned} \quad (j = 1, 2, \dots, t). \quad (5-29)$$

Ha helyettesítjük a második osztályhoz tartozó mintákat a negáltjukkal, ekkor $g(-\underline{w}_j) > 0$ adódik. Legyen tehát

$$\underline{v}_j = \begin{cases} \underline{w}_j, & \text{ha } \delta(\underline{w}_j) = d_1 \\ -\underline{w}_j, & \text{ha } \delta(\underline{w}_j) = d_2 \end{cases} \quad (j = 1, 2, \dots, t). \quad (5-30)$$

Ekkor az (5-29) definíciót úgy módosíthatjuk, hogy minden \underline{v}_j -re a $g(\underline{v}_j) > 0$ feltételt követeljük meg. A továbbiakban a \underline{w}_j jelöléseket használjuk az eredeti, a \underline{v}_j -t az (5-30) szerint módosított mintákra.

Mivel g lineáris diszkriminánsfüggvény, valamilyen \underline{a} súlyvektor és b küszöbsúly mellett előállítható $g(\underline{v}_j) = \underline{a}^T \underline{v}_j - b$ alakban; ezzel az (5-30) feltétel az

$$\underline{a}^T \underline{v}_j \geq b \quad (j = 1, 2, \dots, t) \quad (5-31)$$

alakot ölti. Az így kapott \underline{a} -t *szeperáló* vagy *megoldásvektornak*, a b számot pedig egyszerűen *küszöbnek* is szokás nevezni.

A megoldásvektor (ha létezik), nyilvánvalóan *nem egyértelmű*: létezik több is, amely ugyanazt a döntésfüggvényt határozza meg. Szokás például azt a súlyvektort keresni, amelynek abszolútértéke 1, és minimalizálja a mintavektorok végpontjának (a továbbiakban: mintapontoknak) a szeperáló hipersíktól vett maximális távolságát. Másik lehetőség az (5-31) egyenlőtlenséget kielégítő minimális méretű vektor előállítása.

A lineáris szeperálás általánosítható a következő módon: a $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_t$ mintákat lineárisan szeperálhatónak nevezzük, ha léteznek olyan $g_1(\underline{w}), g_2(\underline{w}), \dots, g_h(\underline{w})$ lineáris diszkrimináns függvények, hogy a

$$g_i(\underline{w}_j) > g_m(\underline{w}_j), \text{ ha } \delta(\underline{w}_j) = d_i \text{ és } i \neq m$$

reláció fennáll minden \underline{w}_j mintára.

5.2.3.3 Lineáris diszkriminánsfüggvények előállítása

Az igen nagy számú módszer közül - fontosságuk és gyakoriságuk miatt - ötöt mutatunk be részletesebben.

1. *Gradiensmódszer.* Legyen adott a $J(\underline{a})$ kritériumfüggvény, keressük azt az \underline{a} megoldásvektort, amely kielégíti az

$$\underline{a}^T \underline{v}_j > 0 \quad (j = 1, 2, \dots, t) \quad (5-32)$$

egyenlőtlenségeket, és minimalizálja a kritériumfüggvényt.

Tegyük fel, hogy a J minden változója szerinti parciális differenciálhányadosa létezik és folytonos, továbbá jelölje $\nabla J(\underline{a})$ a gradiensvektort. Válasszunk egy tetszőleges $\underline{a}(1)$ kiindulási vektort, és keressük a megoldást a következő rekurzív algoritmus szerint:

$$\underline{a}(\tau + 1) = \underline{a}(\tau) - \alpha_\tau \cdot \nabla J(\underline{a}(\tau)) \quad (\tau = 1, 2, 3, \dots); \quad (5-33)$$

ahol α_τ adott pozitív számok, amelyek a lépésnagyságot befolyásolják.

Az (5-33)-mal definiált *gradiensmódszer* képezi az alapját a továbbiakban ismertetendő eljárásoknak is.

2. *Perceptronmódszer.* A *perceptronkritérium-függvényt* a hibásan osztályozott mintapontoknak a szeparáló hipersíktól mért távolsága összegével definiáljuk:

$$J_P(\underline{a}) = \sum_{\underline{a}^T \underline{v} < 0} (-\underline{a}^T \underline{v}). \quad (5-34)$$

A célfüggvény minimalizálásával azt a szeparáló hipersíkot keressük, amelytől az esetleges "rosszul" osztályozott mintapontok a *lehető legkisebb* távolságra vannak.

Alkalmazzuk a gradiensmódszert a J_P célfüggvényre. Az (5-33) és az (5-34) alapján adódik:

$$\underline{a}(\tau + 1) = \underline{a}(\tau) + \alpha_{\tau} \frac{\Sigma}{\underline{a}^T(\tau) \underline{v} < 0} \underline{v} \quad (\tau = 1, 2, 3, \dots).$$

Geometriailag ez azt jelenti, hogy az $\underline{a}(\tau)$ -t módosítjuk azoknak a mintapontoknak a hozzávételével, amelyeket az általa definiált hipersík rosszul szeparál. Igazolható, hogy ha a mintahalmaz lineárisan szeparálható, akkor a perceptronmódszer véges sok lépésben konvergál. Ha a minták nem szeparálhatók, a módszer divergens.

Több osztály esetére a módszert a következőképpen általánosíthatjuk. Jelölje h az osztályok számát és

$$\{W\} = \underline{w}_1, \underline{w}_2, \dots, \underline{w}_t, \underline{w}_1, \underline{w}_2, \dots, \underline{w}_t, \underline{w}_1, \dots$$

a mintavektorok ciklikus felsorolásával keletkező sorozatot. Jelölje ennek j -edik elemét \tilde{w}_j , továbbá jelölje $\underline{a}_i(\tau)$ az iteráció során a τ -adik lépésben az i -edik osztályhoz definiált súlyvektort.

Ha a \tilde{w}_j -t az $\{\underline{a}_i(\tau)\}$ súlyvektorrendszer j öd osztályozza, vagyis ha \underline{w}_j eredetileg az m -edik osztályhoz tartozik, és egyidejűleg

$$\underline{a}_m^T(\tau) \tilde{w}_j > \underline{a}_i^T(\tau) \tilde{w}_j \quad (i, m = 1, 2, \dots, h; j = 1, 2, \dots, t)$$

is fennáll minden $i \neq m$ esetén, akkor a súlyvektorokat nem változtatjuk:

$$\underline{a}_i(\tau + 1) = \underline{a}_i(\tau) \quad (i = 1, 2, \dots, h).$$

Ha viszont a döntésfüggvények szerint \tilde{w}_j -t az i -edik osztályba kell sorolni, azaz

$$\underline{a}_m^T(\tau)\tilde{w}_j < \underline{a}_i^T(\tau)\tilde{w}_j \quad (i, m = 1, 2, \dots, h; j = 1, 2, \dots, t);$$

akkor legyen

$$\underline{a}_i(\tau + 1) = \underline{a}_i(\tau) - \tilde{w}_j$$

$$\underline{a}_m(\tau + 1) = \underline{a}_m(\tau) + \tilde{w}_j$$

$$\underline{a}_s(\tau + 1) = \underline{a}_s(\tau) \quad (s = 1, 2, \dots, h; s \neq i, m).$$

3. *Ho-Kashyap-módszer.* A Ho-Kashyap-módszer esetén a minimalizálando függvény

$$J_s(\underline{a}, \underline{b}) = \|\underline{V} \underline{a} - \underline{b}\|^2$$

ahol $\underline{V}^T = [\underline{v}_1, \underline{v}_2, \dots, \underline{v}_t]$ az (5-30) szerint módosított mintákból alkotott mátrix, \underline{a} és \underline{b} sulyvektorok, és $\underline{b}(\tau) > 0$. A J_s gradiensvektorai:

$$\nabla_{\underline{a}} J_s = 2\underline{V}^T (\underline{V} \underline{a} - \underline{b})$$

$$\nabla_{\underline{b}} J_s = -2(\underline{V} \underline{a} - \underline{b}).$$

Belátható, hogy \underline{b} tetszőleges értéke mellett J_s az \underline{a} -ra nézve akkor minimális, ha

$$\underline{a} = \underline{V}^- \underline{b} = (\underline{V}^T \underline{V})^{-1} \underline{V}^T \underline{b},$$

ahol \underline{V}^- -szal a \underline{V} mátrix általánosított inverzét jelöltük. Ekkor az \underline{a} szerinti gradiens nyilvánvalóan 0, és $\underline{b}(\tau)$ -ra a következő algoritmus adódik (τ a lépésszám):

$$\underline{b}(1) > 0;$$

$$\underline{b}(\tau + 1) = \underline{b}(\tau) + 2\beta \underline{e}^+(\tau);$$

ahol

$$\underline{a}(\tau) = \underline{V}^- \underline{b}(\tau); \underline{e}(\tau) = \underline{V} \underline{a}(\tau) - \underline{a}(\tau); \underline{e}^+(\tau) = \frac{1}{2}(|\underline{e}(\tau)| + \underline{e}(\tau)).$$

A Ho-Kashyap-módszerrel végső soron azt a szeparáló hipersíkot keressük, amelytől mérve az egyes minták távolságának négyzetösszege minimális. Ha a minták lineárisan szeparálhatók, és $0 < \beta < 1$, akkor a módszer véges sok lépésben konvergál a megoldásvektorhoz.

4. *Lineáris programozási módszer alkalmazása.* A lineáris programozás alapfeladata a következő: Keressük azt az $\underline{u}^T = [u^{(1)}, u^{(2)}, \dots, u^{(n)}]$ vektort, amely minimalizálja az $f = \underline{\alpha}^T \underline{u}$ célfüggvényt, az

$$\underline{A} \underline{u} \geq \underline{\beta}$$

feltétel mellett, ahol $\underline{\alpha}$ $n \times 1$ -es költségvektor, $\underline{\beta}$ $m \times 1$ méretű vektor; és \underline{A} $m \times n$ méretű mátrix. A feladatra az ismert szimplex módszer $\underline{u} \geq 0$, nemnegatív komponensű megoldást ad.

Alkalmazzuk a szimplex módszert a mérési minták szeparálhatóságának eldöntésére. Legyenek megadva a $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_t$ minták, az (5-30) feltétel mellett. Keresünk olyan \underline{a} szeparáló vektort, mely adott $b_j > 0$ ($j = 1, 2, \dots, t$) számok mellett kielégíti az

$$\underline{a}^T \underline{v}_j \geq b_j > 0 \quad (j = 1, 2, \dots, t)$$

feltételeket.

Fogalmazzuk meg a következő lineáris programozási feladatot: Minimalizálandó az $f = \varphi$ célfüggvény az

$$\begin{aligned} \underline{a}^T \underline{v}_j + \varphi &\geq b_j & (j = 1, 2, \dots, t); \\ \varphi &\geq 0; \end{aligned}$$

feltételek mellett. Tekintve, hogy az eljárás pozitív megoldást szolgáltat, vezessük be az \underline{a} helyett az

$$\underline{a}^+ = \frac{1}{2}(|\underline{a}| + \underline{a}) \quad \text{és} \quad \underline{a}^- = \frac{1}{2}(|\underline{a}| - \underline{a})$$

vektorokat. Ekkor a lineáris programozási feladat a következő módon alakítható át: minimalizálandó az f célfüggvény az

$$\begin{aligned} \underline{a}^{+T} \underline{v}_j - \underline{a}^{-T} \underline{v}_j + \varphi &\geq b_j & (j = 1, 2, \dots, t); \\ \varphi &\geq 0 \end{aligned}$$

feltételek mellett.

A szokásos jelölési módot alkalmazva a feladatot az alábbiak szerint írhatjuk át: minimalizálandó az $f = \underline{\alpha}^T \underline{u}$ célfüggvény az

$$\begin{aligned} \underline{A} \underline{u} &\geq \underline{\beta}; \\ \underline{u} &\geq 0 \end{aligned}$$

feltételek mellett, ahol

$$\underline{A} = \begin{bmatrix} \underline{w}_1^T & -\underline{w}_1^T & 1 \\ \underline{w}_2^T & -\underline{w}_2^T & 1 \\ \vdots & \vdots & \vdots \\ \underline{w}_t^T & -\underline{w}_t^T & 1 \end{bmatrix}; \quad \underline{u} = \begin{bmatrix} \underline{a}^+ \\ \underline{a}^- \\ \varphi \end{bmatrix}; \quad \underline{\alpha} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}; \quad \underline{\beta} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_t \end{bmatrix}.$$

A felírt lineáris programozási feladatnak pontosan akkor van $\underline{u} \geq 0$ megoldása, ha a minták lineárisan szeparálható rendszert alkotnak. A szimplex módszer segítségével a kérdés eldönthető, és egyúttal egy szeparáló \underline{a} vektor is előáll.

5. *Fisher-módszer.* A szokásos jelöléseket használva, a kétosztályos esetre definiáljuk a mintaátlagokat (várhatóértékeket) és a szórásmatrixokat a következő módon:

$$\underline{\mu}_1 = \frac{1}{|C_1|} \sum_{w_j \in C_1} w_j;$$

$$\underline{\mu}_2 = \frac{1}{|C_2|} \sum_{w_j \in C_2} w_j;$$

$$\underline{\Omega}_1 = \frac{1}{|C_1|} \sum_{w_j \in C_1} (w_j - \underline{\mu}_1)(w_j - \underline{\mu}_1)^T;$$

$$\underline{\Omega}_2 = \frac{1}{|C_2|} \sum_{w_j \in C_2} (w_j - \underline{\mu}_2)(w_j - \underline{\mu}_2)^T;$$

ahol C_1 , illetve C_2 jelöli a két osztályt és $|C_1|$, illetve $|C_2|$ az elemeik számát; továbbá legyen

$$\underline{\Omega} = \underline{\Omega}_1 + \underline{\Omega}_2 \quad \text{és} \quad \underline{M} = (\underline{\mu}_1 - \underline{\mu}_2)(\underline{\mu}_1 - \underline{\mu}_2)^T.$$

Adott \underline{a} vektor esetén most a célfüggvényt a

$$J_R(\underline{a}) = \frac{\underline{a}^T \underline{M} \underline{a}}{\underline{a}^T \underline{\Omega} \underline{a}}$$

képlet írja le; ezt szokás általánosított *Raleigh-hányados*nak nevezni.

Egy tetszőleges \underline{a} vektor esetén az osztályok \underline{a} irányu szeparáltságát a

$$\alpha = \underline{a}^T (\underline{\mu}_1 - \underline{\mu}_2)$$

skalársorozat segítségével definiáljuk. Az \underline{a} vektort úgy kívánjuk megválasztani, hogy α a lehető legnagyobb legyen. Tekintve, hogy ez az érték az \underline{a} egyszerű növelésével nő, a maximumot a minták \underline{a} irányu szórásához viszonyítva fogjuk keresni. Ezt a kívánalmat fejezi ki a $J_R(\underline{a})$ függvény. Másszóval a Fischer-módszer esetén azt az irányt keressük meg, amelyik szerint a mintákat a lehető legjobban el lehet választani.

Igazolható ([A-4]; 111 old.), hogy a $J_R(\underline{a})$ függvény maximumhelye az

$$\underline{M} \underline{a} = \lambda \underline{\Omega} \underline{a}$$

általánosított sajátértékprobléma megoldásával kapott legnagyobb sajátértéknél van.

Az itt ismertetett módszerek mindegyike általánosítható többosztályos esetre is, azonban erre nem térünk ki. Részletesebb ismertetés található például [A-4]-ben (130.-178. old.).

5.2.4 TÁVOLSÁGMÉRŐ MÓDSZEREK

Az 5.2.2 alpontban az osztályozás során a mintákról rendelkezésre álló valószínűségi jellegű információkat használtuk fel. Az 5.2.3 alpontban a mintákat valamilyen célfüggvény szerint legjobban elválasztó lineáris diszkriminánsfüggvényeket kerestünk.

A továbbiakban a mintákról, gyakoriságokról, eloszlásokról semmit sem tételezünk fel; a döntések a saját-ságvektorok és a valamilyen módon előre kiválasztott mintavektorok távolságának mérésén alapulnak. A döntés során elvileg tetszőleges távolságfüggvényt lehet használni, gyakorlatban azonban leginkább a nyolcas (lásd

F3.2.1 alpont) és az euklidészi metrikát szokták alkalmazni.

5.2.4.1 "Legközelebbi szomszéd" módszerek

Tegyük fel, hogy ismert a

$$T = \{(\underline{w}_j, \vartheta_j) \mid \underline{w}_j \in \mathbb{R}^n, \vartheta_j \in D, j = 1, \dots, t\}$$

tananyag, azaz a \underline{w}_j mintákból és a hozzájuk tartozó $\vartheta_j \in D$ tanításokból álló *tanítópárok* (besorolási minták) halmaza. A "legközelebbi szomszéd" (LKS) segítségével történő osztályozás alapelve a következő:

Egy objektumot akkor sorolunk az i -edik osztályba, ha a mintapontok közül hozzá legközelebb eső pont az i -edik osztályba tartozik. Képlettel:

$$\delta(\underline{z}) = d_i,$$

ha

$$\rho(\underline{z}, \underline{w}_i) \leq \rho(\underline{z}, \underline{w}_j) \text{ és } (\underline{w}_i, d_i) \in T \quad (i, j = 1, 2, \dots, t); \quad (5-36)$$

ahol $\rho: \mathbb{R}^n \times \mathbb{R}^t \rightarrow \mathbb{R}$ adott távolságfüggvény. Ha a feltétel több osztályra is teljesül, közülük tetszőlegesen választhatunk.

Megfogalmazhatjuk az eljárást diszkriminánsfüggvények segítségével is, a függvényeket például a következő módon definiálva:

$$g_i(\underline{z}) = - \min_{(\underline{w}_i, d_i) \in T} (\rho(\underline{z}, \underline{w}_i)).$$

Feltételezve a mintavektorok függetlenségét és azonos eloszlását, elmondhatjuk a következőket:

- Egy döntés akkor hibás, ha a szabály alkalmazásával az objektumot "rossz" osztályba soroltuk. Igazolható [A-4] 101.old.), hogy a tanítópárok számának növekedésével a hibás döntések száma az

$$R_{LKS} = M \left(1 - \sum_{i=1}^m P^2(d_i | \underline{z}) \right)$$

értékhez konvergál, melyet az LKS-módszer *veszteségének* tekintünk.

- Ugyancsak bebizonyítható ([A-4] 102.old.), hogy ha R_B jelöli a Bayes-döntés várható veszteségét, akkor

$$R_B \leq R_{LKS} \leq 2R_B;$$

vagyis a módszer hibája (természetesen) aszimptotikusan nagyobb, mint a Bayes-döntésé.

Ez a reláció tulajdonképpen azt mutatja, hogy a LKS-módszer - egyszerűsége ellenére - rendkívül hatékony, hiszen hibája *legfeljebb kétszerese* az optimális Bayes-módszerrel elérhetőnek, azaz a döntésben szerepet játszó információk jelentős részét a *legközelebbi mintavektor hordozza*.

Az LKS-módszer általánosítása az (m, n) -LKS módszer, melyet a következő módon definiálunk:

A \underline{z} sajáttségvektor alapján egy objektumot akkor sorolunk a d_i osztályba, ha a hozzá legközelebb eső m tanítópár közül legalább n a d_i osztályhoz tartozik. Ha ez egyetlen d_i -re sem teljesül, akkor nem döntünk; ha többre is, a leggyakoribbat választjuk. A 32. képen a Bayes-féle (a), a doboz- (lásd következő alpont) (b), az LKS- (c) és a (3,2)-LKS (d) módszer döntési régióit ábrázoltuk, 2-dimenziós sajáttségvektorok esetén, az összehasonlítás kedvéért. (Felhívjuk a figyelmet arra, hogy az (m, n) -LKS-módszer *rosszabb* eredményt is adhat, mint az

alapeljárás. A képen pl. az LKS-módszer esetén a bejelölt pontok voltak a tanítópontok; a (3,2)-LKS módszer-nél a legsötétebb területek azokat a pontokat jelzik, ahol a módszer nem adott döntést.)

Visszatérve az alapeljáráshoz: ha az euklidészi, vagy a (4/8)-távolságot használjuk, a szükséges műveletszám a dimenziószámtól *lineárisan* függ, s ez lényegesen kevesebb, mint a Bayes-módszer esetén. Ha a módszer realizálása során minden sajátvektor távolságát kiszámítjuk mindegyik mintavektortól, akkor ezek számával lineárisan nő az elvégzendő műveletszám. Léteznek azonban olyan új módszerek, melyekkel a számítások jelentősen egyszerűsíthetők. Ezek általában bizonyos előfeldolgozás, előrendezés árán csökkentik az eredetinek töredékére az egyes döntések számításigényét, ami esetleg a minták számától csaknem függetlenné is válhat.

5.2.4.2 A dobozmódszer

A LKS-módszerhez hasonló és ugyancsak távolságmérésen alapul a *doboz-*(box-, vagy téglatest-) *módszer*, amelynek tananyaga

$$T = \{(\underline{w}_j, \underline{a}_j, \vartheta_j) \mid \underline{w}_j \in \mathbb{R}^n, \underline{a}_j \in \mathbb{R}^n, \vartheta_j \in D; (j = 1, 2, \dots, t)\}.$$

Ezt úgy kell értelmeznünk, hogy a *tanítóhármás* egy \underline{w}_i középpontu, a koordinátatengelyekkel párhuzamos oldalélű, $2a_i^{(r)}$ ($r = 1, \dots, n$) oldalhosszuságú, n -dimenziós téglatestet határoz meg. Az \underline{a}_i vektor a gyakorlatban az i -edik osztályhoz tartozó szórásvektor függvénye (pl. konstansszorosa) szokott lenni.

Mindezek előrebocsátása után a döntési szabály a következő:

- ha van olyan $i \in \{1, 2, \dots, t\}$ index, amelyre

$$w_i^{(r)} - \alpha_i^{(r)} \leq z^{(r)} \leq w_i^{(r)} + \alpha_i^{(r)} \quad (r = 1, 2, \dots, n)$$

és

$$\vartheta_i = d_i,$$

akkor a $\delta_D(\underline{z}) = d_i$ döntést hozzuk;

- ha nincs ilyen index, akkor a \underline{z} vektort osztályozhatatlannak tekintjük; azaz egy kitüntetett d_0 döntés mellett $\delta_D(\underline{z}) = d_0$ döntést hozunk; végül
- ha több index is megfelelne, és ezek különböző döntéseket jelentenek, akkor valamilyen egyéb szempont alapján döntünk; ilyen lehet például a középponttól vett távolság, vagy egyszerűen a sorrend.

Tehát a dobozmódszer alkalmazása esetén végül is olymódon osztályozunk, hogy ha a vizsgált sajátságvektor valamelyik tanítóhármassal meghatározott téglatestbe esik, akkor az ehhez tartozó osztályba soroljuk (lásd 33. kép).

5.2.5 A POTENCIÁLFGGVENY-MÓDSZER

Legyenek ismét adottak a $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_t$ mintavektorok. Tegyük fel, hogy csak két osztályunk van, és legyen ismert a $K(\underline{z}, \underline{w}_j)$ függvény, amelyet potenciálfggvénynek nevezünk. Defináljuk a

$$g(\underline{z}) = \sum_{j=1}^t q_j K(\underline{z}, \underline{w}_j) \quad (5-37)$$

függvényt: ez lesz a diszkriminánsfüggvény. Az (5-37) kifejezés értelmében a sajátságvektorokat térbeli pontszerű töltéseknek feleltetjük meg úgy, hogy a mintapon-

tokban elhelyezkedő q_j "töltésmennyiség" pozitív, ha a minta az első osztályba tartozik és negatív, ha a másodikba. A $K(\underline{z}, \underline{w}_j)$ függvény a \underline{z} és a \underline{w}_j pont közötti *potenciálkülönbséget* reprezentálja, $g(\underline{z})$ pedig a \underline{z} -beli *eredő potenciált*.

A \underline{z} pontot az első osztályba soroljuk, ha $g(\underline{z}) > 0$, a másodikba, ha $g(\underline{z}) < 0$. A diszkriminánsfüggvényt rekurzív eljárással definiáljuk; tulajdonképpen "megtanuljuk" a q_j együtthatókat. Általában - nem mindenkor - feltezzük, hogy a $K(\underline{z}, \underline{w}_j)$ függvényeknek \underline{w}_j -ben van maximum, és monoton csökkenően 0-hoz tartanak, ha

$$\|\underline{z} - \underline{w}_j\|^2 \rightarrow \infty.$$

A tanulóeljárás során aszerint korrigáljuk a $g(\underline{z})$ diszkriminánsfüggvényt, hogy hogyan osztályozza a \underline{w}_j -t:

$$g'(z) = \begin{cases} g(\underline{z}) + K(\underline{z}, \underline{w}_j), & \text{ha } \delta_p(\underline{w}_j) = d_1 \text{ és } g(\underline{w}_j) \geq 0; \\ g(\underline{z}) - K(\underline{z}, \underline{w}_j), & \text{ha } \delta_p(\underline{w}_j) = d_2 \text{ és } g(\underline{w}_j) \leq 0; \\ g(\underline{z}) & \text{különben; } (j = 1, 2, \dots, t). \end{cases}$$

A módszernek létezik többosztályos általánosítása is, ismertetése megtalálható például [A-6]-ban.

5.2.6 KLASZTEREZÉS

A klaszterezés tanító nélküli osztályozás. Alkalmazására akkor kerül sor, ha az osztályba soroláshoz semmiféle előzetes információnk nincs: az *osztályokat is* a sajátságvektorok alapján kell meghatároznunk.

Egy klaszterbe azokat az adatokat gyűjtjük össze, amelyek valamilyen értelemben erősen *hasonlítanak* egymásra; a különböző klaszterbe kerülők viszont erősen *különböznek* egymástól. Valójában - mint erre az angol elnevezés is utal (cluster = csomó, nyaláb) - a hisztogram sűrűsö-

dési pontjait keressük meg; ezek környékén definiálunk egy-egy klasztert, és gyűjtjük egybe a sajáttságvektorokat. Egy adathalmaznak nyilvánvalóan többféle klaszterezése is lehetséges. Függetlenül attól, hogy mivel mérjük a sajáttságvektorok hasonlóságát; mivel a klaszterek szeparáltságát, tömörségét; befolyásolja az adattér esetleges induló particiója is.

Nehéz eldönteni egy klaszterezésről - és nem is elsődlegesen matematikai, vagy számítástechnikai probléma - hogy megfelel-e egy adott célnak, hogy olyan klasztereket ismert-e fel, melyek használhatók. Éppen ezért a klaszterezés eredményét mindenkor megfelelő kritikával szabad csak elfogadni és felhasználni.

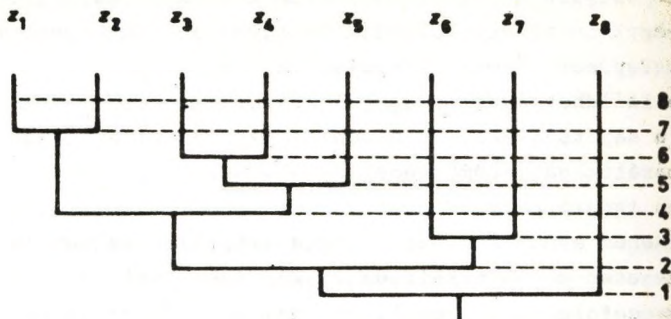
5.2.6.1 Hierarchikus módszerek

A *hierarchikus klaszterező* eljárások alapelve a következő:

Tegyük fel, hogy van N mérési adatunk (sajáttságvektorunk). Induláskor tekintsünk minden adatot önálló klaszternek. Ezek után vonjuk össze a legközelebbi két klasztert, és számítsuk ki a maradék klaszterek távolságát. Ismét vonjuk össze a két legközelebbit stb. Az eljárást addig ismételjük, míg az előírt számú klaszterig nem jutunk. Az eljárást az 5-2. ábra szemlélteti.

A hierarchikus klaszterezés során a *véges, diszjunkt* Z_1 és Z_2 halmazok távolságára az alábbi távolságdefiniciók használatosak (ρ tetszőleges metrika):

$$a) \quad \rho_{\min}(Z_1, Z_2) = \min \left(\rho(\underline{z}, \underline{w}); \begin{array}{l} \underline{z} \in Z_1 \\ \underline{w} \in Z_2 \end{array} \right);$$



5-2. ábra. Hierarchikus klaszterezés elve

$$b) \rho_{\max}(Z_1, Z_2) = \max_{\substack{z \in Z_1 \\ w \in Z_2}} (\rho(z, w));$$

$$c) \rho_{\text{átl}}(Z_1, Z_2) = \frac{1}{c_1 c_2} \sum_{z \in Z_1} \sum_{w \in Z_2} \rho(z, w),$$

ahol $c_1 = |Z_1|$, illetve $c_2 = |Z_2|$ a Z_1 , illetve a Z_2 klaszter elemszáma;

$$d) \rho_{\text{centr}}(Z_1, Z_2) = \rho(\underline{\mu}_1, \underline{\mu}_2),$$

$$\text{ahol } \underline{\mu}_1 = \frac{1}{c_1} \sum_{z \in Z_1} z, \quad \underline{\mu}_2 = \frac{1}{c_2} \sum_{w \in Z_2} w.$$

Az a) esetben *legközelebbi szomszéd*, a b) esetben *minimax* típusú klaszterezésről beszélünk. A hierarchikus klaszterezés tárgyának a klaszterek számával négyzetesen arányos, mivel a klaszterek távolságait tárolni kell; így valójában csak kis elemszám esetén használható. (Részletesebben ismertetés található pl. [5-18]-ban.)

5.2.6.2 Klaszterezés célfüggvény alapján

Az osztályozás jóságát mérhetjük valamilyen *célfüggvény* segítségével. Ilyenkor általában a célfüggvény szélső értéke (minimuma vagy maximuma) szolgáltatja számunkra a keresett legjobb besorolást.

Legyen $Z = \{z_1, z_2, \dots, z_N\}$ a mérések halmaza, és tekintsük egy tetszőleges particionálását:

$$\bigcup_{i=1}^h Z_i = Z,$$

ahol

$$Z_i \cap Z_m = \emptyset, \quad \text{ha } i \neq m; \quad (i, m = 1, 2, \dots, h).$$

Jelölje a továbbiakban

$$\underline{\mu}_i = \frac{1}{c_i} \sum_{z_j \in Z_i} z_j, \quad \text{illetve } c_i = |Z_i| \quad (i = 1, 2, \dots, h) \quad (5-38)$$

az i -edik klaszterbe eső sajátságvektorok *átlagát* (az osztályközpontot), illetve *számukat*. Ezekkel a jelölésekkel a "legkisebb négyzetes hiba" célfüggvényt a következő módon definiáljuk:

$$J_{1knh} = \sum_{i=1}^h \sum_{z_j \in Z_i} \|z_j - \underline{\mu}_i\|^2 \quad (j = 1, 2, \dots, N) \quad (5-39)$$

A J_{1knh} célfüggvény a klaszterközpontok és a klaszterelemek távolságának négyzetösszege, ami az (F3-8)-ban definiált euklidészi norma négyzetének felel meg.

Ha (5-39)-et átírjuk a

$$J_h = \frac{1}{2} \sum_{i=1}^h c_i u_i \quad (5-40)$$

alakba, ahol

$$u_i = \frac{1}{c_i} \sum_{\underline{z}, \underline{w} \in Z_i} \rho(\underline{z}, \underline{w}), \quad (5-41)$$

akkor általánosításra alkalmas formulához jutunk, mivel az u_i függvények tetszőlegesen választhatók (lásd 5.2.6.1 alpont). Tekintsük pl. a $h: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ függvényt, amelyet *hasonlósági függvénynek* nevezünk, ha - egy rögzített $h_0 \in \mathbb{R}^+$ szám esetén - rendelkezik az alábbi 3 tulajdonsággal:

$$\begin{aligned} h(\underline{z}, \underline{w}) &\leq h_0; \\ h(\underline{z}, \underline{z}) &= h_0; \\ h(\underline{z}, \underline{w}) &= h(\underline{w}, \underline{z}). \end{aligned}$$

A hasonlósági függvény felhasználásával legyen

$$u_i = \frac{1}{c_i} \sum_{\underline{z}, \underline{w} \in Z_i} h(\underline{z}, \underline{w}). \quad (5-41/a)$$

Ezt (5-40)-be helyettesítve, a kapott J_h célfüggvénnyel ugyancsak mérhető a klaszterezés "jósága". Ügyeljünk arra, hogy míg az (5-39) függvény minimumát kell keresni, ez esetben az (5-40) maximalizálása jelenti a megoldást.

A J_{lknh} célfüggvényhez az előzőektől eltérő módon is eljuthatunk, szoros kapcsolatban az 5.2.2.3 alpontban definiált empirikus szórásmátrixszal. Legyen

$$\underline{\mu} = \frac{1}{N} \sum_{j=1}^N \underline{z}_j, \quad (j = 1, 2, \dots, N)$$

$$\underline{\Omega}_i = \frac{1}{c_i} \sum_{\underline{z}_j \in Z_i} (\underline{z}_j - \underline{\mu}_i)(\underline{z}_j - \underline{\mu}_i)^T$$

a teljes mérés-halmaz átlaga, illetve az i -edik osztály empirikus szórás-mátrixa, ahol $\underline{\mu}_i$ -t az (5-38) definiálja. Legyen továbbá

$$\underline{\underline{\Omega}}_W = \sum_{i=1}^h c_i \underline{\underline{\Omega}}_i;$$

$$\underline{\underline{\Omega}}_B = \sum_{i=1}^h c_i (\underline{\mu}_i - \underline{\mu})(\underline{\mu}_i - \underline{\mu})^T;$$

$$\underline{\underline{\Omega}}_T = \sum_{z_j \in Z_i} (\underline{z}_j - \underline{\mu})(\underline{z}_j - \underline{\mu})^T.$$

Némi számolással belátható, hogy $\underline{\underline{\Omega}}_T = \underline{\underline{\Omega}}_W + \underline{\underline{\Omega}}_B$. Ezekkel a jelölésekkel (5-39) a

$$\begin{aligned} J_{lknh} &= \sum_{i=1}^h \sum_{z_j \in Z_i} (\underline{z}_j - \underline{\mu}_i)(\underline{z}_j - \underline{\mu}_i)^T = \\ &= \sum_{i=1}^h \text{tr}(\underline{\underline{\Omega}}_i) = \text{tr}(\underline{\underline{\Omega}}_W) \quad (j = 1, 2, \dots, N) \end{aligned} \quad (5-42)$$

alakba írható, vagyis a célfüggvény azonos a $\underline{\underline{\Omega}}_W$ mátrix nyomával amit a $\text{tr}(\cdot)$ szimbólummal jelöltünk. A szórás-mátrix felhasználásával (5-42)-höz hasonlóan definiálhatjuk a *determináns-célfüggvényt* is:

$$J_{\det} = \left| \sum_{i=1}^h \underline{\underline{\Omega}}_i \right|. \quad (5-43)$$

Az (5-42) számos további, különböző szempontokból előnyös célfüggvény előállítására is lehetőséget nyújt. Például a

$$J_{\lambda} = \text{tr}(\underline{\underline{\Omega}}_W^{-1} \underline{\underline{\Omega}}_B) = \sum_{i=1}^n \lambda_j$$

célfüggvény invariáns a lineáris transzformációkkal szemben. (A $\lambda_1, \dots, \lambda_n$ számok az $\Omega_W^{-1} \Omega_B$ mátrix sajátértékei.)

5.2.6.3 Iteratív módszerek

A célfüggvény kiválasztása után az osztályozás tulajdonképpen diszkrét optimalizálási problémává válik; meg kell határozni a Z halmaz egy olyan particionálását, amely a célfüggvény szélső értékét (minimumát vagy maximumát) szolgáltatja. Kicsi adathalmaz esetén ez egyszerűen megoldható az összes lehetséges partició ellenőrzésével; ami nagyobb halmazok esetén már elképzelhetetlen, lévén az n elemű halmaz összes szóba jövő particióinak száma (lásd [A-4] 226. old.):

$$\frac{1}{n!} \sum_{i=1}^n \binom{n}{i} (-1)^{(n-i)} i^n.$$

A probléma pl. iteratív módszerekkel hidalható át. Az alábbiakban ezekre mutatunk három példát.

a) *A négyzetes hiba minimalizálása*

Előzetesen vizsgáljuk meg, hogy az (5-39) célfüggvény értéke hogyan változik meg, ha egy sajátságvektort egyik klaszterből a másikba átsorolunk. Jelölje az m -edik osztály legkisebb négyzetes hibáját

$$J_m = \sum_{\substack{z_j \in Z_m \\ -j}} \left\| \underline{z}_j - \underline{\mu}_m \right\|^2; \quad (m = 1, 2, \dots, h; \quad j = 1, 2, \dots, N).$$

Há a Z_m részhalmazhoz hozzá vesszük a \underline{z}^* pontot, a $\underline{\mu}_m$ átlag, illetve a J_m értéke rendre a következőképpen változik:

$$\underline{\mu}_m^* = \underline{\mu}_m + \frac{\underline{z}^* - \underline{\mu}_m}{c_m + 1}; \quad (5-44/a)$$

$$\begin{aligned} J_m^* &= \sum_{\underline{z}_j \in Z_m} \|\underline{z}_j - \underline{\mu}_m^*\|^2 + \|\underline{z}^* - \underline{\mu}_m^*\|^2 = \\ &= J_m + \frac{c_m}{c_m + 1} \cdot \|\underline{z}^* - \underline{\mu}_m\|^2. \end{aligned} \quad (5-44/b)$$

Hasonlóképpen, ha az i -edik részhalmazból kivesszünk egy elemet, az új átlag, ill. az új függvénykomponens a következő lesz:

$$\underline{\mu}_i^* = \underline{\mu}_i - \frac{\underline{z}^* - \underline{\mu}_i}{c_i - 1}; \quad (5-45/a)$$

$$J_i^* = J_i - \frac{c_i}{c_i - 1} \|\underline{z}^* - \underline{\mu}_i\|^2. \quad (5-45/b)$$

Egy pontnak az i -edik klaszterből az m -edikbe történő átsorolásával a célfüggvény értéke akkor csökken, ha az m -edik függvénykomponens kevésbé nő, mint amennyivel az i -edik csökken. Tehát a feltétel

$$J_m^* - J_m = \frac{c_m}{c_m + 1} \|\underline{z}^* - \underline{\mu}_m\|^2 < \frac{c_i}{c_i - 1} \|\underline{z}^* - \underline{\mu}_i\|^2 = J_i - J_i^*.$$

A legnagyobb csökkenést akkor érjük el, ha

$$\frac{c_m}{c_m + 1} \|\underline{z}^* - \underline{\mu}_m\|^2 = \text{minimum.}$$

A fentiek alapján az algoritmus a következő lesz:

1. Válasszunk egy tetszőleges induló partíciót; számítsuk ki a célfüggvény J_{1knh} értékét és az átlagokat: $\underline{\mu}_1, \underline{\mu}_2, \dots, \underline{\mu}_h$.
2. Válasszunk ki a következő vizsgálandó pontot, és tegyük fel, hogy a Z_i -hez tartozik.
3. Ha $|Z_i| = c_i = 1$, folytatás 5-től; különben számítsuk ki az

$$\epsilon_m = \begin{cases} \frac{c_m}{c_m + 1} \|z_j - \underline{\mu}_m\|^2, & \text{ha } m \neq i; \\ \frac{c_i}{c_i - 1} \|z_j - \underline{\mu}_i\|^2, & \text{ha } m = i \end{cases}$$

változásokat az $i = 1, 2, \dots, h$ esetre.

4. Ha van olyan r index, amelyre $\epsilon_r < \epsilon_m$ ($m = 1, 2, \dots, h$) és $r \neq i$, akkor soroljuk át a z_j vektort Z_r -be és számítsuk ki $\underline{\mu}_r$ és J_r , illetve $\underline{\mu}_i$ és J_i új értékét (5-44), illetve (5-45) alapján. Különben folytatás 5-től.
5. Ha a J_{1knh} célfüggvény nem változott az utolsó s menetben, az eljárás befejeződött; különben folytatás 2-től.

b) *Súlypontkereső eljárás*

1. Legyen $\tau = 0$, és válasszunk ki tetszőlegesen $h_0 >$ számu kezdeti klaszterközéppontot (súlypontot):

$$\underline{\mu}_1, \underline{\mu}_2, \dots, \underline{\mu}_{h_0}$$

2. A τ -adik iterációs lépésben a sajátságvektorokat a következőképpen soroljuk be:

$$\underline{z}_j \in Z_i(\tau), \text{ ha } \|\underline{z}_j - \underline{\mu}_m(\tau)\|^2 \geq \|\underline{z}_j - \underline{\mu}_i(\tau)\|^2;$$

$$(i, m = 1, 2, \dots, h; j = 1, 2, \dots, N).$$

3. Számítsuk ki az új klaszterközéppontokat:

$$\underline{\mu}_m(\tau + 1) = \frac{1}{c_m(\tau)} \sum_{\underline{z}_j \in Z_m(\tau)} \underline{z}_j; \quad (m = 1, 2, \dots, h_\tau);$$

ahol $c_m(\tau) = |Z_m(\tau)|$, a klaszter elemszáma.

4. Ha a $\underline{\mu}_m(\tau + 1) = \underline{\mu}_m(\tau)$ ($m = 1, 2, \dots, h_\tau$), vagyis a klaszterközéppontok nem változtak, akkor az algoritmus véget ért; különben legyen $\tau = \tau + 1$, és folytatás 2-től.

Vegyük észre, hogy ha az új klaszterközéppontokat minden pont besorolása után újraszámítjuk, ez az eljárás lényegében csak a besorolási kritériumban tér el az előbitől.

c) ISODATA-eljárás (Interactive Self-Organizing Data Analysis Techniques)

Először egy általános vázat ismertetünk. Legyen δ egy tetszőleges döntésfüggvény, és L egy olyan leképezés, mely a klaszterekből előállítja az új besoroláshoz szükséges osztályjellemzőket.

1. Válasszunk $h_0 \geq 1$ számú induló klasztert, az $a_1(0), a_2(0), \dots, a_{h_0}(0)$ osztályjellemzőkkel; ($\tau = 0$).

2. A δ döntésfüggvény segítségével soroljuk be az adatokat a megadott h_τ számú osztály valamelyikébe.
3. Az egymáshoz (valamilyen távolságfüggvény szerint) *túl közeli* klasztereket vonjuk össze; a *túl kicsiket* szüntessük meg; a (valamilyen értelemben) *túl lazákat* vágjuk szét. A felszámolt klaszterek elemeit minden alkalommal soroljuk be újra.
4. Származtassuk az új osztályjellemzőket az L leképezés segítségével:

$$a_1(\tau + 1), a_2(\tau + 1), \dots, a_{h_\tau}(\tau + 1).$$

5. Ha az osztályjellemzők az előző ciklus óta nem sokat változtak, vagy az előírt számú ciklus lefutott, az eljárás véget ért; egyébként folytatás 2-től.

Az eljárás során a döntést gyakran a "legközelebbi szomszéd" elv alapján hozzuk és a kiválasztott klaszterjellemző az átlag.

A módszer alapváltozatában [5-2] a 3. lépés nem szerepel, vagyis a klaszterek száma állandó. Igazolható, hogy az eredeti ISODATA-módszer konvergál, és pedig a már vizsgált J_{lknh} célfüggvény minimumhelyéhez.

A 3. lépésben, a klaszterek összevonásánál, leggyakrabban a középpontok távolságát vizsgáljuk és az aránytalanul közeliakat vonjuk össze. *Kicsinek* minősítünk általában egy klasztert, ha elemszáma az átlagos elemszámnál jelentősen kisebb. A klaszterek *lazaságáról* például szórásuk alapján dönthetünk. A kezdeti klaszterek megválasztása *jelentősen befolyásolhatja* a klaszterezés végeredményét; általában a mintatér egyenletes kifestésére szokás törekedni.

A klaszterezés elméletének részletesebb tanulmányozásához ajánljuk pl. az [A-2], [A-6], [A-8], [5-18] műveket.

A 34. képen példát mutatunk az ISODATA típusu klaszterezésre. A feladatot az [1-31] rendszerrel oldottuk meg. Anélkül, hogy belemennénk a klaszterezés eredményének értelmezésébe, ismertetjük a futás néhány paramétereit és részeredményét. A kép egy LANDSAT felvételnek nagyon kicsi - 128×128-as méretű - részlete, mely Tata környékét ábrázolja. A programot úgy paramétereztük, hogy gyakorlatilag tetszőleges számú ciklust végre tudjon hajtani. Maximálisan 16 klasztert engedélyeztünk, és lehetővé tettük az egészen kicsi klaszterek megtartását is; továbbá a klaszterek összevonására és vágására is "hajlamossá" tettük az eljárást. A 34. képen az 1., a 4., a 11. és a 20. ciklus eredménye látható; az utolsóban 15 klaszter adódott eredményként. Belőlük - az egymástól mért távolságuk vizsgálata után - különböző összevonások segítségével - végül 7 klasztert nyertünk. Az 1., 3., 6. és 7. klaszter különböző mezőgazdasági kultúrákkal azonosítható, közülük például az első klaszter kukorica/silókukorica. A 2. klaszter vízfelület, a 4. erdő; míg az 5. klaszter a lakott településeken, kisparcellás területeken fordul elő. Az 5-1. táblázatban megadjuk az első és a második klaszter 1., 4., 11. és 20. ciklusbeli centrumát, szórását és elemszámát. Az 5-1. táblázatot és az ábrát összevetve világos, hogy az eljárás az első néhány ciklusban még jelentős változásokat eredményezett, a későbbiekben az egyes osztályok lényegesen már nem módosultak.

Az 5-2. táblázatban az 1. és a 2. klaszter centrumának elmozdulását mutatjuk be. A távolságokat a ρ_E (euklidészi) távolságfüggvény alapján számítottuk.

5-1. táblázat. A 34. kép 1. és 2. klasztere jellemzőinek változása

		1. klaszter				2. klaszter			
		1. sáv	2. sáv	3. sáv	4. sáv	1. sáv	2. sáv	3. sáv	4. sáv
1. ciklus	Átlag	53,87	67,14	114,81	112,38	42,97	45,49	85,62	85,20
	Szórás	4,66	9,98	7,29	7,13	2,60	5,22	16,61	22,66
	Elemsz.	4427	-	-	-	4666	-	-	-
4. ciklus	Átlag	50,51	60,99	111,70	111,27	42,94	44,61	56,85	44,39
	Szórás	8,94	13,33	6,44	5,33	2,95	5,87	15,08	19,40
	Elemsz.	5513	-	-	-	919	-	-	-
11. ciklus	Átlag	50,11	59,22	110,87	110,78	42,25	42,81	48,59	33,52
	Szórás	5,93	12,30	8,12	6,30	1,67	2,74	9,09	11,10
	Elemsz.	5528	-	-	-	649	-	-	-
20. ciklus	Átlag	49,78	58,57	110,35	110,46	42,29	42,80	48,55	33,48
	Szórás	6,34	13,06	8,63	6,55	1,67	2,74	9,06	11,06
	Elemsz.	5574	-	-	-	648	-	-	-

5-2. táblázat. Klaszterközéppont elmozdulása

Ciklus	1. klaszter		2. klaszter	
	Centrum elmozdulása	Elemzám-változás	Centrum elmozdulása	Elemzám-változás
0 → 1	18,05	-	22,86	-
1 → 4	7,73	+ 1086	48,94	- 3747
4 → 11	2,05	+ 15	13,73	- 270
11 → 20	0,91	- 254	0,07	- 1

A kiindulási értékek a következők voltak:

$$[48,06 \quad 67,32 \quad 123,12 \quad 26,75]^T,$$

illetve

$$[42,20 \quad 67,85 \quad 85,92 \quad 8,90]^T.$$

5.2.7 A KÖRNYEZETI INFORMÁCIÓK FIGYELEMBEVÉTELE

Az eddig ismertetett statisztikus (döntéseméleti) módszerek nem, vagy csak nagyon kevésbé használják ki, hogy a mintatér elemei képpontok; a képi információval, a "képszerűséggel" nem foglalkoznak. A "képszerűség" tulajdonképpen abban rejlik, hogy az egyes képpontok környezetükkel együtt ábrázolnak egy tárgyat, hogy - a döntéseméleti modellek alapfeltevésével ellentétben - a környezetüktől egyáltalán nem függetlenek [5-12].

A képi jelleg figyelembevételének egy módja lehet, hogy a döntés során a pont környezetéből származó információkat is felhasználjuk. Az ilyen típusú módszerek közül emelünk ki az alábbiakban néhányat anélkül, hogy részletesebb ismertetésre vállalkoznánk.

Előfeldolgozás. Az előfeldolgozó módszerek közös alap gondolata, hogy valamilyen módon belefoglaljuk a sajtóságvektorba a környezeti információt is, miáltal az osztályozás során automatikusan figyelembe vesszük. A bevétel sokféleképpen történhet; szokásos és egyszerű eljárás a környezet pontjaiból származtatott valamilyen statisztikai jellemző, például az átlag választása. Ide sorolható az 5.2.1.1 alpontban ismertetett sajtóságdefiníció (-származtató) eljárások némelyike is.

Előszegmentálás. Az osztályozandó területet valamilyen módon előzetesen szegmentáljuk, majd a spektrálisan homogénnek tekinthető és feltehetően egy tárgyat ábrázoló objektumot (minden pontját) egyszerre soroljuk be valamelyik osztályba.

Utőfeldolgozó módszerek. Végrehajtjuk az osztályozást a szokásos módon, majd - már csak az eredményképpel foglalkozva - a környezeti információk alapján esetleg korrigáljuk a döntést. Legegyszerűbb példa erre a többségi elvű szűrés (lásd pl. 22. kép). Egy pont alkalmasan választott környezetében összeszámoljuk a különböző osztályokba sorolt képpontokat, s ha ezek között valamelyik osztály relatív gyakorisága meghalad egy előre adott értéket, a vizsgált pontot átsoroljuk ehhez az osztályhoz. A módszer finomítható, például az egyes osztályokhoz különböző elfogadási arányt választhatunk.

A témáról további részletek található pl. [A-7]-ben.

5.3 SZINTAKTIKUS ALAKFELISMERÉS

A statisztikus alakfelismerési módszerek az alakfelismerési problémák egy részében nehézkesen alkalmazhatók, nem nyújtanak a feladat megoldásához eszközt. Jelentős nehézségeket okoz például az összetett alakzatok kezelése; ezek nehezen írhatók le bármilyen sajáttságvektorral.

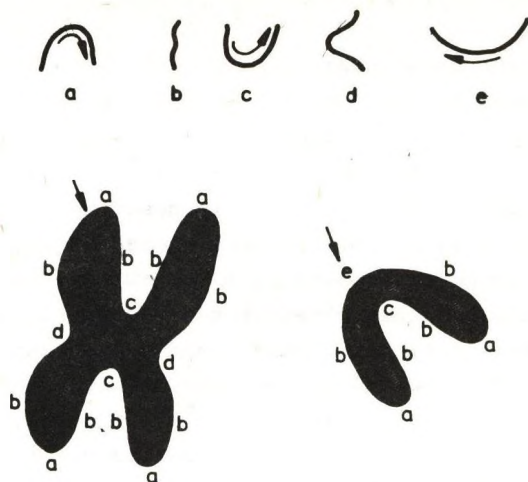
A *szintaktikus alakfelismerési* módszerek más oldalról közelítik meg a problémát; alapelvük a következő: *Bontosuk fel* a vizsgált objektumot olyan egységekre, amelyek egyszerűbb strukturájuak, majd ezeket ismét egyszerűbbekre, egészen addig, amíg olyan elemi objektumokhoz nem jutunk, melyeket már nem célszerű tovább bontanunk, s melyeket a képen biztosan fel tudunk ismerni. Ezeket az elemi objektumokat *képelemeknek* nevezzük. Együtt állapítsuk meg a képelemek közötti kapcsolatokat és a segítségükkel írjuk le az objektumokat.

Az alakfelismerési eljárás során egy osztályt a hozzá tartozó objektumok *leírásának halmaza* definiál [A-5]. Az osztályozás ezek után abból áll, hogy elkészítjük az aktuális objektum leírását, és megvizsgáljuk, hogy tarthat-e valamelyik osztályhoz.

Az objektumok leírására példaként gyakran említik a kromoszómák ilyen jellegű leírását, amelyet [5-15] alapján az 5-3. ábrán mutatunk. A két különböző típusu kromoszóma az "abc**bab**db**bab**cb**ab**db", illetve az "eb**ab**cb**ab**" jel-sorozattal írható le, a nyíllal jelölt szakasztól kiindulva.

A szintaktikus alakfelismerés eszközrendszere a formális nyelvtanok elméletére épül [A-5]. A $G = \langle T, N, S, P \rangle$ négyest akkor tekintjük formális nyelvtannak, ha

- T és N véges, diszjunkt halmazok (esetünkben T a képelemek, N a szimbólumok halmaza);
- $S \in N$ a kezdőjel;
- P az átírási szabályok véges halmaza.



5-3. ábra. Kromoszómák szintaktikus leírása

Egy átírási szabály TUN -beli elemekből összetett, véges hosszúságú jelsorozatokból áll, és úgy értelmezendő, hogy valahányszor az első jelsorozat előfordul, helyettesítendő a másodikkal.

Egy T -beli jelsorozatról akkor mondjuk, hogy levezethető a G nyelvtan segítségével, ha létezik olyan szabálysorozat P -ben, amellyel - S -ből kiindulva - egymás utáni helyettesítésekkel eljutunk hozzá. A G -vel levezethető T -beli jelsorozatok alkotják a G nyelvtan által generált $L(G)$ nyelvet.

A szintaktikus alakfelismerési eljárás ezek alapján a következő lépésekből áll:

- a képelemek meghatározása;
- az osztályokat definiáló jelsorozatok megállapítása;

- egy-egy nyelvtan megadása minden osztályhoz, amelyik a lehető legpontosabban definiálja az osztályhoz tartozó jelsorozatokat;
- az osztályozandó objektum leírásának meghatározása;
- elemzés, hogy melyik nyelvtan generálja az objektumot leíró jelsorozatot (ha egyáltalán van ilyen).

Elvileg a feladatok világosak és egyszerűek, a megvalósításuk azonban számos nehézségbe ütközik. (Igy pl. a képelemkészlet kialakítása alapos előkészítő munkát igényel; nem kevésbé komoly apparátusra van szükség a döntési eljárás megvalósításához.)

A nehézségek általános megoldása egyelőre nem ismeretes, mindazonáltal a szintaktikus alakfelismerésnek máris számos érdekes és sikeres alkalmazása van. A felmerülő problémák és részletek iránt érdeklődő olvasóknak ajánljuk az [5-13], [5-15] és az [5-17] művet.

5.4 TEXTURAELEMZÉS

A texturaelemzési módszerek áttekintését megnehezíti az a körülmény, hogy a texturának nincs elfogadott, precíz definíciója. Ebből következik, hogy ugyanazt a texturát különbözőképpen leírva más-más eredményhez juthatunk.

A továbbiakban az 1.1.2.2 alpontban adott közelítő definícióból indulunk ki. Ennek értelmében *homogén*nak nevezük az olyan texturákat, amelyekben bizonyos *statisztikus ismétlődés* található (lásd alább), amit [5-5] alapján "önhasonlóságnak" is nevezhetünk. Ezen azt értjük, hogy a kép összes "elég nagy" (lásd alább) részképe texturálisan hasonló, függetlenül a helyzetüktől, az alakjuktól és a méretüktől.

Sajnos, ez ideig a hasonlóság fogalmát sem sikerült

egyértelműen meghatározni. Egyes rendszerekben a szubjektív emberi döntést fogadják el. Célszerűbb azonban megállapítani a texturális sajátságok objektív mérésén alapuló kritériumokat; ezek a fentiek alapján nyilvánvalóan függeni fognak a leirási mód megválasztásától.

A texturát *inhomogén*nek mondjuk, ha találhatók olyan "elég nagy" képrészek, amelyek az elfogadott kritériumok szerint nem hasonlók. Első közelítésben ezt úgy képzelhetjük el, hogy valamilyen homogén texturába hibás elemek keveredtek. Ez lényegesen megnehezíti az elemzést, és a kétféle textura egységes tárgyalása egyelőre távolról sincs megoldva. (Egy kísérleti módszer leírása [5-4]-ben található.)

Mint említettük (lásd 5.1 pont), a texturaelemzés folyamata is két fázisra bontható, ezekről lesz szó a következő két alpontban.

5.4.1 TEXTURÁLIS SAJÁTSÁGVEKTOROK ELŐÁLLÍTÁSA

Texturális tulajdonságokat csak olyan képrészekhez lehet hozzárendelni, amelyek "elég nagyok" abban az értelemben, hogy meg lehessen határozni a szükséges (a sajátságvektorok összetevőit adó) statisztikus jellemzőiket.

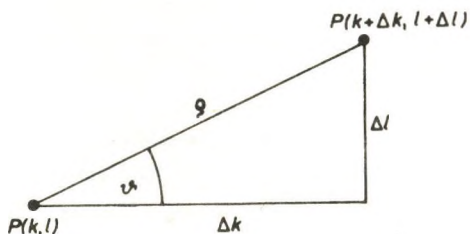
Ezeket származtathatjuk egyszerűen a *hisztogram*ból; ilyen pl. az átlag, a szórás, a medián, a maximum/minimum stb. Képezhetünk sajátságvektort a *futási hosszak*ból is (= az egymás utáni azonos világosságkódú képpontok száma valamilyen irányban); ugyanis a világosságkódok geometriai eloszlása összefüggésben van a texturával.

A leggyakrabban használt "sajátságforrás" az *együtt-előfordulási mátrix* (co-occurrence matrix), amelynek $\beta_{ij}(\rho, \theta)$ eleme azt adja meg, hogy egy q_i világosságkódú

képponttól (a $+X$ tengelyhez viszonyított) ϑ szögben ρ távolságra milyen valószínűséggel (pontosabban: milyen relatív gyakorisággal) találunk q_j világosságkódokat. Az 5-4. ábra alapján ezt a

$$\beta_{i,j}(\rho, \vartheta) = \frac{1}{N} \text{card}(\{P(k, l), P(k + \Delta k, l + \Delta l)\} | q(k, l) = q_j \wedge q(k + \Delta k, l + \Delta l) = q_j)$$

képletből számíthatjuk ki, ahol N a képpontok számát, $\text{card}(A)$ pedig az A halmaz elemeinek számát jelenti. Fontos megjegyezni, hogy a diszkrét képsíkon ρ és ϑ is csak diszkrét értékű lehet, és nem függetlenek egymástól. (Leggyakrabban a $\vartheta = n \cdot 45^\circ$ értékekkel számolnak, de ez általában nem eredményez kellő pontosságot.)



5-4. ábra. Együtt-előfordulási mátrix eltolásvektorának szemléltetése

Mivel b bites világosságkódok esetén az együtt-előfordulási mátrix mérete egy (ρ, ϑ) pár esetén $2^b \times 2^b$, [4-20] alapján többféle tömörített sajátást szokás leszámaztatni. Ezek közül az *uniformitást* (U), az *entrópiát* (E), az *n -edrendű nyomatékokat* (M_n), a *korrelációs együttthatót* (C) és az *autokorrelációs függvényt* ($R(s, t)$) említjük meg, amelyeket rendre az alábbi összefüggésekkel definiálunk:

$$U = \sum_{q_i=0}^J \sum_{q_j=0}^J \beta_{ij}^2(\rho, \vartheta); \quad (5-46)$$

$$E = - \sum_{q_i=0}^J \sum_{q_j=0}^J \beta_{ij}(\rho, \vartheta) \log_2 \beta_{ij}(\rho, \vartheta); \quad (5-47)$$

$$M_n = \sum_{q_i=0}^J \sum_{q_j=0}^J \beta_{ij}(\rho, \vartheta) |q_i - q_j|^n, \quad (5-48)$$

(az elsőrendű nyomatékot *kontrasztnak*, a másodrendűt *inerciának* is nevezik);

$$C = \sum_{q_i=0}^J \sum_{q_j=0}^J \frac{q_i q_j \beta_{ij}(\rho, \vartheta) - \mu_i \mu_j}{\sigma_i \sigma_j}, \quad (5-49)$$

(ahol μ_i és σ_i , illetve μ_j és σ_j a mátrix i -edik sorának, illetve j -edik oszlopának átlaga és szórása);

$$R(s, t) = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} q(k, l) q(k-s, l-t), \quad (5-50)$$

($s = -K, -K+1, \dots, K$; $t = -L, -L+1, \dots, L$; és $q(m, n) = 0$, ha $m < 0$, illetve $m \geq K$ vagy $n < 0$, illetve $n \geq L$).

Az együtt-előfordulási mátrix egyszerűsített változata a *különbségi hisztogram* (gray-level difference histogram) [4-19]. A hisztogram $\delta_m(\rho, \vartheta)$ eleme annak relatív gyakoriságát adja meg, hogy a q_i világossággódu képpontoktól ϑ szögben, ρ távolságra éppen $q_j = q_i + m$ világossággóduakat találunk. Az előbbi jelölésekkel ezt a

$$\delta_m(\rho, \vartheta) = \frac{1}{N} \text{card}\{(P(k, l), P(k + \Delta k, l + \Delta l)) \mid |q(k, l) - q(k + \Delta k, l + \Delta l)| = m\}$$

képlet fejezi ki ($0 \leq m \leq J = 2^b - 1$). A hisztogramból az (5-46)-(5-48) összefüggések értelemszerű alkalmazásával ugyancsak leszámaztathatunk tömörített sajátságokat.

5.4.2 HOMOGEN TEXTURÁK VIZSGÁLATA

A továbbiakban abból indulunk ki, hogy a vizsgálandó képet fel lehet bontani legalább két, az 5.4.1 alpont értelmében "elég nagy" részképre. Feltesszük továbbá, hogy a megtalált sajátságvektorok kielégítően leírják a kép lényeges texturális tulajdonságait, valamint hogy rendelkezésünkre állnak az ezeknek megfelelő kritériumok, a képrészek hasonlóságának eldöntésére. (Ez utóbbiak vagy eleve (a priori) adottak, vagy meghatározhatók a képanyag előzetes, interaktív elemzése során.) Ezek után a kép texturáját akkor mondjuk *homogén*nek, ha *bármely két eleme hasonló*. Ezzel kapcsolatban két fontos megjegyzést teszünk:

1. A hasonlóság szimmetrikus és reflexív, de nem tranzitív reláció [5-5].
2. Egy kép homogénnek minősített texturája más képleírók és a hozzájuk tartozó hasonlósági kritériumok alapján ihomogénnek minősülhet!

Elterjedt felfogás szerint a fentiekben értelmezett részképeket azonosnak tekintik a texturaelemekkel, amelyek két lépésben különíthetők el:

1. Kijelölik az értékes és a háttérpontokat (pl. két szintre vágással).

2. Megkeresik (csoportosítják) az egy alakzathoz tartozó képpontokat, valamilyen ismert szabályok szerint. (Ez utóbbi lépést hierarchikus texturák esetében minden hierarchiaszinten végre kell hajtani.)

A texturaelemeket azonban nem mindig lehet egyértelműen (esetleg egyáltalán nem is lehet) elkülöníteni, mivel a fogalmuk erősen kapcsolódik az emberi texturaérzékeléshez (lásd 1.1.3.4 alpont); ez visszavezet a fenti 2. megjegyzésben említett problémához. Jelenleg tehát azt kell mondanunk, hogy "természetes" texturaleírás nem létezik, csak egy feladat megoldására többé vagy kevésbé alkalmas leírások konstruálhatók.

5.4.2.1 A texturák alaptulajdonságai

A texturák kvalitatív minősítésére sok, pontosan nem definiált fogalom használatos. Eredeti célkitűzésünknek megfelelően csak a texturák síkbeli képét, illetve "sima" felületek texturáját vizsgáljuk; vagyis a harmadik dimenzióban megjelenő "érdességet" nem vesszük figyelembe. Ilyen texturák *kvantitatív* jellemzésére 3 tulajdonságot szokás értékelni:

1. A *finomság* (vagy durvaság) a texturaelemek átlagos méretével írható le. Beszélünk még "sűrű", illetve "laza" texturáról is; a sűrűséget a texturaelemek átlagos méretének és átlagos távolságának hányadosaként értelmezzük.
2. A *szabdlyosság* (vagy véletlenszerűség) a texturaelemek strukturáját leíró bonyolult és sokrétű fogalom, amelynek általános mértékét még nem sikerült definiálni. Akkor használjuk, ha a képen ismétlődő, determinisztikus minta található. Azon-

ban nem csak a szigorú periodicitásra kell gondolni (amikor is az egész kép leszámaztatható egy részkép ismétlésével), hanem a *globális* vagy a *lokális geometriában*, illetve a világosságkódok síkbeli eloszlásában mutatkozó szabályosságokra is [5-1]. (A globális geometria a texturaelemek síkbeli elhelyezkedését, a lokális az alakjukat és az irányítottóságukat írja le.) Mindhárom komponens valószínűségi változóként kezelhető, amelyek azonban nem függetlenek egymástól.

3. Az *anizotrópia* az újabb kutatásokban került előtérbe [5-3]. Tulajdonképpen ez a szabályosságnál alacsonyabb szintű strukturaleírás és arról ad számot, hogy léteznek-e a texturaelemek elrendezésében kitüntetett irányok. (Minden, globális geometriájában szabályos minta anizotróp! Az izotróp texturák szerkezete irányfüggetlen.) Az anizotrópia fontossága abban van, hogy szoros kapcsolatban áll a texturák szimmetriájával, ami kulcsfontosságú a bonyolult objektumok matematikai leírásában; de jelenléte megkönnyíti az emberi észlelést is.

Megjegyezzük, hogy a digitális képekre alkalmazott irányfüggő leírók bizonyos szórást mutatnak, amit a kiértékeléskor figyelembe kell venni.

5.4.2.2 Elemzési módszerek

A texturaelemzési módszerek [5-9] alapján 3 csoportba sorolhatjuk:

1. a *képpont alapúak* (pixel-based) a világosságkódok,
2. a *lokális sajátosság alapúak* (local feature-based) a világosságkódok eloszlásában található lokális átmeneti tartományok (pl. élek),

3. a *tartomány alapúak* (region-based) pedig az előzetesen meghatározott texturaelemek (elemi tartományok)

eloszlási statisztikáit, illetve geometriai/topológiai jellemzőit használják a textúra leírására. Az előbbiek a texturaelemzés statisztikus, az utóbbiak a szintaktikus (vagy strukturális) ágához tartoznak (lásd 5-1. ábra).

Az előzőekből nyilvánvaló, hogy a tartomány alapú módszereknek szembe kell nézniük a texturaelemek elkülönítésének nehézségeivel. Emiatt inkább az első két csoportba tartozó módszerek terjedtek el. Az alábbiakban [5-5] alapján ezeket tekintjük át röviden.

A *képpont alapú módszerek* a finomabb, illetve sűrű texturák elemzésére a legalkalmasabbak. A különböző eljárások összehasonlítása során a *kis eltolások* ($\rho = 1, 2$) melletti *együtt-előfordulási mátrixokból* származtatott jellemzőkkel lehetett a legjobb felismerőképességet elérni. Durvább, de sűrű, mozaikszerű texturák esetében az (5-48) szerint számítható $M_2(\rho, \vartheta)$ inercia akkor a legkisebb, ha ρ kb. megegyezik a texturaelemek (előzetesen meg nem határozott!) átlagos méretével. Laza texturáknál a momentumok szélső értékei inkább a meredekségfüggvény (lásd 2-12. ábra) maximumhelyeinek távolságával függnek össze. Kimutatták, hogy szabályos, periodikus texturák esetében az inercia ρ szerint periodikus, és periódusa megegyezik a texturáéval. Így problémát okozhat, hogy az inercia akkor is kis értékű, ha ρ a periódus egész számu többszöröse. A periodicitást a χ^2 próbával ellenőrizve a szabályosság, az együtt-előfordulási mátrix ϑ szögtől való függését vizsgálva pedig az anizotrópia objektív mérésére nyílik lehetőség.

A szabályosságot és az anizotrópiát a *különbségi hisztogramból* leszámaztatott sajátosságok alapján is lehet mérni, így egyszerűbb eljárásokhoz jutunk.

Korábban gyakran használták az (5-50)-nel definiált *autokorrelációs függvényt*, mivel Fourier-transzformáltja megegyezik a kép Fourier-transzformáltja spektrális sűrűségével (= a Fourier-együtthatók négyzetösszegével) [4-20]. Ebben a durvább mintáknál az alacsonyfrekvenciás tagok dominálnak; egyébként a spektrális sűrűség érzékeny az anizotrópiára is, és periodikus texturáknál éles lokális maximumhelyei vannak. Mindezen előnyei mellett a módszer hátránya a túlzott zavarérzékenység, a túl meredek átmenetek zavaró hatása és a nagy számításigény.

A *lokális sajátság alapú* módszereket a nagyobb texturaelemekből felépülő, durvább texturák elemzésére célszerű használni, mivel a képpont alapúak ilyenkor zavaró módon tükrözik az elemek belső szerkezetét is. Leggyakrabban a texturaélek (lásd 4.4.1.4 alpont) eloszlását vizsgálják.

Ezekben a módszerekben kulcsszerepet játszik az *általánosított együtt-előfordulási mátrix* (generalized co-occurrence matrix) [5-8], amelynek elemei a (világosságkódok helyett) bizonyos tulajdonságú lokális sajátságok (pl. a képpontok helyett élek) között fennálló síkbeli viszonyokat írják le. Ez a mátrix jól felhasználható az anizotrópia és a periodicitás mérésére.

A texturaelemzés teljesebb áttekintése és néhány új elemzési módszer leírása [5-5]-ben található, amelynek bőséges irodalomjegyzéke segítséget nyújt a részletek iránt érdeklődőknek is.

FÜGGELÉK

A Függelékben a könyvünkben leírtak megértéséhez szükséges matematikai ismereteket foglaltuk össze röviden, a teljességre való törekvés nélkül. Ez azt jelenti, hogy - bár ügyeltünk a matematikai precizitásra - a problémákat nem vizsgáljuk teljes általánosságban, hanem csak a képfeldolgozás gyakorlatában szükséges mélységben, és csak a számunkra érdekes speciális eseteket tárgyaljuk. Ahol tulságosan bonyolult lenne, nem foglalkozunk precíz definíciókkal vagy bizonyításokkal sem.

A könyvünkben feltételezzük az alapvető halmazelméleti és vektoralgebrai fogalmak ismeretét. Kiemelünk azonban néhány jelölést, amelyeket gyakran használunk:

$\mathbb{N}; \mathbb{R}$

az egész, illetve a valós számok halmaza (\mathbb{R} -t testnek is nevezik, mivel az elemeivel végzett négy aritmetikai alapművelet nem vezet ki a halmazból);

$m, n \in \mathbb{N}; \lambda \in \mathbb{R}$

m és n egész számok, λ valós szám;

$\mathbb{R}^2; \mathbb{R}^n$

a folytonos (euklidészi) sík, illetve n -dimenziós tér;

$A = \{a_1, a_2, \dots, a_n\}$

az A halmaz az a_1, a_2, \dots, a_n elemekből áll;

$A \subset B$	A halmaz részhalmaza a B halmaznak;
$A \times B$	A és B halmaz Descartes-szorzata: az összes lehetséges (a, b) számpár- ból álló halmaz;
$\forall a \in A$	minden olyan a , amelyik eleme az A halmaznak;
$L: A \mapsto B$	az L leképezés az A halmazhoz a B -t rendeli;
$L: a \mapsto b$	az L leképezés az a halmazelemhez a b halmazelemet rendeli;
\wedge	"logikai ÉS": az összekapcsolt fel- tételek vagy események egyidejűleg teljesülnek;
\vee	"logikai VAGY": az összekapcsolt fel- tételek közül egyidejűleg legalább az egyik teljesül;
	"feltéve, hogy"; a feltétel(ek) a vonal jobb oldalán áll(nak);
\Leftrightarrow	"akkor és csak akkor": szükséges és elegendő feltétel vagy egymással ek- vivalens állítások.

A Függelék 4 részből áll, ezekben rendre a mátrix-
algebra, a valószínűségszámítás, a diszkrét geometria,
illetve a függvénykalkulus egyes témaköreit tekintjük át.
A részletek iránt érdeklődő olvasóknak ajánljuk az igen
jó magyar nyelvű szakirodalmat (pl. [F-2], [F-3], [F-8],
[F-9], [F-11]).

F1. MÁTRIXALGEBRA

A képfeldolgozásban a mátrixok hasznos segédeszközt jelentenek, mert segítségükkel tömören és elegánsan lehet leírni a digitális képekkel végzendő műveleteket.

F1.1 A MÁTRIX FOGALMA ÉS FŐBB TIPUSAI

Legyen T egy adott test (pl. a valós számtest); az

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (a_{ij} \in T; \quad i = 1, \dots, m; \\ j = 1, \dots, n)$$

táblázatot a T test feletti *mátrix*nak nevezzük: az a_{ij} adatok a *mátrix* elemei. A sorokat felülről lefelé, az oszlopokat balról jobbra számozzuk; az i -edik sor j -edik elemét a_{ij} jelöli. Egy mátrixot akkor tekintünk adottnak, ha egyrészt ismert sorainak és oszlopainak száma, másrészt minden eleme adott. Jelölésére az

$$\underline{A} = [a_{ij}]_{m \times n}$$

szimbólumot fogjuk használni, és - ha nem vezet félre-

értésre - az $m \times n$ indexet elhagyjuk. Az $\underline{A} = [a_{i,j}]$ és a $\underline{B} = [b_{i,j}]$ mátrix akkor és csak akkor egyenlő, ha egyrészt azonos típusuak (vagyis sor- és oszlopszámuk megegyezik), másrészt minden elemükre $a_{i,j} = b_{i,j}$.

A mátrixok felépítése és szimmetriatulajdonságai alapján a következő fontosabb típusokat különböztetjük meg (az elmondandókat, a könnyebb megértés kedvéért, az F1-1. ábrán adott példákkal illusztráljuk):

- Egy mátrix *transzponáltját* úgy nyerjük, hogy felcseréljük a sorait és az oszlopait; jelölése: \underline{A}^T (pl. $\underline{E}^T, \underline{G}^T, \underline{H}^T, \underline{L}^T, \underline{Z}^T$).
- Ha egy mátrixban minden elem helyére a konjugáltját írjuk, a *konjugált mátrixhoz* jutunk (pl. $\underline{Z}^*, \underline{L}^*$).
- A *négyzetes* vagy kvadratikus mátrix sorainak és oszlopainak száma egyenlő. Az n sorból és oszlopból álló négyzetes mátrixot n -edrendűnek, a főátlóban levő $a_{i,i}$ elemeit pedig *diagonális elemeknek* nevezük; összegük a mátrix *nyoma*. (Pl. az \underline{F} harmadrendű mátrix diagonális elemei: 1,5,6).
- *Diagonálmátrix* az olyan négyzetes mátrix, amelynek csak a főátlóban vannak 0-tól különböző elemei, vagyis minden $a_{i,j}$ ($i \neq j$) elem zérus. Ha a főátlóban minden elem 1 - vagyis $a_{i,i} = 1$ ($i = 1, 2, \dots, n$) - a diagonálmátrixot *egységmátrixnak* nevezzük (pl.: \underline{E}).
- *Szimmetrikusnak*, illetve *hermitikusnak* nevezzük az olyan négyzetes mátrixot, amely egyenlő a transzponáltjával, illetve a konjugált transzponáltjával. (Pl. $h_{i,j} = h_{j,i}$, illetve $z_{i,j} = z_{j,i}^*$; vagyis $\underline{H} = \underline{H}^T$, illetve $\underline{Z} = \underline{Z}^{*T}$).
- *Ortogonálisnak*, illetve *unitérnek* nevezzük az olyan négyzetes mátrixot, amelynek transzponáltjával, illetve a konjugált transzponáltjával képzett szorzata egységmátrix. (Pl. $\underline{G} \underline{G}^T = \underline{E}$; $\underline{H} \underline{H}^T = \underline{E}$; $\underline{Z} \underline{Z}^{*T} = \underline{E}$.)

- Az oszlopmátrix vagy oszlopvektor egyetlen oszlopból $\underline{P} = p$, a sormátrix vagy sorvektor pedig egyetlen sorból $\underline{S} = s$ áll. Nyilvánvaló, hogy egy oszlopmátrix transzponáltja sormátrix és fordítva.

$$\underline{F} = \begin{bmatrix} 1 & 7 & 8 \\ 3 & 5 & 9 \\ 2 & 4 & 6 \end{bmatrix}; \quad \underline{F}^T = \begin{bmatrix} 1 & 3 & 2 \\ 7 & 5 & 4 \\ 8 & 9 & 6 \end{bmatrix}; \quad \underline{F}^{-1} = \begin{bmatrix} -0,6 & -1 & 2,3 \\ 0 & -1 & 1,5 \\ 0,2 & 1 & -1,6 \end{bmatrix};$$

$$\underline{G} = \begin{bmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{bmatrix}; \quad \underline{G}^T = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix};$$

$$\underline{H} = \begin{bmatrix} 0,6 & 0,8 \\ 0,8 & -0,6 \end{bmatrix} = \underline{H}^T; \quad \underline{L} = \begin{bmatrix} 3 & -2i \\ 1 & 4i \end{bmatrix}; \quad \underline{L}^T = \begin{bmatrix} 3 & 1 \\ -2i & 4i \end{bmatrix};$$

$$\underline{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \underline{L}^* = \begin{bmatrix} 3 & 2i \\ 1 & -4i \end{bmatrix}; \quad \underline{L}^{*T} = \begin{bmatrix} 3 & 1 \\ 2i & -4i \end{bmatrix};$$

$$\underline{Z} = \begin{bmatrix} \sqrt{5}/3 & 2/3i \\ -2/3i & -\sqrt{5}/3 \end{bmatrix} = \underline{Z}^{*T}; \quad \underline{Z}^* = \begin{bmatrix} \sqrt{5}/3 & -2/3i \\ 2/3i & -\sqrt{5}/3 \end{bmatrix} = \underline{Z}^T;$$

$$\underline{P} = p = \begin{bmatrix} 2 \\ 9 \end{bmatrix}; \quad \underline{S} = s = [7 \ 4].$$

F1-1. ábra. Példamátrixok ($i = \sqrt{-1}$)

F1.2 MŰVELETEK MÁTRIXOKKAL

F1.2.1 ÖSSZEADÁS, KIVONÁS ÉS SZORZÁS

Az összeadás és a kivonás csak azonos típusu mátrixokra értelmezhető. A

$$\underline{\underline{C}} = [c_{ij}]_{m \times n}$$

mátrixot az $\underline{A} = [a_{ij}]_{m \times n}$ és a $\underline{B} = [b_{ij}]_{m \times n}$ ($a_{ij}, b_{ij} \in T$)

mátrix összegének (különbségének) nevezzük, ha

$$c_{ij} = a_{ij} \pm b_{ij} \quad (i = 1, \dots, m; j = 1, \dots, n).$$

Az így értelmezett mátrixösszeadás kommutatív és asszociatív.

A $\lambda \in T$ skalár és az \underline{A} mátrix szorzatán a

$$\lambda \underline{A} = [\lambda a_{ij}]_{m \times n} \quad (a_{ij} \in T; i = 1, \dots, m; j = 1, \dots, n.)$$

mátrixot értjük. A skalárral való szorzás kommutatív, asszociatív és disztributív.

A $\underline{C} = [c_{ij}]_{m \times p}$ mátrixot az $\underline{A} = [a_{is}]_{m \times n}$ és a

$\underline{B} = [b_{sj}]_{n \times p}$ ($a_{is}, b_{sj} \in T$) mátrix szorzatának nevezzük

- és az $\underline{A} \underline{B}$ szimbólummal jelöljük - ha

$$c_{ij} = \sum_{s=1}^n a_{is} b_{sj}. \quad (\text{F1-1})$$

Hangsúlyozni kell a következőket:

- az $\underline{A} \underline{B}$ szorzat csak akkor van értelmezve, ha \underline{A} -nak ugyanannyi oszlopa van, mint ahány sora van \underline{B} -nek (vagyis \underline{A} és \underline{B} a megadott sorrendben *konformdábilis*);
- mátrix mátrixszal való szorzása általában nem kommutatív;
- sorvektor és oszlopvektor szorzata ebben a sorrendben *szám* (= *skalárszorzat*); fordított sorrendben *mátrix* (= *diadikus szorzat*).

F1.2.2 OSZTÁS NÉGYZETES MÁTRIXSZAL

A számok körében az osztást az osztó reciprokával - inverzével - való szorzásként is értelmezhetjük. Ahhoz, hogy ezt a négyzetes mátrixok körében is megtehessük, értelmezzük az inverz mátrix fogalmát. Az \underline{A}^{-1} mátrixot az \underline{A} négyzetes *mátrix inverzének* nevezzük, ha

$$\underline{A} \underline{A}^{-1} = \underline{A}^{-1} \underline{A} = \underline{E}.$$

Ismeretes, hogy az inverz mátrix akkor és csak akkor létezik, ha a mátrix elemeiből alkotott

$$|\underline{A}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

determináns nullától különböző. A determinánst az $|\underline{A}|$ szimbólummal jelöljük, mivel a mátrixok körében ugyanazt a szerepet játssza, mint az abszolútérték a valós és a komplex számok körében. (Részletes definíciója megtalálható a megfelelő irodalomban; röviden: a determináns az összes olyan szorzat előjeles összege, amely a mátrix minden sorából és oszlopából pontosan egy elemet tartalmaz. Belátható, hogy ezek száma $n!$.)

Például (lásd F1-1. ábra): $|\underline{F}| = 1(30 - 36) - 3(42 - 32) + 2(63 - 40) = 10$; $|\underline{G}| = 1$.

Az inverz mátrix meghatározása céljából vezessük be az *adjungált mátrix* fogalmát. Jelöljük az \underline{A} mátrix a_{ij} elemzéséhez tartozó *előjeles al-determinánst* (= az adott elemet tartalmazó sor és oszlop elhagyásával nyert mátrix determinánsát) $|\underline{A}_{ij}|$ -vel. Az ezekből képzett

$$\text{adj}(\underline{A}) = \begin{bmatrix} |\underline{A}_{11}| & |\underline{A}_{21}| & \dots & |\underline{A}_{n1}| \\ |\underline{A}_{12}| & |\underline{A}_{22}| & \dots & |\underline{A}_{n2}| \\ \dots & \dots & \dots & \dots \\ |\underline{A}_{1n}| & |\underline{A}_{2n}| & \dots & |\underline{A}_{nn}| \end{bmatrix}$$

mátrixot az \underline{A} mátrix *adjungáltjának* nevezzük. (Láthatjuk, hogy $\text{adj}(\underline{A})$ az $|\underline{A}_{ij}|$ előjeles al-determinánsokból alkotott mátrix transzponáltja.)

Bizonyítás nélkül közöljük az *inverz mátrix* kiszámításának módját:

$$\underline{A}^{-1} = \frac{\text{adj}(\underline{A})}{|\underline{A}|} \quad (|\underline{A}| \neq 0). \quad (\text{F1-2})$$

Az (F1-2) képlettel példaképpen kiszámítottuk az F1-1. ábrán szereplő \underline{F} mátrix \underline{F}^{-1} inverzét (lásd az ábrán; könnyen igazolható, hogy $\underline{F} \underline{F}^{-1} = \underline{F}^{-1} \underline{F} = \underline{E}$).

Megjegyezzük, hogy ha egy valós elemű mátrix ortogonális, illetve egy komplex elemű unitér, akkor az inverze egyenlő a *transzponáltjával*, illetve a konjugált transzponáltjával (lásd pl. \underline{G} , illetve \underline{Z}). Ha még a mátrix szimmetrikus, illetve hermitikus is, akkor *önmaga inverze*. (Lásd pl. \underline{H} , illetve \underline{Z} mátrix az F1-1. ábrán).

F2. VALÓSZÍNŰSÉGSZÁMITÁS

A valószínűségszámítás a matematika egyik modern és napjainkban talán legnagyobb gyakorlati jelentőségű ága. Feladata a *véletlen tömegjelenségek* törvényszerűségeinek vizsgálata. Véletlen jelenségről akkor beszélünk, ha egy folyamat vagy kísérlet lehetséges eredményeit (kimeneteket) az általunk ismert körülmények, okok *nem* határozzák meg egyértelműen.

A digitális képek feldolgozási módszereinek egy része - elsősorban az alakfelismerés területén - feltételezi a valószínűségszámítás alapfogalmainak és alapvető tételeinek ismeretét. Ezért ebben a fejezetben röviden összefoglaljuk a legfontosabb tudnivalókat.

F2.1 ALAPFOGALMAK ÉS ALAPTÉTELEK

F2.1.1 ESEMÉNYTÉR ÉS VALÓSZÍNŰSÉGI MEZŐ

Jelölje E (kszi) azt a halmazt, amelyben egy kísérlet összes lehetséges kimenetének egy és csakis egy halmazelem felel meg. A E halmaz tetszőleges számú eleméből képzett részhalmazok az *események*, összességük az *eseménytér*.

Az események között értelmezhetünk bizonyos műveleteket és bevezethetünk bizonyos fogalmakat. Legyen A és B két esemény:

- Két esemény *összege* az az esemény, mely akkor következik be, ha a kettő közül *legalább az egyik* bekövetkezik. Az összegzési művelet azonos a halmazegyesítéssel (*unióképzéssel*); jelölése ennek megfelelően: $A \cup B$. (Szokásos még az $A+B$ jelölés is, de mivel félreértéseket okozhat, nem használjuk.)
- Két esemény *szorzata* azt az eseményt jelenti, hogy a kettő *egyidejűleg* bekövetkezik. Ez a művelet a halmazelméleti *metszet* megfelelője, jelölése: $A \cap B$. (Bár szokásos, az $A \cdot B$, illetve AB jelölést a fenti okból nem használjuk.)
- Az A és B események *különbsége* az az esemény, hogy A bekövetkezik és *egyidejűleg* B nem. Jelölése: $A \setminus B$, (A szokásos $A - B$ jelölést ugyancsak nem használjuk.)
- Az A esemény *ellentettje* az az esemény, hogy A nem következik be, jelölése \bar{A} . Nyilvánvalóan $\bar{\bar{A}} = A$. Az események különbsége az ellentett segítségével is kifejezhető: $A \setminus B = A \cap \bar{B}$.
- Két eseményt *kizáró eseménynek* tekintünk, ha szorzatuk a lehetetlen esemény: $A \cap B = \emptyset$.
- *Teljes eseményrendszer* olyan páronként *kizáró* események halmaza, amelyek *összege* a biztos esemény.

Az eseményteret *eseményalgebrának* nevezzük és a továbbiakban A -val jelöljük, ha teljesül a következő 3 fel-tétel:

- a) Az \emptyset eleme a Θ (théta) üres halmaz és a Ξ alaphalmaz:

$$\emptyset, \Xi \in A.$$

\emptyset -t *lehetetlen eseménynek* nevezzük, mivel bekövetkezése azt jelentené, hogy a kísérlet eredménye

nem egyezik meg egyik lehetséges kimenetével sem. Ξ -t viszont *biztos eseménynek* hívjuk, mivel bekövetkezése azt jelenti, hogy az eredmény a lehetséges kimenetek egyike.

- b) Az A tetszőleges (akár végtelen) számú elemének egyesítésénél (uniójával) képzett halmazok szintén A elemei; vagyis $A_1, A_2, \dots \in A$ esetén $\bigcup_i A_i \in A$.

Az A *elemi eseményei* azok az események, amelyeket nem lehet előállítani más események uniójaként.

(Kivétel \emptyset , amelyet nem tekintünk elemi eseménynek.)

- c) Az A tetszés szerinti elemeiből képzett különbség-halmaz(ok) ugyancsak eleme(i) A -nak; azaz $A_1, A_2 \in A$ esetén $A_1 \setminus A_2 \in A$.

Valamilyen A esemény akkor következik be, ha a végzett kísérlet eredménye az A részhalmazhoz tartozik; ezzel egyidejűleg mindazok az események is bekövetkeznek, amelyeknek a szóban forgó esemény az elemük.

Tekintsük például a játékkockával végzett dobási kísérletet. A kockadobás lehetséges kimeneteinek halmaza: $\Xi = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, ahol például x_1 azt az elemi eseményt jelöli, hogy a kockával 1-et dobunk.

Legyen A_0 a Ξ halmaz részhalmazainak következő rendszere:

$$A_0 = \{\emptyset, \{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_1, x_2, x_3, x_4, x_5, x_6\} = \Xi\}.$$

Az A_0 halmaz láthatóan kielégíti a fenti három feltételt, tehát eseményalgebra; ebben azonban az egyes kockadobások már nem elemi események.

Végezzünk el egy kísérletet n -szer egymás után, azonos körülmények közt. Tegyük fel, hogy valamely általunk megfigyelt esemény ezalatt k -szor következik be. Ekkor k -t

az esemény gyakoriságának, a k/n hányadost *relatív gyakoriságnak* nevezzük. A kísérletek számának növelésével (ha a körülmények változatlanok) a relatív gyakoriság általában stabilitást mutat. Azt a számot, amely körül a relatív gyakoriság ingadozik, az esemény *valószínűségének* nevezzük. A valószínűség 0 és 1 közötti valós szám.

A továbbiakban szükségünk lesz a valószínűség pontosabb definíciójára, ezt a következőkben adjuk meg:

Legyen Ξ tetszőleges eseménytér, A pedig Ξ -beli eseményalgebra. Ekkor a

$$P: A \rightarrow [0, 1]$$

valós értékű leképezést *valószínűségi mértéknek* vagy röviden *valószínűségnek* nevezzük, ha teljesül a következő 3 axióma:

1. Minden $A \in \mathcal{A}$ eseményre $0 \leq P(A) \leq 1$.
2. A biztos esemény valószínűsége 1; vagyis $P(\Xi) = 1$.
3. Legyen A_1, A_2, \dots legfeljebb megszámlálhatóan végtelen sok, egymást páronként kizáró esemény A -ban. Ekkor az *egyesítésükkel* keletkezett eseménynek (amely a fenti b) feltétel szerint szintén eleme A -nak!) valószínűsége az egyes események valószínűségének *összege*. Képlettel:

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i); \text{ ha } A_i \cap A_j = \emptyset, \forall i \neq j.$$

Az ilymódon definiált (Ξ, \mathcal{A}, P) hármast *valószínűségi mezőnek* nevezzük.

A fentiekből következik, hogy $P(\emptyset) = 0$. A továbbiakban 0 -val, illetve 1 -vel jelöljük és lehetetlen, illetve biztos eseménynek nevezzük mindazokat az eseményeket, amelyeknek valószínűsége 0 , illetve 1 .

Felhívjuk a figyelmet arra, hogy a lehetetlen és a biztos eseménynek ez a definíciója eltér az előzőtől. Lehet ugyanis olyan valószínűségi mezőt definiálni, amelyben egyrészt a 0 valószínűségű (lehetetlen) események is bekövetkezhetnek, másrészt 1 valószínűségű (biztos) eseményhez jutunk akkor is, ha Ξ -ből elhagyunk 0 valószínűségű eseményeket. (Ha pl. egy céltábla eltalálása a biztos esemény, akkor egy kijelölt pont eltalálása lehetetlen esemény, mivel 0 a valószínűsége - de bekövetkezhet. Hasonlóan biztos esemény, mivel 1 a valószínűsége, hogy eltaláljuk a céltáblát akkor is, ha legfeljebb megszámlálhatóan végtelen sok pontját kizárjuk belőle (vagyis az ide eső találatot érvénytelennek tekintjük), mégis lehetséges, hogy ez esetben nincs találatunk.)

Az eddigiek alapján könnyen beláthatók az eseményekre vonatkozó alábbi összefüggések:

$$\begin{aligned} \bar{0} &= I, & \bar{I} &= 0, & \bar{\bar{A}} &= A, \\ AUA &= A, & AU\bar{A} &= I, & AUI &= I, & AUO &= A, \\ A\bar{A} &= A, & A\bar{A}\bar{A} &= 0, & A\bar{A}I &= A, & A\bar{A}O &= 0. \end{aligned}$$

Ugyancsak könnyen adódnak a következők:

- tetszőleges A eseményre $P(\bar{A}) = 1 - P(A)$;
- ha A_1, \dots, A_n teljes eseményrendszer, akkor

$$\sum_{i=1}^n P(A_i) = 1; \quad (\text{F2-1})$$

- tetszőleges A és B eseményre fennáll:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (\text{F2-2})$$

Végül belátható még, hogy tetszőleges B eseményre nézve, amelyre $P(B) \neq 0$, a

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (\text{F2-3})$$

definícióval adott $P(\cdot|B)$ halmazfüggvény ugyancsak valószínűségi mérték A -n. A $P(A|B)$ valószínűséget az A esemény (B feltétel mellett) feltételes valószínűségének, a $P(\cdot|B)$ függvényt pedig a B feltétel mellett feltételes valószínűségi mértéknek nevezzük.

Az A esemény akkor független a B eseménytől, ha bekövetkezése nem függ B bekövetkezésétől. Természetesen ekkor B is független A -tól; képlettel:

$$P(A|B) = P(A); \quad P(B|A) = P(B).$$

(F2-3)-ból következik, hogy független eseményekre

$$P(A \cap B) = P(A) \cdot P(B). \quad (\text{F2-4})$$

F2.1.2 VALÓSZÍNŰSÉGSZÁMITÁSI TÉTELEK

1. *Teljes valószínűség tétele.* Alkossanak az $\{A_1, \dots, A_n\}$ események teljes eseményrendszert, és tegyük fel, hogy $P(A_i) \neq 0$, ($i = 1, 2, \dots, n$). Ekkor tetszőleges B eseményre fennáll:

$$P(B) = \sum_{i=1}^n P(B|A_i) \cdot P(A_i). \quad (\text{F2-5})$$

A tétel egyszerűen belátható, ha felidézzük, hogy $P(B) = P(B \cap I)$; továbbá $I = A_1 \cup A_2 \cup \dots \cup A_n$. Ezekből (F2-3) átrendezésével a 3. axióma alapján adódik (F2-5).

2. *Bayes-tétel.* Legyen $\{A_1, \dots, A_n\}$ teljes eseményrendszer, és $P(A_i) \neq 0$ ($i = 1, 2, \dots, n$). Ekkor tetszőleges B eseményre, amelyre $P(B) \neq 0$, igaz a következő összefüggés:

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^n P(B|A_i) \cdot P(A_i)} \quad (\text{F2-6})$$

Ugyanis (F2-3) alapján felírhatjuk, hogy

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j \cap B)}{P(A_j)} \cdot \frac{P(A_j)}{P(B)} = P(B|A_j) \cdot \frac{P(A_j)}{P(B)},$$

amiből (F2-5) behelyettesítésével már következik az (F2-6) állítás.

A Bayes-tétel lényegében azt mondja ki, hogy ha ismerjük a B esemény feltételes valószínűségeit egy teljes eseményrendszerre nézve, akkor az egyes komponensesemények B -re vonatkozó feltételes valószínűségeit is kiszámíthatjuk. A $P(A_j)$ valószínűségeket "a priori", a $P(A_j|B)$ valószínűségeket "a posteriori" valószínűségeknek is nevezik.

3. *Valószínűségek szorzástétele.* Tetszőleges $\{A_1, \dots, A_n\}$ eseményrendszer mellett

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_n) &= \\ &= P(A_n | A_1 \cap \dots \cap A_{n-1}) \cdot P(A_{n-1} | A_1 \cap \dots \cap A_{n-2}) \cdot \dots \cdot P(A_1). \end{aligned} \quad (\text{F2-7})$$

F2.2 VALÓSZINŰSÉGI VÁLTOZÓK

Legyen adva egy (Ξ, A, P) valószínűségi mező. Rendeljünk hozzá Ξ minden eleméhez, vagyis az egyes eseményekhez egy-egy valós számot. Ily módon valószínűségi változókhoz jutunk, amelyek adott számérték(ek)et valamilyen esemény(ek) bekövetkezésekor (az eseményével azonos valószínűséggel) vesznek fel. Pontosabban a

$$\zeta: \Xi \rightarrow \mathbb{R}$$

függvényt *valószínűségi változónak* nevezzük, ha tetszőleges $z \in \mathbb{R}$ esetén a

$$\Xi(\zeta < z) = \{x \in \Xi \mid \zeta(x) < z\}$$

halmaz A -ban van. Hasonlóan értelmezhető az

$$\underline{v}: \Xi \rightarrow \mathbb{R}^n$$

valószínűségi vektorváltozó fogalma is.

(Megjegyezzük, hogy a "valószínűségi változó" elnevezést nem tartjuk helyesnek, de megtartjuk, mivel általában ezt használják. Helyesebb volna "eseményváltozó függvény"-nek hívni. Ugyanis - mivel értelmezési tartománya a Ξ , értékkészlete az \mathbb{R} halmaz - véletlenszerűen - bár adott valószínűséggel - bekövetkező eseményekhez rendel valós számokat.) Könyvünkben a valószínűségi változókat ζ (dzéta), illetve v (űpszilon) görög kisbetűkkel jelöljük.

Definiáljuk a következő függvényt:

$$S_{\zeta}(z) = P(\Xi(\zeta < z)) = P(\{x \in \Xi \mid \zeta(x) < z\}). \quad (F2-8)$$

Ezt az $S_\zeta: \mathbb{R} \rightarrow \mathbb{R}$ függvényt nevezzük a ζ valószínűségi változó *eloszlásfüggvényének*. Hasonlóan értelmezhető a valószínűségi vektorváltozó együttes eloszlásfüggvénye is. A továbbiakban - ahol félreértést nem okozhat - (F2-8) helyett általában az

$$S(z) = P(\zeta < z) \quad (\text{F2-8/a})$$

rövidített írásmódot használjuk. Az eloszlásfüggvény alaptulajdonságai:

- monoton növekvő;
- $\lim_{z \rightarrow -\infty} S(z) = 0, \quad \lim_{z \rightarrow +\infty} S(z) = 1;$
- \mathbb{R} -en mindenütt jobbról folytonos és balról van határértéke.

Igaz az az állítás is, hogy ha S olyan függvény, amely rendelkezik a fenti három tulajdonsággal, akkor létezik olyan valószínűségi változó, melynek ez az eloszlásfüggvénye.

A valószínűségi változók két nagy csoportját szokás külön kiemelni, az egyiket a diszkrét, a másikat az (abszolút) folytonos valószínűségi változók alkotják.

Egy valószínűségi változót *diszkrét* (eloszlásu) *valószínűségi változónak* nevezünk, ha *értékkészlete megszámlálható*. Legyenek a ζ lehetséges értékei $z_1, z_2, \dots;$ ekkor a

$$p_i = P(\zeta = z_i) \quad (i = 1, 2, \dots) \quad (\text{F2-9})$$

valószínűségeket. - nem egészen precizen - ζ *eloszlásának* hívjuk. Ezek felhasználásával a diszkrét valószínűségi változó eloszlásfüggvényét - amely most nyilván lépcsős-függvény - az

$$S(z) = P(\zeta < z) = \sum_{z_i < z} p_i \quad (\text{F2-10})$$

alakban írhatjuk fel.

Ha az *eloszlásfüggvény* a teljes számegyenesen folytonos, továbbá - véges sok hely kivételével - differenciálható, akkor a ζ valószínűségi változót (abszolút) folytonosnak nevezzük. Ha az eloszlásfüggvény derivált függvénye legalább szakaszonként folytonos, akkor az

$$s(z) = S'(z) \quad (\text{F2-11})$$

függvényt a folytonos valószínűségi változó *sűrűségfüggvényének* nevezzük. Érvényesek a következő összefüggések:

$$S(a) = P(\zeta < a) = \int_{-\infty}^a s(z) dz;$$

$$S(b) - S(a) = P(a \leq \zeta < b) = \int_a^b s(z) dz;$$

$$\int_{-\infty}^{+\infty} s(z) dz = 1.$$

F2.2.1 NEVEZETES ELOSZLÁSOK

Az alábbiakban felsorolunk néhány nevezetes eloszlást; folytonos esetben megadjuk a sűrűségfüggvényt is.

F2.2.1.1 Diszkrét eloszlások

1. Legyen a valószínűségi változó értékkészlete a $\{0, 1, \dots, n\}$ halmaz, és legyen $0 < p < 1$ valós szám.

A ζ valószínűségi változó n -edrendű, p paraméterű *bínomiális eloszlásu*, ha az (F2-9)-ben definiált eloszlása:

$$P(\zeta = i) = \binom{n}{i} p^i (1-p)^{n-i}. \quad (\text{F2-12})$$

2. Legyen az értékkészlet ismét a $\{0, 1, \dots, n\}$ halmaz, továbbá legyen $M \leq N$ két adott természetes szám. A ζ valószínűségi változó *hipergeometrikus eloszlásu*, ha az (F2-9)-ben definiált eloszlása.

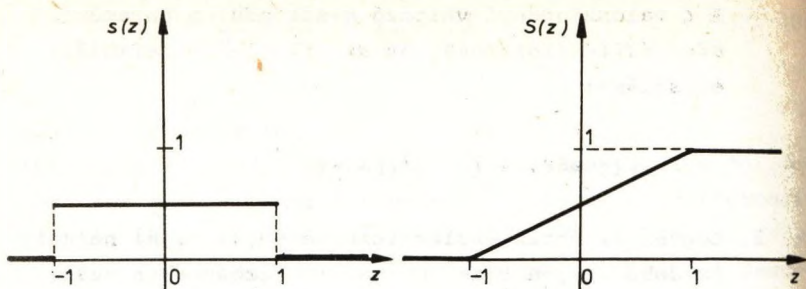
$$P(\zeta = i) = \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}. \quad (\text{F2-13})$$

3. Legyen most az értékkészlet a nemnegatív egész számok halmaza \mathbb{E} , és legyen $\lambda > 0$ valós szám. A ζ valószínűségi változó *Poisson-eloszlásu*, ha az (F2-9)-ben definiált eloszlása:

$$P(\zeta = i) = \frac{\lambda^i e^{-\lambda}}{i!}. \quad (\text{F2-14})$$

F2.2.1.2 Folytonos eloszlások

1. Legyenek $a, b \in \mathbb{R}$ valós számok, és $a < b$. A ζ valószínűségi változó *egyenletes eloszlásu* (lásd F2-1. ábra) az (a, b) intervallumon, ha sűrűségfüggvénye:



F2-1. ábra. Az egyenletes eloszlás sűrűség- (a) és eloszlásfüggvénye (b), a $[-1, 1]$ intervallumban

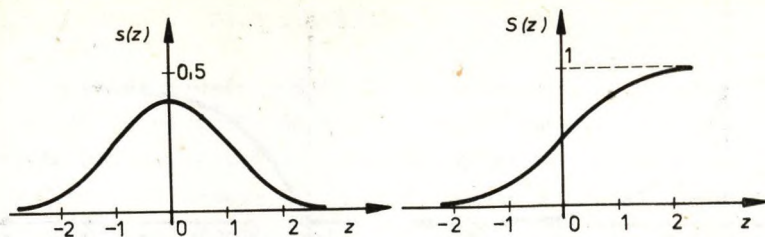
$$s(z) = \begin{cases} 0, & \text{ha } z \leq a \text{ vagy } z \geq b; \\ \frac{1}{b-a}, & \text{ha } a < z < b. \end{cases} \quad (\text{F2-15/a})$$

Ekkor az eloszlásfüggvénye:

$$S(z) = \begin{cases} 0, & \text{ha } z \leq a; \\ \frac{z-a}{b-a}, & \text{ha } a < z < b; \\ 1, & \text{ha } z \geq b. \end{cases} \quad (\text{F2-15/b})$$

2. Legyenek μ, σ valós számok, $\sigma > 0$. A ζ valószínűségi változó (μ, σ) paraméterű *normális eloszlású* (lásd F2-2. ábra) - jelölése $N(\mu, \sigma)$ - ha sűrűségfüggvénye:

$$s(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) \quad (z \in \mathbb{R}). \quad (\text{F2-16/a})$$



F2-2. ábra. A $\mu = 0$, $\sigma = 1$ paraméterű normális eloszlás sűrűség- (a) és eloszlásfüggvénye (b)

Ekkor az eloszlásfüggvénye:

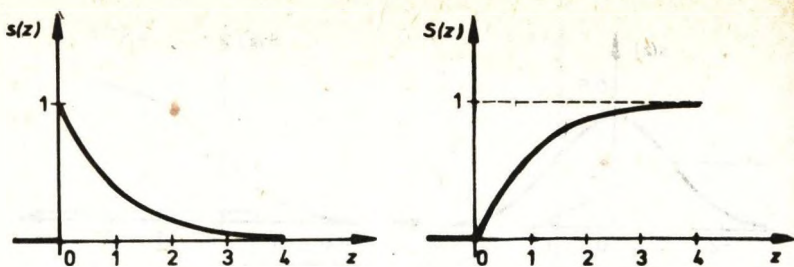
$$S(z) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right) dt \quad (z \in \mathbb{R}). \quad (\text{F2-16/b})$$

3. Legyen $\lambda > 0$ valós szám. A ζ valószínűségi változó *exponenciális eloszlásu* (lásd F2-3. ábra), ha sűrűségfüggvénye:

$$s(z) = \begin{cases} \lambda e^{-\lambda z}, & \text{ha } z \geq 0; \\ 0, & \text{ha } z < 0; \end{cases} \quad (\text{F2-17/a})$$

illetve eloszlásfüggvénye:

$$S(z) = \begin{cases} 1 - e^{-\lambda z}, & \text{ha } z \geq 0; \\ 0, & \text{ha } z < 0. \end{cases} \quad (\text{F2-17/b})$$



F2-3. ábra. A $\lambda = 1$ paraméterű exponenciális eloszlás sűrűség- (a) és eloszlásfüggvénye (b)

F2.2.2 VALÓSZÍNŰSÉGI VÁLTOZÓK TELJES FÜGGETLENSÉGE

Legyenek ζ_1, \dots, ζ_n valószínűségi változók. Jelölje $S_i(z_i)$ a ζ_i eloszlásfüggvényét, és $S(z)$ az együttes eloszlásfüggvényét. A ζ_i valószínűségi változókat teljesen függetleneknek nevezzük, ha az együttes eloszlásfüggvényü előáll eloszlásfüggvényeik szorzataként, azaz

$$S(z_1, z_2, \dots, z_n) = S_1(z_1) \cdot S_2(z_2) \dots S_n(z_n). \quad (\text{F2-18})$$

Valószínűségekkel fogalmazva, ha

$$P(\zeta_1 < z_1, \zeta_2 < z_2, \dots, \zeta_n < z_n) = P(\zeta_1 < z_1) P(\zeta_2 < z_2) \dots P(\zeta_n < z_n).$$

Megjegyezzük, hogy (abszolút) folytonos valószínűségi változó esetén ekkor a sűrűségfüggvényekre is hasonló összefüggés érvényes:

$$s(z_1, z_2, \dots, z_n) = s_1(z_1) s_2(z_2) \dots s_n(z_n).$$

F2.3 STATISZTIKAI JELLEMZŐK

A valószínűségi változók jellemzésére - az eloszlás típusán és paraméterén túl - számos esetben szükség van egyéb paraméterekre is; ezeket együttesen *statisztikai jellemzőknek* nevezzük. Az alábbiakban a képfeldolgozás szempontjából fontosakat tekintjük át.

F2.3.1 MOMENTUMOK

Az eloszlás- és sűrűségfüggvények statisztikai jellemzői közül a momentumokat és a centrális momentumokat használjuk leggyakrabban. Az ezékt definíciókhoz további fogalmak bevezetésére volna szükség, számunkra azonban - a felmerülő esetekben - kevésbé általános értelmezés is elegendő. Ennek megfelelően az egyes jellemzőket külön-külön definiáljuk a diszkrét és az (abszolút) folytonos esetre.

Az n -edik momentumot, illetve az n -edik centrális momentumot folytonos esetben a sűrűségfüggvény ismeretében a következőképpen lehet meghatározni:

$$M_n(\zeta) = \int_{-\infty}^{+\infty} z^n s(z) dz \quad (n = 1, 2, \dots); \quad (\text{F2-19/a})$$

illetve

$$D_n(\zeta) = \int_{-\infty}^{+\infty} (z - M_1(\zeta))^n s(z) dz \quad (n = 2, 3, \dots). \quad (\text{F2-19/b})$$

Diszkrét esetben a fenti integrálok helyett összegzéssel kaphatjuk meg a momentumokat, illetve a centrális momentumokat.

Az (F2-9) szerinti jelölésekkel:

$$M_n(\zeta) = \sum_{i=1}^{\infty} z_i^n p_i, \quad (\text{F2-20/a})$$

illetve

$$D_n(\zeta) = \sum_{i=1}^{\infty} (z_i - M_1)^n p_i. \quad (\text{F2-20/b})$$

A momentumok közül a leggyakrabban használatos az $n = 1$ esetben adódó *várhatóérték* (M_1), míg a centrális momentumok közül az $n = 2$ esetnek megfelelő *szórásnégyzet* (D_2). Erre a két mennyiségre - fontosságuk miatt - külön jelölést is bevezetünk: a várhatóértéket μ -vel, a szórást σ -val (illetve a szórásnégyzetet σ^2 -tel) jelöljük. Így *folytonos* esetben

$$\mu = M_1(\zeta) = \int_{-\infty}^{+\infty} z s(z) dz, \quad (\text{F2-21/a})$$

$$\sigma = \sqrt{D_2(\zeta)} = \sqrt{\int_{-\infty}^{+\infty} (z - \mu)^2 s(z) dz};$$

illetve *diszkrét* esetben

$$\mu = M_1(\zeta) = \sum_{i=1}^{\infty} z_i p_i, \quad (\text{F2-21/b})$$

$$\sigma = \sqrt{D_2(\zeta)} = \sqrt{\sum_{i=1}^{\infty} (z_i - \mu)^2 p_i}.$$

Érdeemes megjegyezni néhány speciális eloszlású folytonos valószínűségi változó jellegzetes momentumait:

- az (a, b) intervallumon egyenletes eloszlású

$$\text{várhatóértéke } \frac{a + b}{2}, \text{ szórása } \frac{b - a}{2\sqrt{3}};$$

- az $N(\mu, \sigma)$ paraméterű, normális eloszlású várhatóértéke μ , szórása σ ;
- a λ paraméterű exponenciális eloszlású várhatóértéke és szórása $1/\lambda$.

F2.3.2 MEDIÁN

Az a valós számérték, melynél a valószínűségi változó $1/2$ valószínűséggel kisebb, illetve $1/2$ valószínűséggel nagyobb vagy egyenlő, az eloszlás mediánja. Eloszlásfüggvénnyel fogalmazva a ζ mediánja z , ha $S(z) = 1/2$.

F2.3.3 KORRELÁCIÓS EGYÜTTHATÓ

Legyenek ζ és ν valószínűségi változók. Tegyük fel, hogy az elsőrendű momentumaik $M_1(\zeta)$ és $M_1(\nu)$, illetve a másodrendű centrális momentumaik $D_2(\zeta)$ és $D_2(\nu)$ léteznek és végesek. Ekkor a

$$\text{cov}(\zeta, \nu) = M((\zeta - M(\zeta))(\nu - M(\nu))) = M(\zeta\nu) - M(\zeta)M(\nu) \quad (\text{F2-22})$$

értéket, amely ugyancsak létezik, a ζ és az ν valószínűségi változó *kovarianciájának* nevezzük. Ha $\sigma^2(\zeta)$ és $\sigma^2(\nu)$ nem nulla, az ebből nyerhető

$$r(\zeta, \nu) = \frac{\text{cov}(\zeta, \nu)}{\sigma(\zeta)\sigma(\nu)}$$

(F2-23)

érték a ζ és az ν korrelációs együtthatója.

Két valószínűségi változó korrelálatlan, ha $\text{cov}(\zeta, \nu) = 0$.

A korrelációs együttható és a kovariancia alaptulajdonságai közül kiemelünk néhányat:

- $\text{cov}(\zeta, \zeta) = \sigma^2(\zeta)$;
- ha ζ és ν függetlenek, $\text{cov}(\zeta, \nu) = 0$; ha mindkét valószínűségi változó normális eloszlású, akkor az állítás megfordítása is igaz;
- $r(\zeta, \nu) \leq 1$;
- $r(\zeta, \nu) = \pm 1$; akkor és csak akkor, ha $\nu = \pm a\zeta + b$, valamilyen $a > 0$ és b valós számok esetén.

F3. DISZKRÉT GEOMETRIA

Diszkrét képsíkot pl. a következőképpen konstruálhatunk:

- értelmezzük a $\{p\}$ folytonos képsíkot mint egy olyan Descartes-féle koordináta-rendszer pontjainak halmazát, melyet az abszcisszatengelyre tükröztünk;
- legyen $\{\pi\}$ diszkrét képsík azon pontok halmaza, amelyek egyrészt elemei a $\{p\}$ folytonos képsíknak, másrészt koordinátái egész számok.

Igy:

$$\{\pi\} = \{P(x, y) \mid P(x, y) \in \{p\} \wedge x \in \mathbb{N} \wedge y \in \mathbb{N}\}.$$

A $\{\pi\}$ diszkrét képsík elemeit *képpontoknak* nevezzük. Ezek - a fenti jelölés értelmében - a $\{p\}$ folytonos képsík azon pontjai, amelyeknek mindkét koordinátájuk egész szám. Nyilvánvaló, hogy a $\{p\}$ egy korlátos részhalmazának megfeleltetett $\{\pi\}$ diszkrét képsík véges számú képpontból áll. A korlátos képsíkot mindig *téglalap alakunak* tekintjük; evégett szükség szerint kiegészítjük "üres" képpontokkal. Ilyenkor - a tv-technikában szokásos megjelenítési mód miatt - a koordináta-rendszer kezdőpontját a *bal felső* sarokba helyezzük; továbbá az x tengelyt balról jobbra, az y tengelyt *felülről lefelé* irányítjuk úgy, hogy a téglalap megfelelő oldalai rájuk essenek.

F3.1 TOPOLÓGIAI JELLEMZŐK

Képzeljük el a folytonos $\{p\}$ képsikot rugalmas gumilemeznek. Topológiaiainak nevezzük azokat a sajátságokat és relációkat, amelyek tetszőleges, anizotrop (= iránytól függő, különböző irányokban általában különböző mértékű) erőhatásokra bekövetkező torzulások esetén is érvényben maradnak.

Ez a modell a digitális képsikra is átvihető. A topológiai jellemzők közül számunkra a szomszédság és az összefüggőség fontos, ezekkel foglalkozunk részletesebben.

F3.1.1 SZOMSZÉDSÁG

Ahhoz, hogy egy halmaz elemei között szomszédsági relációkat állapíthassunk meg, először *rendeznünk* kell őket, és meg kell határozni a megengedett haladási irányokat. Ezek után két halmazelemet *szomszédosnak* nevezünk, ha valamelyik megengedett irányban a *lehető legközelebb* vannak egymáshoz, vagyis *nincs közöttük más* halmazelem az adott rendezési szabály szerint.

A diszkrét képsik esetében a rendezettséget a *képpkoordináták* adják, ezek szabják meg a lehetséges haladási irányokat is. (Az x irányu képpkoordináta a képpont *oszlop*, az y irányu a *sorszám*. A rövidség kedvéért könyvünkben koordinátán mindig képpkoordinátát értünk.)

Tekintsük a (k, l) koordinátájú $P(k, l) \in \{\pi\}$ képpontot:

- a) Ha csak a négy fővilágtáj szerinti haladást engedjük meg, az F3-1/a ábra szerinti $P_0(k + 1, l)$, $P_1(k, l - 1)$, $P_2(k - 1, l)$, illetve $P_3(k, l + 1)$ négy képpontot a $P(k, l)$ képpont *négyesszomszédainak* (röviden: (4)-szomszédainak), a négy pontot együttesen (4)-szomszédságnak nevezzük.

$k-2$	$k-1$	k	$k+1$	$k+2$	
					$l-2$
		P_1			$l-1$
	P_2	P	P_0		l
		P_3			$l+1$
					$l+2$

a)

$k-2$	$k-1$	k	$k+1$	$k+2$	
					$l-2$
	P_3	P_2	P_1		$l-1$
	P_4	P	P_0		l
	P_5	P_6	P_7		$l+1$
					$l+2$

b)

F3-1. ábra. Négyes- és nyolcasszomszédság

b) Ha a haladás a mellékvilágtájak irányában is megengedett, a $P(k, l)$ képpontnak nyilvánvalóan nyolc *nyolcasszomszédja* (röviden: (8)-szomszédja) lesz:

$$P_0(k+1, l), P_1(k+1, l-1), P_2(k, l-1), P_3(k-1, l-1),$$

$$P_4(k-1, l), P_5(k-1, l+1), P_6(k, l+1), P_7(k+1, l+1);$$

ezek együttesen a (8)-szomszédságot alkotják (lásd F3-1/b ábra).

A mondottakból kitűnik, hogy a (8)-szomszédság magában foglalja a (4)-szomszédságot.

F3.1.2 UTVONAL

A (4/8)-utvonal képpontok véges sorozata, amelyben - az elsőt kivéve - mindegyik pont (4/8)-szomszédja az előtte állónak. Az utvonal akkor *egyszeres*, ha bármelyik elemének legfeljebb két szomszédja tartozik bele. *Végpon-*

toknak nevezzük azokat az elemeket, amelyeknek pontosan egy (4/8)-szomszédjuk van.

Rendeljük hozzá egy képpont szomszedságához az F3-1. ábrán a pontok indexelésére használt $\{0 - 3\}$, illetve $\{0 - 7\}$ számokat. Ezeket *Freeman-féle iránykódvektoroknak*, röviden *iránykódoknak* nevezzük, és α_i -vel jelöljük. Ha $\alpha_i \equiv \alpha_m \pmod{4/8}$, akkor $\alpha_i = \alpha_m$.

Ábrázoljunk egy egyszeres utat úgy, hogy - a kezdőponttól kiindulva - felsoroljuk a következő pontra mutató iránykódokat. Ekkor az utvonal *lánckódját* kapjuk, amely a haladás irányát is előírja. Az F3-2/a, illetve F3-2/b ábrán megadjuk egy *négyes-*, illetve egy *nyolcasutvonal* mindkét irányu lánckódját.

Jelölje egy egyszeres (4/8)-utvonal egyik irányu ($P_1 \rightarrow P_2$) lánckódját az $\alpha_1, \alpha_2, \dots, \alpha_n$ iránykódsorozat, akkor a fordított irányu lánckódot az

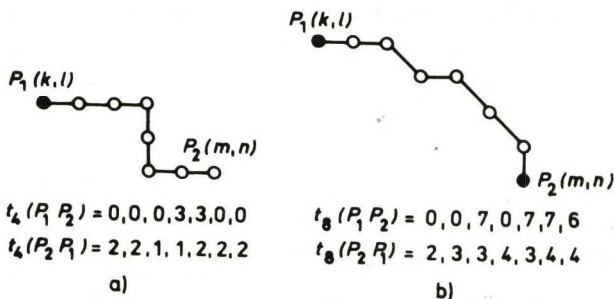
$$\alpha'_i = \alpha_{n-i} + 2 \pmod{4},$$

illetve az

$$\alpha'_i = \alpha_{n-i} + 4 \pmod{8}$$

(F3-1)

összefüggés adja (4)-, illetve (8)-utvonal esetén.



F3-2. ábra. Négyes- és nyolcasutvonal

F3.1.3 ÖSSZEFÜGGŐSÉG

A diszkrét képsík pontjait *értékes* pontoknak nevezzük, ha valamilyen értékelhető információt hordoznak; a többi pont *nem értékes vagy háttérpont*. (Az értékes pontokat különböző tulajdonságaik alapján különböző osztályokba is sorolhatjuk, lásd pl. a 4. és az 5. fejezetet; ez azonban nem érinti az alábbiakat.)

Két képpontot (egy osztályozásra nézve) (4/8)-*összefüggőnek* mondunk, ha létezik olyan (4/8)-*utvonal*, amelynek

- mindkét pont eleme, *és*
- a vizsgált két pont közé eső minden pontja értékes (illetve ugyanahhoz az osztályhoz tartozik).

Megjegyezzük, hogy mivel a *digitális kép* a $\{\pi\}$ digitális képsíknak egy véges tartománya, szükségképpen vannak olyan pontjai, amelyeknek *legalább egy szomszédjuk nem tartozik* a képhez. Ezeket a pontokat (kép)*határpontoknak* nevezzük.

Azok a nem értékes képpontok, amelyek nem függnek össze egyetlen határponttal sem, *lyuka(ka)t* alkotnak.

Ha egy értékes képpont minden szomszédja háttérpont vagy fordítva (azaz ha egy képpontnak egyik szomszédja sem tartozik vele azonos osztályba), *iszlandt* (elszigetelt) pontról beszélünk.

Az értékes (illetve különböző osztályokhoz tartozó) képpontokból álló tartományok határpontjai, amelyeknek legalább egyik szomszédjuk nem értékes (illetve másik osztályhoz tartozik), a *konturpontok*. Az összefüggő konturpontok halmaza a *kontur(vonal)*. A (4)-összefüggő tartományok konturvonalala (8)-összefüggő és fordítva.

Az F3-3. ábrán látható 4 értékes pont

- (4)-szomszédtság esetén négy különálló tartomány,

amelyek egy izolált pontból álló "lyukat"
fognak körül;

- (8)-szomszédság esetében egyetlen tartományt alkot, amelyben nincs lyuk.

	0	0	0	0	0
	0	0	1	0	0
	0	1	0	1	0
	0	0	1	0	0
	0	0	0	0	0

F3-3. ábra. Négyes- és nyolcas-összefüggőség

F3.2 GEOMETRIAI JELLEMZŐK

A geometriai jellemzők hosszúság- és szögmértékeken alapulnak; ezek a "gumilemez" torzulása esetén megváltoznak. A szögmértékeknek a képfeldolgozásban kisebb jelentőségük van, ezért részletesen csak a távolságmértékeket, illetve ezekkel összefüggésben a normákat tárgyaljuk.

F3.2.1 DISZKRÉT METRIKÁK

Értelmezzük a távolságot a diszkrét képsíkon. Ahhoz, hogy a ρ függvény diszkrét képsíkbeli távolságfüggvény

(=metrika) lehessen, tetszőleges $P_1, P_2, P_3 \in \pi$ pontra teljesítenie kell a következő feltételeket:

I. pozitív definit:

$$\rho(P_1, P_2) \geq 0 \text{ és } \rho(P_1, P_2) = 0 \Leftrightarrow P_1 = P_2;$$

II. szimmetrikus:

$$\rho(P_1, P_2) = \rho(P_2, P_1);$$

III. érvényes a háromszög-egyenlőtlenség:

$$\rho(P_1, P_3) \leq \rho(P_1, P_2) + \rho(P_2, P_3);$$

IV. reguláris:

$$\rho(P_2, P_3) = \rho(P_1, P_3) - 1 \mid \rho(P_1, P_2) = 1 \wedge \rho(P_1, P_3) \geq 2.$$

(A képlet lényeges tartalma, hogy ilyen P_2 pont létezik. A regularitás azt fejezi ki, hogy a diszkrét képsík képpontjai egyenletes sűrűséggel helyezkednek el; folytonos metrikák esetében ez magától értetődik.)

Folytonos esetben a $P_1(k, l) \in \{p\}$ és a $P_2(m, n) \in \{p\}$

pontok *euklidészi* távolsága:

$$\rho_E(P_1, P_2) = \sqrt{(k - m)^2 + (l - n)^2}. \quad (F3-2)$$

Bár ez - bizonyos közelítéssel - a diszkrét képsíkra is átvihető, ehhez jobban illeszkednek a következő távolságmértékek:

a) A $P_1(k, l)$ és a $P_2(m, n) \in \{\pi\}$ pontok *négyestávolságdát* a (4)-szomszédság alapján a

$$\rho_A(P_1, P_2) = |k - m| + |l - n| \quad (F3-3)$$

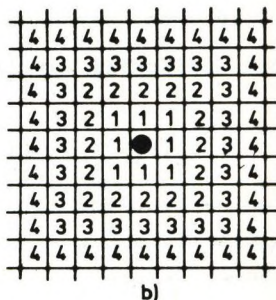
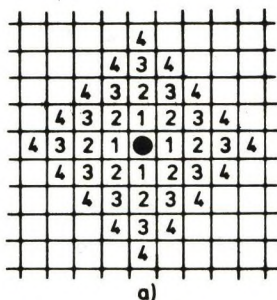
metrika definiálja. Mivel ilyenkor csak koordináta-irányokban lehet haladni, ezt a távolságot - New York jellegzetes városnegyedéről - *manhattan-távolságnak* is nevezik. A definíciókból következik, hogy a (4)-szomszédok (4)-távolsága 1, hi-

szen ez a lehető legkisebb diszkrét távolságérték (lásd F3-4/a ábra).

- b) Ugyanezen két pont *nyolcastávolsága* a (8)-szomszéd-ságon alapuló

$$\rho_8(P_1, P_2) = \max(|k - m|, |l - n|) \quad (\text{F3-4})$$

metrikából származhat, amelyet *sakktávolságnak* is neveznek, mivel ez esetben a lehetséges haladási irányok azonosak a vezéréivel a sakktáblán. A definíció szerint a (8)-szomszédok (8)-távolsága ugyancsak 1 (lásd F3-4/b ábra). Vegyük még észre, hogy két pont (4/8)-távolsága megegyezik az őket összekötő legrövidebb (4/8)-utvonal leírásához szükséges iránykódok számával.



F3-4. ábra. Négyes- és nyolcastávolságok

Az F3-4. ábrán megjelöltük az adott ponttól 1-4 távolságra levő pontokat, mindkét metrika szerint. Láthatjuk, hogy mindkettő *erősen irányfüggő*, és *rosszul közelíti* az euklidészi.

- c) Jó eredmény érhető el a ρ_E és a ρ_8 metrika kombinálásával. Használjuk lokálisan az euklidészi távolságfüggvényt, és ennek megfelelően rendeljük

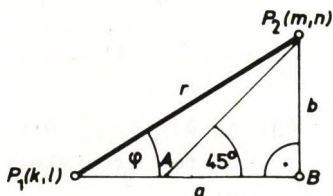
a (8)-as szomszédságu diszkrét képsík páros számú iránykódjaihoz az 1, a páratlan számúakhoz a $\sqrt{2}$ mennyiséget.

Ennek megfelelően a $P_1(k, l)$ és $P_2(m, n)$ pontok közötti E8-távolságot a

$$\rho_{E8}(P_1, P_2) = \sum_{i=0}^{n-a} \rho_{\alpha_i} \quad \rho_{\alpha_i} \{1, \sqrt{2}\} \quad (E3-5)$$

összefüggéssel definiáljuk, ahol a ρ_{α_i} ($i = 0, \dots, n-1$) sorozat a P_1 és a P_2 pontokat összekötő legrövidebb (8)-utvonálhoz tartozó iránykódsorozatnak megfeleltetett euklidészi távolságok sorozata.

Vizsgáljuk meg ρ_E és ρ_{E8} viszonyát. Legyen $P_1(k, l)$ és $P_2(m, n)$ két $\{\pi\}$ -beli pont (lásd F3-5. ábra).



F3-5. ábra. E8-távolság meghatározása

Legyen:

$$\overline{P_1 B} = \max(|k - m|, |l - n|) = \rho_8(P_1, P_2) = a,$$

$$\overline{P_2 B} = \min(|k - m|, |l - n|) = b,$$

$$\overline{AB} = \overline{P_2 B} \text{ (ui.: } ABP_2\Delta \text{ egyenlő szárú),}$$

$$\overline{P_1 A} = a - b.$$

Igy a $\overline{P_1 P_2}$ szakaszt $\overline{P_1 A} = a - b$ tengely irányu és $\overline{AB} = b$ átlós irányu elemi lépésre (iránykódra) bontottuk fel. (A $\overline{P_1 P_2}$ szakasz ezeket az iránykódokat meghatározott szabályok szerint tartalmazhatja, lásd az F3.2.2 alpontban.) Az előbbi távolság-hozzárendelés szerint:

$$\rho_{E8} = \sum_{i=1}^{a-b} 1 + \sum_{i=1}^b \sqrt{2} = (a - b) + b\sqrt{2},$$

másrészt

$$\rho_E = \sqrt{a^2 + b^2}.$$

Áttérve polárkoordinátákra, megkapható a két metrika aránya (elegendően nagy távolságok esetére):

$$\frac{\rho_{E8}}{\rho_E} = \cos \varphi + (\sqrt{2} - 1) \sin \varphi.$$

Ha az egyenes orientációja a $[0, \pi/4]$ intervallumban egyenlő valószínűséggel vehet fel értékeket, akkor a két metrika arányának átlaga:

$$\frac{\rho_{E8}}{\rho_E} = \frac{4}{\pi} \int_0^{\pi/4} [\cos \varphi + (\sqrt{2} - 1) \sin \varphi] d\varphi = 1,055.$$

Ez a ρ_{E8} metrika célszerű korrekciójára ad lehetőséget:

$$\rho_{E8k} = \frac{1}{1,055} \rho_{E8} = 0,948 \rho_{E8}.$$

A ρ_{E8k} metrika a diszkrét képsikon +2,5% és -5,3% közötti hibával valósítja meg az euklidészi metrikát.

Az (F3-2)-(F3-5) metrikával rendelkező képsíkot *metrikus képsíknak* nevezzük. (Meggjegyezzük, hogy ez a fogalom tetszőleges ponthalmazokra kiterjeszthető, ha ezeken létezik tetszőleges, az I-III, illetve még a IV. feltételt kielégítő távolságfüggvény.)

F3.2.1.1 Diszkrét normák

A norma a távolságfogalom általánosítása tetszőleges dimenziószámú vektorterekre.

A Z halmazt a T számtest (a képfeldolgozásban: \mathbb{R}) feletti *vektortérnek* nevezünk, ha teljesülnek az alábbi feltételek:

- Értelmezve van az *összeadási* művelet és nem vezet ki a halmazból, továbbá kommutatív és asszociatív. Pl. ha $\underline{u}, \underline{v}, \underline{z} \in Z$ akkor $\underline{v} + \underline{w} \in Z$; $\underline{v} + \underline{w} = \underline{w} + \underline{v}$;
 $(\underline{v} + \underline{w}) + \underline{z} = \underline{v} + (\underline{w} + \underline{z})$.
- Létezik *nullvektor*: $\underline{0} \in Z$; bármely $\underline{z} \in Z$ esetén $\underline{z} + \underline{0} = \underline{z}$.
- Minden $\underline{z} \in Z$ vektornak létezik *ellentettje*: $-\underline{z}$; úgy, hogy $\underline{z} + (-\underline{z}) = \underline{0}$.
- Vektor *skalárral való szorzása* nem vezet ki a halmazból, továbbá kommutatív, asszociatív és disztributív. Pl. ha $\underline{v}, \underline{z} \in Z$ és $\alpha, \beta \in \mathbb{R}$, akkor $(\alpha\beta)\underline{z} = \alpha(\beta\underline{z}) = \beta(\alpha\underline{z})$, valamint \mathbb{R} egységelemére $1 \cdot \underline{z} = \underline{z}$; $(\alpha + \beta)\underline{z} = \alpha\underline{z} + \beta\underline{z}$; $\alpha(\underline{v} + \underline{z}) = \alpha\underline{v} + \alpha\underline{z}$.

Felhívjuk a figyelmet arra, hogy ez a vektorfogalom nem egyezik meg pontosan az F1.1 pontban definiált sor-, illetve oszlopvektor figalmával; feltételezi ugyanis egy *ortonormált bázis* létezését, amelynek minden elemére igaz, hogy

$$\underline{u}_i^T \underline{u}_j = \begin{cases} 0, & \text{ha } i \neq j; \\ 1, & \text{ha } i = j; \end{cases} \quad (i, j = 1, 2, \dots, n); \quad (\text{F3-6})$$

ahol az $\underline{u}_i^T \underline{u}_j$ skalárszorzatot az (F1-1) képlet definiálja és n a tér dimenziószáma. (Kissé pontosabban lásd az F4.3.1 alpontot.) Ekkor bármely $\underline{z} \in \mathbb{Z}$ vektor a

$$\underline{z} = \sum_{j=1}^n z^{(j)} \underline{u}_j \quad (\text{F3-7})$$

alakba írható, ahol $z^{(j)} \in \mathbb{R}^n$.

Ezek után a norma a vektortér elemeinek általánosított hosszát jelenti abban az értelemben, hogy a bázisvektorok által meghatározott koordinátatengelyeken az egység *nem* jelent szükségképpen hosszúságegységet. A $\|\cdot\|$ szimbólummal jelölt normafüggvénynek ki kell elégítenie az alábbi feltételeket, minden $\underline{v}, \underline{z} \in \mathbb{Z}$ vektor és $\lambda \in \mathbb{R}$ skalár esetén:

- α) $\|\underline{z}\| \geq 0$; $\|\underline{z}\| = 0$ akkor és csak akkor, ha $\underline{z} = 0$;
- β) $\|\lambda \underline{z}\| = |\lambda| \cdot \|\underline{z}\|$;
- γ) $\|\underline{v} + \underline{z}\| \leq \|\underline{v}\| + \|\underline{z}\|$ (háromszög-egyenlőtlenség).

Ha egy n dimenziós

$$\underline{z}^T = (z^{(1)}, z^{(2)}, \dots, z^{(n)})$$

vektor összetevői valós számok ($z^{(j)} \in \mathbb{R}$, $j = 1, 2, \dots, n$), akkor az \mathbb{R}^n teret (a \underline{z} helyvektorokból álló) vektortérként is felfoghatjuk. Ebben a következő diszkrét normákat értelmezzük:

- euklidészi norma: $\|\underline{z}\|_E = \sqrt{\sum_{j=1}^n z^{(j)2}}$; (F3-8)

(könnyen belátható, hogy ez azonos az (F1-1)-ben definiált $\underline{z}^T \underline{z}$ skalárszorzat négyzetgyökével);

- *abszolútérték* norma: $\| \underline{z} \|_a = \sum_{j=1}^n |z^{(j)}|$; (F3-9)

- *maximum* norma: $\| \underline{z} \|_n = \max_{1 \leq j \leq n} (|z^{(j)}|)$. (F3-10)

Természetesen más normafüggvények is definiálhatók, amelyek eleget tesznek az α - γ) feltételeknek.

Ha létezik normafüggvény, a vektorteret normálnak nevezzük. Vegyük észre, hogy a

$$\rho(\underline{v}, \underline{z}) = \| \underline{v} - \underline{z} \|; \quad (\underline{v}, \underline{z} \in Z)$$

definiációjával a normált teret metrikussá tehetjük; pl. az (F3-8)-(F3-10) normából rendre az (F3-2)-(F3-4) metrikához jutunk.

Megjegyezzük még, hogy könyvünkben további két *folytonos* normadefiniációt is használunk: az egyiket az $I \in \mathbb{R}$ zárt, folytonos intervallum felett értelmezett *folytonos, valós értékű* függvények $C(I)$, a másikat a *négyzetesen integrálható* függvények L_2 halmazára; mindkettőről belátható, hogy a fenti értelemben vektortérnek tekinthető.

Az $f \in C(I)$ függvényekre definiált normafüggvény

$$\| f \|_C = \max\{|f(x)| \mid x \in I\}; \quad (\text{F3-11})$$

az L_2 téré pedig

$$\| f \|_{L_2} = \sqrt{\int_{-\infty}^{+\infty} f^2(x) \, dx}, \quad (x \in \mathbb{R}, f \in L_2) \quad (\text{F3-12})$$

alaku; ez utóbbit az F4.3.1 alponban tárgyaljuk részletebben.

F3.2.2 DISZKRÉT EGYENESEK

Kimutatható [F-6], hogy egy iránykódsorozat akkor és csak akkor reprezentál *racionális meredekségű diszkrét egyenest*, ha

- legfeljebb kétféle iránykód fordul benne elő és ezek mod(8) szomszédosak;
- kétféle iránykód esetében egyikük csak egyszeresen fordul elő (azaz mindkét szomszédja a másik féle iránykód);
- az egyszeres iránykódok sűrűségeloszlása a lehető legegyszerűsebb.

Be lehet bizonyítani, hogy ezek a feltételek általános esetben is elégségesek. (A diszkrét egyenes előállítását részletesen lásd a 3.3.6 alpontban.) Megjegyzendő, hogy a diszkrét egyenes nem feltétlenül egyértelmű (lásd F3-6. ábra).



F3-6. ábra. Diszkrét egyenes többértelmősége

F3.2.3 DISZKRÉT GÖRBÉK

F3.2.3.1 Görbejellemezők

Legyen az $x = f(t)$ és $y = g(t)$ ($0 \leq t \leq 1$) folytonos, paraméteresen megadott görbe. A fontosabb folytonos jellemzők diszkrét megfelelői a következők:

1. *Meredekség*. Folytonos esetben a $P_a(x(t_a), y(t_a))$ pontban definíció szerint:

$$\left. \frac{dy}{dx} = \frac{\frac{dg}{dt}}{\frac{df}{dt}} \right|_{t=t_a}$$

Diszkrét esetben megállapodás szerint minden görbepont-hoz az orientációnak megfelelően utána következő iránykód meredekségét rendeljük. (8)-szomszédság esetén összesen nyolcféle meredekség lehetséges, a meredekség diszkrét esetben ennél "lokálisabban" nem fogalmazható meg. A diszkrét görbék "kvázilokális" meredeksége az adott intervallumra eső lokális meredekségek számtani közepeként értelmezhető.

2. *Görbe ivhossza.* Folytonos esetben, pl. a (t_a) és (t_b) paraméterekkel meghatározott pontok közt definíció szerint:

$$\int_{t_a}^{t_b} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \int_{t_a}^{t_b} \sqrt{\left(\frac{df}{dt}\right)^2 + \left(\frac{dg}{dt}\right)^2} dt.$$

Diszkrét esetben a görbe ivhossza értelemszerűen a szomszédos görbepontok közötti lokális távolságok összegzésével állítható elő. A ρ_{E8} metrika alkalmazása esetén kétféle lokális távolság van ($\sqrt{2}$ és 1). Az összegezés kommutatív tulajdonságának felhasználásával belátható, hogy az n iránykódból álló diszkrét görbe ivhossza $(n+1)$ -féle különböző értéket vehet fel.

3. *Görbület.* Folytonos esetben definíció szerint

$$\left. \frac{d^2y}{dx^2} \left[1 + \left(\frac{dy}{dx}\right)^2 \right]^{-3/2} \right|_{t=t_a} =$$

$$= \left(\frac{df}{dt}\right)^{-2} \left(\frac{df}{dt} \cdot \frac{d^2g}{dt^2} - \frac{dg}{dt} \cdot \frac{d^2f}{dt^2}\right) \left[\left(\frac{df}{dt}\right)^2 + \left(\frac{dg}{dt}\right)^2\right]^{-3/2} \Big|_{t=t_a}$$

A (8)-szomszédságu diszkrét esetben összesen hétféle lokális görbület lehetséges; ui. bármely iránykódvektor után ennyiféle - nem ellentett - iránykódvektor következhet. A diszkrét görbék "kvázilokális" görbülete az adott intervallumra eső lokális görbületek számtani közepeként értelmezhető.

F3.2.3.2 Diszkrét görbék előállítás

Legyen $\{p\}$ a folytonos, $\{\pi\}$ a diszkrét képsík. Feleltessük meg a $g \in \{g_p\}$ folytonos vonalnak a $\gamma \in \{\gamma_\pi\}$ diszkrét alakzatot, ahol $\{g_p\}$ a folytonos, $\{\gamma_\pi\}$ a diszkrét görbék halmaza. Határozzuk meg γ -t a *rdcsmetszéses módszerrel* a következő módon:

- detektáljuk a g vonal és az $x = n$, illetve $y = m$ *hálónonalak* metszéspontjait ($n, m \in \mathbb{N}$; legyenek ezek a $P_i(x_i, y_i)$ pontok ($x_i, y_i \in \mathbb{N}$);
- feleltessük meg a P_i pontoknak a következő $P_i^!(\in \gamma)$ pontokat:

$$P_i(x_i, y_i) \rightarrow P_i^!(x_i^!, y_i^!) \begin{cases} x_i^! = \begin{cases} x_i, & \text{ha } x_i \in \mathbb{N} \\ \text{int}(x_i), & \text{különben} \end{cases} \\ y_i^! = \begin{cases} y_i, & \text{ha } y_i \in \mathbb{N} \\ \text{int}(y_i), & \text{különben} \end{cases} \end{cases} ;$$

ahol $\text{int}(x)$ az x -hez legközelebbi egész számot jelenti.

- az *indifferens* metszéspontoknak (melyekhez az int függvény két koordinátát rendel) feleltessük meg a két legközelebbi rácspontot.

Az eljárás a folytonos görbéknek diszkrét koordinátájú *rácspontok* sorozatát felelteti meg. Könnyen belátható, hogy a $g \in \{g_p\}$ folytonos görbéknek a $\gamma \in \{\gamma_\pi\}$ (8)-utvonal felel meg. (A $g(x) \rightarrow \gamma(x)$ és $g(y) \rightarrow \gamma(y)$ projekció a folytonos koordináta-függvényeket páronként szomszédos egész számok sorozatára képezi le, így a (8)-szomszédság feltételei mindkét koordinátára nézve teljesülnek.) A leírt megfeleltetés könnyen kiterjeszthető térgörbékre, valamint felületekre is. Megjegyzendő, hogy mivel projekció (azaz nem egy-egy-értelmű), nem invertálható (lásd [F-7]).

A $g(x, y) = 0$ folytonos, kétdimenziós görbe diszkrét megfelelőjét a következőképpen generálhatjuk:

Tételezzük fel, hogy a kezdőpont a görbén van, és a lépésköz (a diszkrét képsík rácscsillandója) 1.

Tegyük fel, hogy az (x, y) pontban vagyunk; határozzuk meg, hogy melyik (8)-szomszédja van legközelebb a folytonos görbéhez. Definiáljuk a görbe mentén a

$\pm \underline{v} = \left(\mp \frac{\partial g}{\partial y}, \frac{\partial g}{\partial x} \right)$ irányokat. (Ez azt jelenti, hogy ha pl.

a $+\underline{v}$ irányban határozzuk meg a következő görbepontot, akkor Δx előjelének és a $\frac{\partial g}{\partial y}$ parciális derivált előjelének ellentétesnek kell lennie, míg Δy és $\frac{\partial g}{\partial x}$ parciális derivált előjele azonos.) Ha mindkét parciális derivált folytonos, akkor a görbementi irány rögzítése után következő görbepontként már csak három szomszéd jöhet szóba.

Boole-algebra alkalmazásával ezt a következőképpen fogalmazhatjuk meg:

Legyen:

$$D = \begin{cases} 1, & \text{a } +v \text{ irány esetén;} \\ 0, & \text{a } -v \text{ irány esetén;} \end{cases}$$

ekkor

$$\Delta x = \begin{cases} +1, & \text{ha } \left[\left(\frac{\partial g}{\partial y} \geq 0 \right) \wedge (D = 0) \right] \vee \left[\left(\frac{\partial g}{\partial y} < 0 \right) \wedge (D = 1) \right]; \\ -1, & \text{ha } \left[\left(\frac{\partial g}{\partial y} < 0 \right) \wedge (D = 0) \right] \vee \left[\left(\frac{\partial g}{\partial y} \geq 0 \right) \wedge (D = 1) \right], \end{cases}$$

és

$$\Delta y = \begin{cases} +1, & \text{ha } \left[\left(\frac{\partial g}{\partial x} < 0 \right) \wedge (D = 0) \right] \vee \left[\left(\frac{\partial g}{\partial x} \geq 0 \right) \wedge (D = 1) \right]; \\ -1, & \text{ha } \left[\left(\frac{\partial g}{\partial x} \geq 0 \right) \wedge (D = 0) \right] \vee \left[\left(\frac{\partial g}{\partial x} < 0 \right) \wedge (D = 1) \right]. \end{cases}$$

Tehát következő görbepontként az $(x + \Delta x, y)$, $(x, y + \Delta y)$ és az $(x + \Delta x, y + \Delta y)$ pontok közül kell választani, ahol Δx és Δy értékét a görbén való haladási irány és az első parciális deriváltak egyértelműen meghatározzák. A választás a $g(x, y) = 0$ egyenlet minél pontosabb teljesülése alapján történhet. Vezessük be a következő jelöléseket:

$$g^x = g(x + \Delta x, y),$$

$$g^y = g(x, y + \Delta y),$$

$$g^{xy} = g(x + \Delta x, y + \Delta y).$$

Ekkor a választáshoz elegendő az $|g^x|$, $|g^y|$, $|g^{xy}|$ összehasonlítása:

$$x \leftarrow (x + \Delta x)$$

(azaz lépés x irányban), ha

$$\min(|g^x|, |g^y|, |g^{xy}|) = |g^x|;$$

$$y \leftarrow (y + \Delta y)$$

(azaz lépés y irányban), ha

$$\min(|g^x|, |g^y|, |g^{xy}|) = |g^y|;$$

illetve

$$x \leftarrow (x + \Delta x) \text{ és } y \leftarrow (y + \Delta y)$$

(lépés mindkét irányban), ha

$$\min(|g^x|, |g^y|, |g^{xy}|) = |g^{xy}|.$$

A következő görbepont meghatározása előtt aktualizálni kell a parciális deriváltakat és a függvényértékeket az alábbiak szerint:

$$\frac{\partial g}{\partial x} \leftarrow \frac{\partial g}{\partial x} + \frac{\partial^2 g}{\partial x^2} \Delta x + \frac{\partial^2 g}{\partial x \partial y} \Delta y +$$

$$+ \frac{1}{2} \left(\frac{\partial^3 g}{\partial x^3} (\Delta x)^2 + \frac{2 \partial^3 g}{\partial x^2 \partial y} \Delta x \Delta y + \frac{\partial^3 g}{\partial x \partial y^2} (\Delta y)^2 \right) + \dots,$$

$$\frac{\partial g}{\partial y} \leftarrow \frac{\partial g}{\partial y} + \frac{\partial^2 g}{\partial x \partial y} \Delta x + \frac{\partial^2 g}{\partial y^2} \Delta y +$$

$$+ \frac{1}{2} \left(\frac{\partial^3 g}{\partial x^2 \partial y} (\Delta x)^2 + \frac{2 \partial^3 g}{\partial x \partial y^2} \Delta x \Delta y + \frac{\partial^3 g}{\partial y^3} (\Delta y)^2 \right) + \dots,$$

$$g^x \leftarrow g^\alpha + \frac{\partial g}{\partial x} \Delta x + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (\Delta x)^2 + \dots,$$

$$g^y \leftarrow g^\alpha + \frac{\partial g}{\partial y} \Delta y + \frac{1}{2} \frac{\partial^2 g}{\partial y^2} (\Delta y)^2 + \dots,$$

$$g^{xy} \leftarrow g^\alpha + \frac{\partial g}{\partial x} \Delta x + \frac{\partial g}{\partial y} \Delta y +$$

$$+ \frac{1}{2} \left[\frac{\partial^2 g}{\partial x^2} (\Delta x)^2 + \frac{\partial^2 g}{\partial x \partial y} \Delta x \Delta y + \frac{\partial^2 g}{\partial y^2} (\Delta y)^2 \right] + \dots,$$

ahol g^α az előző választás szerinti (g^x , g^y vagy g^{xy}) függvényt jelöli. Természetesen az összes nem zérus magasabbrendű parciális deriváltat is aktualizálni kell (lásd [F-5]).

F4. FÜGGVÉNYKALKULUS

Rendeljünk hozzá egy ff képben előforduló fénysűrűség-értékeket a $[0, J]$ zárt intervallumhoz (vagyis feleltessük meg a feketének a 0 , a fehérnek a J számot, és a közbűlső értékeket határozzuk meg valamilyen interpolációval). Ekkor a képet leírhatjuk a

$$z = f(x, y)$$

kétváltozós függvénnyel, amelynek értelmezési tartományát a kép méretei határozzák meg, és értékkészlete a $[0, J]$ halmaz. Ezáltal helyettesítettük a képet egy térbeli felülettel és lehetőségünk nyílik arra, hogy alkalmazzuk a matematikai analízis módszereit a képfeldolgozási eljárásokban. Tekintve, hogy a digitális képek csak az (x, y) sík meghatározott pontjaiban vannak értelmezve, és a világosságkódok értékkészlete is diszkrét (egész), a differenciálhányadost differenciahányadossal, az integrálokat (végtelen) sorokkal fogjuk közelíteni, ez azonban nem érinti a módszerek lényegét.

F4.1 LINEÁRIS FÜGGVÉNYOPERÁTOROK

F4.1.1 A DIRAC-DELTA

A képek kényelmes leírása céljából vezessük be a *pontforrás* fogalmát: ez a ff kép egy pontjának "fénysűrűségét" jelenti. (Színes kép esetében a képpont színét 3 monokromatikus (R, G, B) pontforrással írjuk le; az alábbiak könnyen általánosíthatók erre az esetre.) A pontforrásokat egzaktul az *általánosított függvények* (disztribúciók) segítségével lehet leírni. Ezek fogalma jól tükrözi azt a tényt, hogy a valóságban *egyetlen* pont fénysűrűsége *nincs* értelmezve, mivel területe 0. Ehelyett - nem egészen precízen - azt mondhatjuk, hogy az általánosított függvényt minden pontban a pont környezetében vett függvényértékek valamilyen "középpértéke" határozza meg.

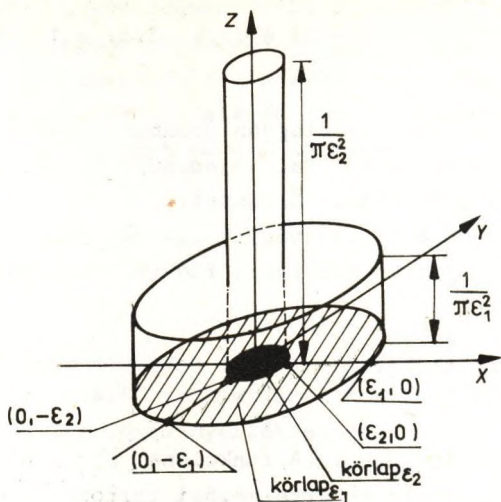
A probléma érzékeltetésére kíséreljük meg először megközelíteni az origóban levő egységnyi pontforrást egységnyi fénysűrűségű képpontok olyan sorozatával, amelyeknek területe 0-hoz tart. Az egyszerűség kedvéért legyenek a képpontok kör alakúak, és ábrázoljuk a fénysűrűségüket a föléljük emelt henger térfogatával (lásd F4-1. ábra). Induljunk ki a "körlap" függvényből, amelynek definíciója

$$\text{körlap}_\epsilon(x, y) = \begin{cases} 1, & \text{ha } \sqrt{x^2 + y^2} \leq \epsilon; \\ 0, & \text{különben.} \end{cases} \quad (\text{F4-1})$$

Ezzel a "fénysűrűség-henger" magasságát a

$$\delta_\epsilon(x, y) = \frac{1}{\pi\epsilon^2} \text{körlap}_\epsilon(x, y), \quad (\text{F4-2a})$$

a képpont fénysűrűségét pedig az



F4-1. ábra. A Dirac-delta közelítése

$$\iint_{-\infty}^{+\infty} \delta_{\varepsilon}(x, y) \, dx dy = \frac{1}{\pi \varepsilon^2} \iint_{\sqrt{x^2+y^2} \leq \varepsilon} dx dy = \frac{1}{\pi \varepsilon^2} \cdot \pi \varepsilon^2 = 1 \quad (\text{F4-2b})$$

képlet írja le. Az (F4-1) és az (F4-2a) definícióból látszik, hogy δ_{ε} értéke 0 az ε sugarú körön kívül, illetve $\frac{1}{\pi \varepsilon^2}$ a kör belsejében; a képpont fénysűrűsége pedig az (F4-2b) szerint mindig 1. Ezekkel a pontforrás "fénysűrűség-hengerének" magasságára a

$$\delta(x, y) = \lim_{\varepsilon \rightarrow 0} \delta_{\varepsilon}(x, y) = \begin{cases} +\infty, & \text{ha } x = 0 \wedge y = 0; \\ 0, & \text{különben;} \end{cases} \quad (\text{F4-3a})$$

értékére pedig formálisan az

$$\iint_{-\infty}^{+\infty} \delta(x, y) \, dx dy = \lim_{\varepsilon \rightarrow 0} \iint_{-\infty}^{+\infty} \delta_{\varepsilon}(x, y) \, dx dy = 1 \quad (\text{F4-3b})$$

eredményt kapjuk. (F4-3a) alapján azonban látható, hogy az (F4-3b) kifejezés bal oldala mindenütt 0, ami *ellentmondásban* van a jobb oldali értékkel.

Tekintsük most a δ_{ε} függvények *gyenge határértékét*, amit tetszőleges $f(x, y)$ folytonos függvény esetében a

$$\lim_{\varepsilon \rightarrow 0} \iint_{-\infty}^{+\infty} \delta_{\varepsilon}(x, y) \cdot f(x, y) \, dx dy \quad (\text{F4-4})$$

funkcionállal definiálunk. (A funkcionál olyan általánosított függvény, amelynek értelmezési tartománya egy függvényosztály; vagyis a függvényosztály egyes elemeihez saját értékészletük egy elemét rendeli hozzá.) Könnyen bebizonyítható (lásd pl. [F-12], 80. old.), hogy (F4-4) értéke $f(0, 0)$; vagyis az (F4-4) funkcionál minden folytonos függvényhez a $(0, 0)$ pontban felvett értékét rendeli. Ezzel a funkcionállal definiáljuk a $\delta(x, y)$ pontforrást, ami nem más, mint a közismert, de matematikailag általában pontatlanul értelmezett *Dirac-delta*:

$$\delta_{\varepsilon}(x, y) \rightarrow \delta(x, y)$$

abban az értelemben, hogy tetszőleges folytonos függvényre

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \iint_{-\infty}^{+\infty} \delta_{\varepsilon}(x, y) \cdot f(x, y) \, dx dy = \\ & = \lim_{\substack{x \rightarrow 0 \\ y \rightarrow 0}} \iint_{-\infty}^{+\infty} \delta(x, y) \cdot f(x, y) \, dx dy = f(0, 0). \end{aligned} \quad (\text{F4-5})$$

Ha $f(x, y)$ a képfüggvény, (F4-5) az origóbeli pontforrást adja, ami éppen az $f(0, 0)$ képfüggvényérték. Ezzel hasznos eszközhöz jutottunk a mintavételezés matematikai leírásához. Ugyanis (F4-5) alapján egy tetszőleges (ξ, η) koordinátájú pontforrást meghatározhatunk a

$$\lim_{\substack{x \rightarrow \xi \\ y \rightarrow \eta}} \iint_{-\infty}^{+\infty} \delta(x - \xi, y - \eta) \cdot f(x, y) \, dx dy = f(\xi, \eta) \quad (\text{F4-6})$$

képlet segítségével. (F4-6) azt fejezi ki, hogy az $f(x, y)$ képet "generáló" pontforrás a képsík (ξ, η) pontjában éppen az $f(\xi, \eta)$ képfüggvényértékkel egyenlő.

F4.1.2 KONVOLUCIÓ

Vegyük észre, hogy az (F4-6) képlet szerint az $e(x, y)$ eredeti képfüggvény felírható pontforrások lineáris összegeként, mivel minden (x, y) pontban

$$e(x, y) = \lim_{\substack{\xi \rightarrow x \\ \eta \rightarrow y}} \iint_{-\infty}^{+\infty} e(\xi, \eta) \cdot \delta(\xi - x, \eta - y) \, d\xi d\eta. \quad (\text{F4-7})$$

Tegyük fel, hogy a leképező rendszer lineáris, és az eredeti képjelet éppen az

$$f(x, y) = T\{e(x, y)\}$$

függvénybe viszi át. Az (F4-7) képlet behelyettesítésével és a linearitás kihasználásával

$$f(x, y) = \iint_{-\infty}^{+\infty} e(\xi, \eta) \cdot \lim_{\substack{\xi \rightarrow x \\ \eta \rightarrow y}} T\{\delta(\xi - x, \eta - y)\} d\xi d\eta.$$

Az integrandusz 2. tényezőjét *pontválasznak* nevezzük, mivel a leképező rendszernek a pontforrásra adott választ írja le. Ez lineáris rendszerekben csak a *koordinátakülönbségektől* függ, ezért vezessük be a

$$g(\xi - x, \eta - y) = \lim_{\substack{\xi \rightarrow x \\ \eta \rightarrow y}} T\{\delta(\xi - x, \eta - y)\} \quad (\text{F4-8})$$

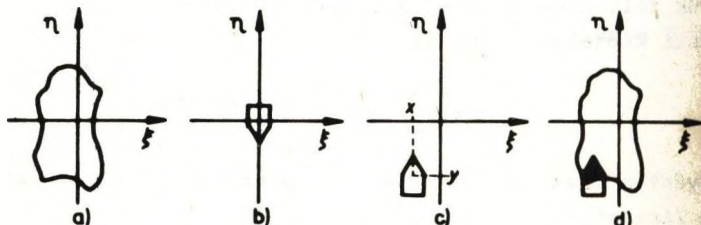
jelölést. A leképezési folyamatot most már leírhatjuk az

$$f(x, y) = \iint_{-\infty}^{+\infty} e(\xi, \eta) \cdot g(\xi - x, \eta - y) d\xi d\eta \quad (\text{F4-9})$$

konvolúcióintegrállal, amelyet a továbbiakban röviden az

$$f(x, y) = e(x, y) \otimes g(x, y) \quad (\text{F4-10})$$

szimbólummal fogunk jelölni.



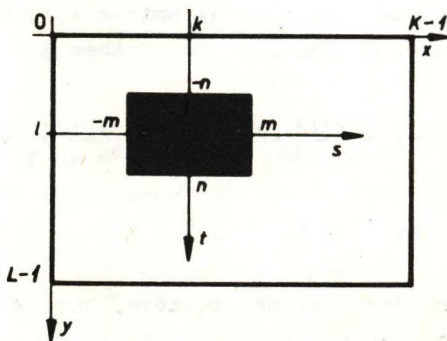
F4-2. ábra. A konvolúcióintegrál szemléltetése

A konvolúcióintegrált az F4-2. ábrán szemléltetjük. Az a) ábrán az eredeti képtartományt látjuk a (ξ, η) síkban, efelett helyezkedik el az $e(\xi, \eta)$ képfüggvény. A b) ábrán mutatjuk meg, hogy a leképező rendszer milyen tartományba képezi le a $(0,0)$ koordinátájú pontforrást. A leképező rendszer úgy működik, hogy minden (x,y) pontot "bejár", s mindegyikben összegzi az eltolt és saját középpontján át tükrözött leképezési tartomány (lásd c) ábra), valamint az eredeti képtartomány közös része feletti (lásd d) ábra) pontforrásokra adott pontválaszait.

A konvolúcióintegrál digitális képek esetében *véges összegzésre* módosul. Ekkor ugyanis a $q(k, l)$ -lel jelölt képfüggvény csak a véges, $K \times L$ méretű diszkrét képsíkon van értelmezve (lásd F3. fejezet), vagyis

$$q(k, l) = 0, \quad \text{ha } k \geq K, \text{ ill. } l \geq L. \quad (\text{F4-11})$$

Az (F4-8) összefüggéssel definiált pontválaszról ugyszintén feltesszük, hogy csak egy véges, téglalap alakú tartomány pontjaiban különbözik nullától. Más szóval megadható egy $T(s, t)$ mátrixszal, amelyet a továbbiakban *konvolúciós szűrőnek* nevezünk és elemeit *súlytényezőkként* értelmezzük (lásd F4-3. ábra). Az ábra jelöléseivel az



F4-3. ábra. Konvolúciós szűrő illesztése egy pontra

(F4-9)-beli konvolúcióintegrál helyébe ekkor az

$$r(k, l) = \sum_{s=-m}^{+m} \sum_{t=-n}^{+n} q(k+s, l+t) \cdot T(s, t) \quad (\text{F4-12})$$

képlettel definiált *konvolúcióösszeg* lép, ahol $r(k, l)$ -lel az átalakított ("megszűrt") digitális képet jelöltük. Szemléletesen ez azt jelenti, hogy egy képpontban a képfüggvény értékét csak a $(2m+1) \cdot (2n+1)$ -es környezetében levő pontforrások befolyásolják, mégpedig éppen a konvolúciós szűrőben szereplő súlyoknak megfelelő mértékben.

A konvolúciós összegzést széles körben alkalmazzák a különböző képszűrési eljárásokban.

F4.2 DIFFERENCIÁLOPERÁTOROK

F4.2.1 A GRADIENS ÉS A LAPLACE-OPERÁTOR

A függvényanalízis számos differenciáloperátort (rotáció, divergencia stb.) használ a függvények tulajdonságainak vizsgálatára, ezek közül a képfeldolgozásban főleg a *gradiensnek* van jelentősége, amit a ∇ (nabla) szimbólummal jelölünk, és kétváltozós esetben a

$$\begin{aligned} \nabla f(x, y) &= \frac{\partial f(x, y)}{\partial x} \underline{e}_x + \frac{\partial f(x, y)}{\partial y} \underline{e}_y = \\ &= f'_x \underline{e}_x + f'_y \underline{e}_y \end{aligned} \quad (\text{F4-13})$$

összefüggéssel definiálunk; feltéve, hogy a parciális differenciálhányadosok léteznek. A képletben \underline{e}_x , illetve

$\frac{e}{y}$ az x , illetve y irányu egységvektort, f'_x , f'_y a megfelelő irányu parciális differenciálhányadosot jelenti.

(Az (x, y) argumentumot a rövidség kedvéért elhagytuk.)

A gradiens vektor, amelynek nagyságát az (F3-8) szerinti

$$\|\nabla f\| = \sqrt{f'_x{}^2 + f'_y{}^2} \quad (\text{F4-14})$$

euklidészi norma; az x tengellyel bezárt szögét pedig a

$$\operatorname{tg} \alpha = \frac{f'_y}{f'_x},$$

illetve a

$$\cos \alpha = \frac{f'_x}{\|\nabla f\|} \quad \text{vagy} \quad \sin \alpha = \frac{f'_y}{\|\nabla f\|} \quad (\text{F4-15})$$

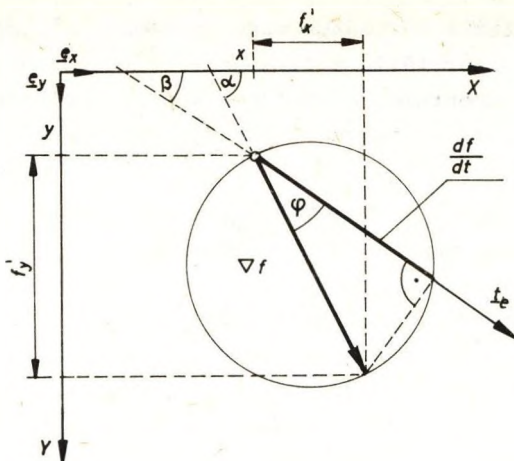
képletből lehet kiszámítani.

Ismeretes, hogy az $f(x, y)$ függvénynek egy tetszőleges \underline{t}_e egységvektorral kijelölt irány szerinti differenciálhányadosát a

$$\frac{df(x, y)}{dt} = \lim_{dt \rightarrow 0} \frac{f(x + dt \cos \beta, y + dt \sin \beta) - f(x, y)}{dt}$$

határérték definiálja (lásd F4-4. ábra). Ha figyelembe vesszük, hogy $dx = dt \cos \beta$, illetve $dy = dt \sin \beta$, akkor az összetett függvénynek differenciálási szabálya szerint:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} = f'_x \cos \beta + f'_y \sin \beta. \quad (\text{F4-16})$$



F4-4. ábra. Iránymenti differenciálhányados szemléltetése

Vegyük észre, hogy az (F4-16) differenciálhányados nagyságát kifejezhetjük a ∇f és a \underline{t}_e skalárszorzatával

$$\frac{df}{dt} = \nabla f \cdot \underline{t}_e = \|\nabla f\| \cos \varphi \quad (\text{F4-17})$$

ahol φ a két vektor által bezárt szög (lásd az F4-4. ábrán). Ez az érték akkor a legnagyobb, ha a két vektor egyirányú ($\beta = \alpha$, vagyis $\varphi = 0$).

Helyettesítsük be (F4-16)-ba (F4-14) és (F4-15) alapján a gradiens iránytényezőit, ekkor megkapjuk a *gradiens irányába eső* differenciálhányados nagyságát:

$$\frac{df}{dg} = f'_x \frac{f'_x}{\sqrt{f_x'^2 + f_y'^2}} + f'_y \frac{f'_y}{\sqrt{f_x'^2 + f_y'^2}} = \sqrt{f_x'^2 + f_y'^2} = \|\nabla f\|, \quad (\text{F4-18})$$

s ez megegyezik (F4-17)-tel a $\varphi = 0$ esetben.

A mondottakból világos, hogy a gradiens jelöli ki a

képfüggvény egy adott pontjában a *legnagyobb változás* irányát, s ennek értékét a *gradiens nagysága* adja. (Szemléletesen szólva: a gradiens iránya megmondja, hogy a képfüggvénnyel meghatározott "terepen" merre van a legmeredekebb "kapaszkodó", vagyis merre tart a "csúcshoz" vezető legrövidebb út; nagysága pedig arról tájékoztat, hogy az egységnyi távolságon mekkora a "szintkülönbség", vagyis milyen meredek az út.)

A másodrendű parciális differenciálhányadosok a szélsőérték helyek meghatározásában játszanak szerepet. A képfeldolgozásban elterjedt a

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} = f''_{xx} + f''_{yy}, \quad (\text{F4-19})$$

illetve a

$$\begin{aligned} \nabla_{\underline{t}}^2 f(x, y) &= \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} + 2 \frac{\partial^2 f(x, y)}{\partial x \partial y} = \\ &= f''_{xx} + f''_{yy} + 2 f''_{xy} \end{aligned} \quad (\text{F4-20})$$

Laplace-operátor használata (a vegyes parciális differenciálhányadosok nélkül vagy azokkal) az élkiemelési és az élkitűzési eljárásokban.

F4.2.2 DISZKRÉT DIFFERENCIAOPERÁTOROK

Digitális képek esetében a differenciálást *differenciálhányadosok* képzésével valósítjuk meg, ezeknek azonban többféle értelmezésük lehetséges. Így pl.:

$$\frac{\partial f(x, y)}{\partial x} \rightarrow \Delta_x(q(k, l)) = q(k, l) - q(k-1, l); \quad (F4-21)$$

illetve

$$\begin{aligned} \frac{\partial^2 f(x, y)}{\partial x^2} \rightarrow \Delta_{xx}^2(q(k, l)) &= \Delta_x(\Delta_x(q(k+1, l))) = \\ &= (q(k+1, l) - q(k, l)) - (q(k, l) - q(k-1, l)) = \\ &= q(k+1) + q(k-1, l) - 2q(k, l). \quad (F4-22) \end{aligned}$$

Ezt az értelmezést az F4-5/a és F4-5/b ábrán szemléltetjük. Eszerint az (F4-13)-ban definiált gradiensvektor digitális megfelelőjének nagyságát az (F4-14) helyett a

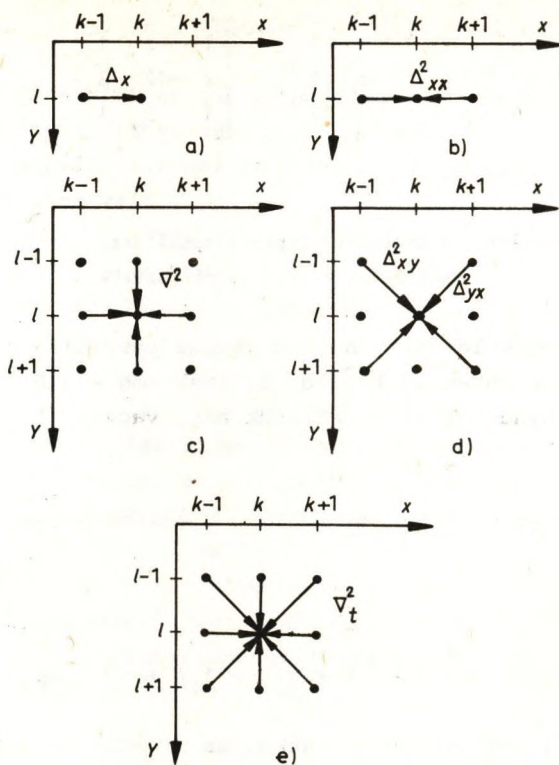
$$\|\nabla q(k, l)\| = \sqrt{\Delta_x^2(q(k, l)) + \Delta_y^2(q(k, l))} \quad (F4-23)$$

képlettel számolhatjuk ki. Mivel a négyzetgyökvonás meg lehetõsen "drága" mûvelet, a képfeldolgozásban a *digitális gradiens nagyságát* inkább az (F3-9)-cel definiált

$$\|\nabla q(k, l)\| = |\Delta_x(q(k, l))| + |\Delta_y(q(k, l))| \quad (F4-23/a)$$

abszolútérték-norma szerint számítjuk ki.

Megjegyezzük, hogy mint ismeretes - a gradiensvektor bármely két, egymásra merõleges irányu differenciálhányadossal is kifejezhetõ. Ezt kihasználva számos módszer alakult ki a digitális gradiens közelítésére; ezek az anyagban foglalkozunk részletesen. Egyszerû számitással kimutatható, hogy (F4-22) alapján az (F4-19)-beli Laplace-operátor digitális megfelelõje a



F4-5. ábra. Differenciahányadosok szemléltetése

$$\begin{aligned}
 \nabla^2 q(k, l) &= \Delta_{xx}^2 (q(k, l)) + \Delta_{yy}^2 (q(k, l)) = \\
 &= q(k+1, l) + q(k, l+1) + q(k-1, l) + q(k, l-1) - 4q(k, l)
 \end{aligned}
 \tag{F4-24}$$

kifejezés, amelyet az F4-5/c ábrán szemléltetünk, és az F4-6/a ábrán megadott konvolúciós szűrővel számíthatunk ki az (F4-12) képlet szerint.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

a)

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

b)

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

c)

F4-6. ábra. A digitális Laplace-operátort megvalósító konvolúciós szűrők

Kézenfekvőnek látszik, hogy a vegyes parciális differenciálhányadosokat az F4-5/d ábrának megfelelő "ferde" differenciahányadosokkal közelítsük meg, vagyis (F4-22) mintájára a

$$\Delta_{xy}^2(q(k, l)) = q(k + 1, l + 1) + q(k - 1, l - 1) - 2q(k, l)$$

illetve a

$$\Delta_{yx}^2(q(k, l)) = q(k + 1, l - 1) + q(k - 1, l + 1) - 2q(k, l)$$

kifejezésekkel definiáljuk. Ezekkel az (F4-20)-nak megfelelő digitális Laplace-operátort a

$$\begin{aligned} \nabla_{\tau}^2 q(k, l) &= q(k + 1, l) + q(k + 1, l - 1) + q(k, l - 1) + \\ &+ q(k - 1, l - 1) + q(k - 1, l) + q(k - 1, l + 1) + \\ &+ q(k, l + 1) + q(k + 1, l + 1) - 8q(k, l) \end{aligned} \quad (\text{F4-25})$$

képlet adja (lásd F4-5/e és F4-6/b ábra).

Megjegyezzük, hogy bár a Laplace-operátorból csak közvetve származtatható le, néha az F4-6/c ábra szerinti konvolúciós szűrőt is használják az (F4-25) kifejezés helyett.

F4.3 FÜGGVÉNYTRANSZFORMÁCIÓK

A gyakorlatban előforduló képfüggvényeket analitikus formában általában nem lehet felírni, matematikai kezelésük nehézkes. A probléma megoldására kedvelt módszer a sorbafejtés valamilyen jól kezelhető és/vagy könnyen programozható *bázisfüggvények* szerint.

Egy $f(x,y)$ függvényt *felbonthatónak* nevezünk, ha a sík egy zárt, derékszögű A tartománya ($a \leq x \leq b$, $c \leq y \leq d$) felett teljesülnek rá a következő feltételek:

1. *szakaszonként folytonos* (ez azt jelenti, hogy vagy folytonos, vagy majdnem mindenütt (=véges, illetve megszámlálhatóan végtelen számú hely kivételével) folytonos, és a szakadási pontokban minden irányban léteznek az egyoldalu határértékei; pl. $f(x + 0, y)$ stb.);
2. *korlátos* (ez következik 1.-ből);
3. *négyzetesen integrálható*, vagyis $f(x,y)$ mérhető (ez a szakaszonként folytonos függvényekre automatikusan teljesül) és

$$\int_{-\infty}^{+\infty} [f(x,y)]^2 dx dy < \infty.$$

Megjegyezzük, hogy ha az A tartomány véges mértékű (mint esetünkben is), 3.-ból $f(x,y)$ integrálhatósága is következik, ez azonban nem szükséges feltétel.

A képfüggvények mindig eleget tesznek ezeknek a feltételeknek, sőt az A tartományon (a képsíkon) kívül értékük mindenütt 0; ezt a digitalizálásnál használjuk ki.

A függvényt sor együtthatóinak meghatározását szokás *transzformációnak*, az eredményül kapott függvényt pedig *transzformált függvénynek* nevezni. Nyilvánvaló, hogy

az eljárás akkor vezet célhoz, ha a felbontás *egyértelmű és invertálható*; ez a követelmény bizonyos korlátozásokat ró a bázisfüggvényekre. Anélkül, hogy részletekbe mennénk, összefoglaljuk a főbb ismerveket (lásd [F-1]).

F4.3.1 ORTOGONÁLIS FÜGGVÉNYRENDSZEREK

A négyzetesen integrálható, véges vagy végtelen

$$\{\Psi_n\} = \Psi_0, \Psi_1, \Psi_2, \dots, \Psi_n, \dots$$

függvénysorozatról akkor mondjuk, hogy *ortogonális függvényrendszer* alkot, ha a sík egy *véges mértékű E tartománya* felett a *skalárszorzatokra* teljesül, hogy

$$\iint_E \Psi_i(x, y) \cdot \Psi_j(x, y) \, dx dy = \begin{cases} N_i \neq 0, & \text{ha } i = j; \\ 0, & \text{különben.} \end{cases} \quad (\text{F4-26})$$

(Emlékeztetünk rá, hogy két vektor skalárszorzata akkor nulla, ha merőlegesek egymásra; (F4-26) ennek általánosítása a függvényekre.) Az $i = j$ esetben a $\sqrt{N_i}$ számot a Ψ_i függvény *normájának* nevezzük és $\|\Psi_i\|$ -val jelöljük (lásd F3.2.1.1 alpont). Képezzük a

$$\{\varphi_n\} = \frac{\Psi_n}{\sqrt{N_n}}$$

függvénysorozatot: ez *ortonormált* lesz, mivel nyilván teljesül rá az (F4-26) feltétel, és ezen kívül

$$\|\varphi_i\| = \sqrt{\iint_E \varphi_i^2(x, y) \, dx dy} = \sqrt{\frac{1}{N_i} \iint_E \Psi^2(x, y) \, dx dy} = 1.$$

Egy (ortonormált) függvényrendszert *zártnak* nevezünk, ha tovább nem egészíthető ki. Más szóval: ha $g(x, y)$ négyzetesen integrálható függvény, és

$$\iint_E \varphi_i(x, y) \cdot g(x, y) \, dx dy = 0$$

minden i -re, akkor majdnem mindenütt $g(x, y) = 0$. Vagyis a $\{\varphi_n\}$ sorozat a négyzetesen integrálható függvények (L_2) vektorterének ortonormált bázisát képezi. A felbontáshoz az egyes φ_i komponensek *együtthatóját* az

$$a_i = \iint_A f(x, y) \cdot \varphi_i(x, y) \, dx dy$$

integrállal kell meghatározni. ($A \subseteq E$ kötelező!) Ezekkel

$$f_m(x, y) = \sum_{i=0}^m a_i \cdot \varphi_i \quad (\text{F4-27})$$

és

$$f(x, y) \sim f_m(x, y). \quad (\text{F4-28})$$

(F4-28)-at úgy olvassuk, hogy f_m *átlagosan konvergál* f -hez, s ezen azt értjük, hogy a *négyzetes eltérése* nullához tart:

$$\lim_{m \rightarrow \infty} \iint_A [f(x, y) - f_m(x, y)]^2 \, dx dy = 0. \quad (\text{F4-29})$$

Ezek után kimondhatjuk a következő tételt:

Zárt, ortonormált függvényrendszer szerint *bármely*, a sik egy zárt A tartománya felett felbontható $f(x, y)$ függvényt *egyértelműen és invertálhatóan* (F4-27) alakú sorba lehet fejteni. A sorfejtés (F4-29) értelmében átlagosan konvergál az $f(x, y)$ -hoz, és adott m esetében a négyzetes eltérés minimális minden más típusú sorbafejtéshez képest.

Egyidejűleg teljesül a

$$\sum_{i=0}^{\infty} a_i^2 \leq \iint_A [f(x,y)]^2 dx dy \quad (F4-30)$$

(Bessel-féle) egyenlőtlenség. Bebizonyítható, hogy (F4-29) konvergenciájához szükséges és elegendő, ha (F4-30)-ban az egyenlőség érvényes.

F4.3.2 FOURIER-TRANSZFORMÁCIÓ

Ismeretes (lásd pl. [F-1]), hogy a $\{\sin nx\}$ és a $\{\cos nx\}$ függvénysorozat ortogonális és zárt a $\{0, 2\pi\}$ intervallumon, ezen alapul a *Fourier-transzformáció*. Nagy számításigénye miatt azonban nehézkes az alkalmazása, ezért egy sereg más, "olcsóbban" futtatható transzformációt is kidolgoztak, ezek közül kettőt ismertetünk a következő alpontokban.

A képfüggvényeket könyvünkben mindenütt kisbetűvel, a transzformált függvényeket a megfelelő nagybetűvel jelöljük; az előbbiekről mindig feltesszük, hogy felbontható az F4.3.1 alpont értelmében.

Egy $f(x,y)$ függvény Fourier-transzformáltját vagy - általánosan - *frekvenciaspektrumát* az

$$F(\omega_x, \omega_y) = \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} f(x,y) \exp[-i2\pi(\omega_x x + \omega_y y)] dx dy \quad (F4-31)$$

képlet definiálja. Itt és a továbbiakban \exp az e -alapu hatványfüggvényt, $i = \sqrt{-1}$ a képzetes egységet jelöli. (Az, hogy a határok a végtelenbe vannak kitolva, nem zavaró, ha az A tartományon kívül $f(x,y)$ -t mindenütt nullának tekintjük.)

Az (F4-31)-ből az eredeti függvényt az

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(\omega_x, \omega_y) \exp[i2\pi(\omega_x x + \omega_y y)] d\omega_x d\omega_y \quad (\text{F4-32})$$

szerinti inverz Fourier-transzformációval kaphatjuk meg.
Ismeretes, hogy

$$\exp[i2\pi(\omega_x x + \omega_y y)] = \cos 2\pi(\omega_x x + \omega_y y) + i \sin 2\pi(\omega_x x + \omega_y y).$$

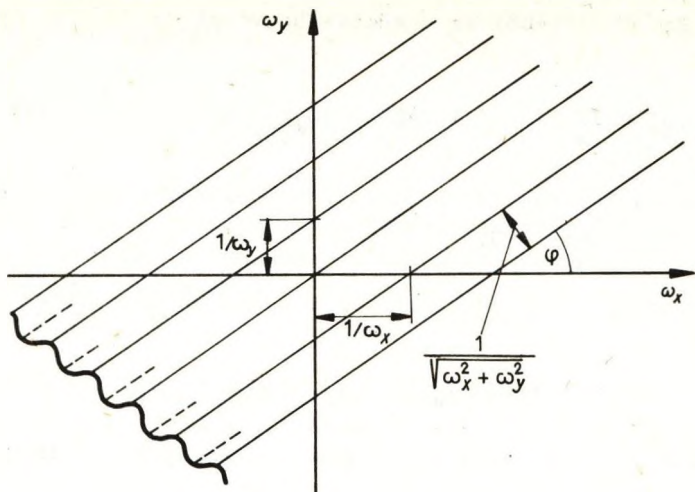
Tekintsük pl. a $\cos(\omega_x x + \omega_y y)$ tagot: ez egy síkhullámot ír le az (x, y) síkban, amelynek x , illetve y irányu frekvenciája ω_x , illetve ω_y , vagyis a megfelelő hullámhosszak ω_x^{-1} , illetve ω_y^{-1} (lásd F4-7. ábra). A hullámfront az x tengellyel

$$\varphi = \arctg\left(\frac{\omega_x}{\omega_y}\right)$$

szöget zár be, s az erre merőleges irányban a síkfrekvencia $\sqrt{\omega_x^2 + \omega_y^2}$. Ugyanez érvényes értelemszerűen a szinusz-hullámu összetevőre is.

Ezek szerint az (F4-31)-gyel definiált $F(\omega_x, \omega_y)$ nem más, mint az $f(x, y)$ -nak trigonometrikus függvények szerinti felbontásában a $\sqrt{\omega_x^2 + \omega_y^2}$ síkfrekvenciájú tag együtthatója: azt mondja meg, hogy az adott síkhullám mekkora amplitudóval szerepel a felbontásban. Az $1/4\pi^2$ szorzóra azért van szükség, mert a trigonometrikus függvények nem normáltak.

A digitális képek felbontásánál a helyzet egy kicsit más, mivel a képfüggvény csak a diszkrét (k, l) ($k, l = 0, \pm 1, \pm 2, \dots$) pontokban van értelmezve. Ennek megfelelően frekvenciaspektruma sem folytonos, hanem csak a diszkrét



F4-7. ábra. Koszinusz típusú síkhullám

(u, v) ($u, v = 0, \pm 1, \pm 2, \dots$) frekvenciák fordulnak benne elő [F-4].

Jelölje a felbontandó kép méretét x , illetve y irányban K , illetve L , akkor (F4-31) a következőképpen módosul:

$$F(u, v) = \frac{1}{KL} \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} f(k, l) \exp[-i2\pi(\frac{ku}{K} + \frac{lv}{L})] \quad (\text{F4-33})$$

Ezekkel az együtthatókkal a felbontás

$$f(k, l) = \sum_{u=0}^{K-1} \sum_{v=0}^{L-1} F(u, v) \exp[i2\pi(\frac{ku}{K} + \frac{lv}{L})]. \quad (\text{F4-34})$$

Fontos megjegyezni, hogy ebben az esetben az $F(u, v)$ együtthatóju

$$\cos\left[2\pi\left(\frac{ku}{K} + \frac{lv}{L}\right)\right], \quad \sin\left[2\pi\left(\frac{ku}{K} + \frac{lv}{L}\right)\right]$$

bázisfüggvények csak *durva közelítései* a folytonos \cos , illetve \sin függvénynek: a magasabb frekvenciákon négy-szöghullámba mennek át.

Ennek érzékeltetésére megadjuk az egydimenziós

$$\cos\left(2\pi\frac{ku}{K}\right), \quad \sin\left(2\pi\frac{ku}{K}\right)$$

bázisfüggvények alakját $K = 16$ esetre. (Lásd F4-8. ábra; a valóságban az F4-7. ábrához hasonló síkhullámmal törté-
nik a közelítés, amelynek ez a "keresztmetszete" a hul-
lámfrontra merőleges irányban.)

F4.3.3 HADAMARD-TRANSZFORMÁCIÓ

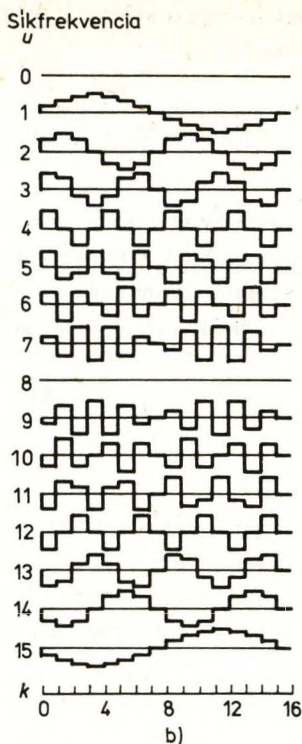
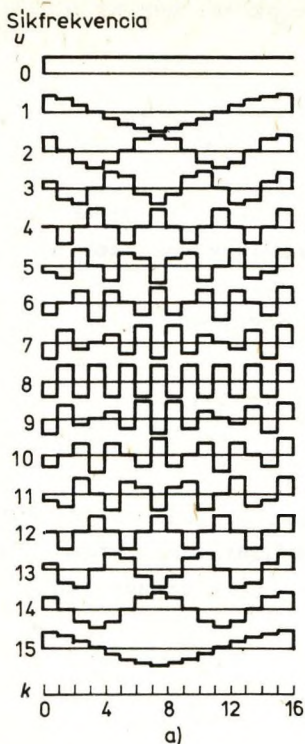
A Hadamard-transzformáció a normalizált Hadamard-mát-
rixokon alapul. Ezek ± 1 elemekből álló *négyzetes* mátrixok,
és érvényes a

$$\underline{H} \cdot \underline{H}^T = \underline{E}$$

összefüggés (lásd F1.2.1 alpont)

A legkisebb ortonomált Hadamard-mátrix másodrendű:

$$\underline{H}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$



F4-8. ábra. Egydimenziós, digitális Fourier-bázisfüggvények

A transzformációhoz csak az $N = 2^n$ ($n = 0, 1, 2, 3, \dots$) rendű mátrixokat használják; ezek rekurzívan építhetők fel a

$$\underline{H}_{2N} = \frac{1}{\sqrt{2}} \begin{bmatrix} \underline{H}_N & \underline{H}_N \\ \underline{H}_N & -\underline{H}_N \end{bmatrix}$$

képlet szerint. Így pl.:

$$\underline{H}_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix};$$

$$\underline{H}_8 = \frac{\sqrt{2}}{4} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

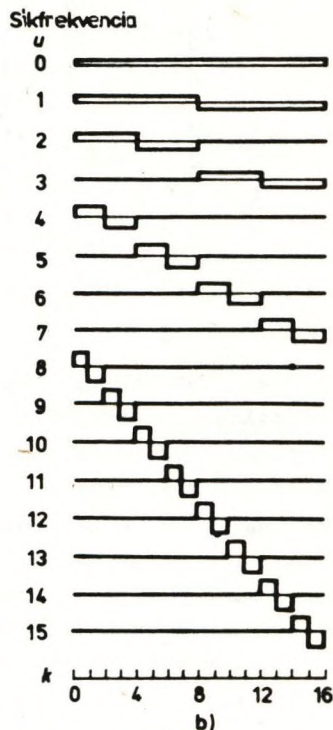
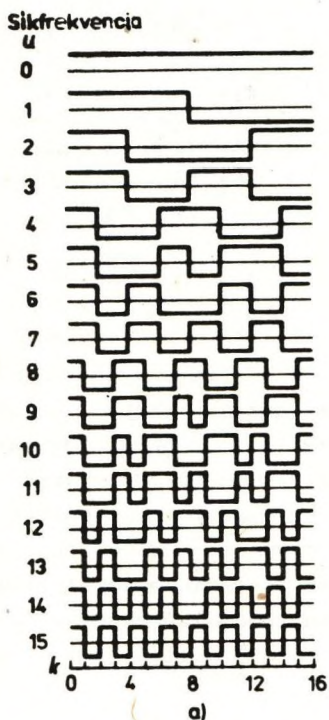
stb.

A K -adrendű Hadamard-mátrixok sorait $u = 0, 1, \dots, K-1$ frekvenciájú négyszög hullámokkal ábrázolhatjuk, ezeket Walsh-függvényeknek nevezik. Eszerint az Hadamard-transzformáció nem más, mint egy függvény felbontása Walsh-függvények szerint. A gyakorlatban - a bemutatott helyett - a mátrixok, illetve Walsh-függvények "rendezett" sorozatát használják, amelyekben a frekvencia (ami megfelel a mátrixok egyes soraiban az előjelváltások számának) soronként nő. Legyenek a képméretetek $K = 2^m$, illetve $L = 2^n$; ezekkel a sorfejtés együtthatóit az

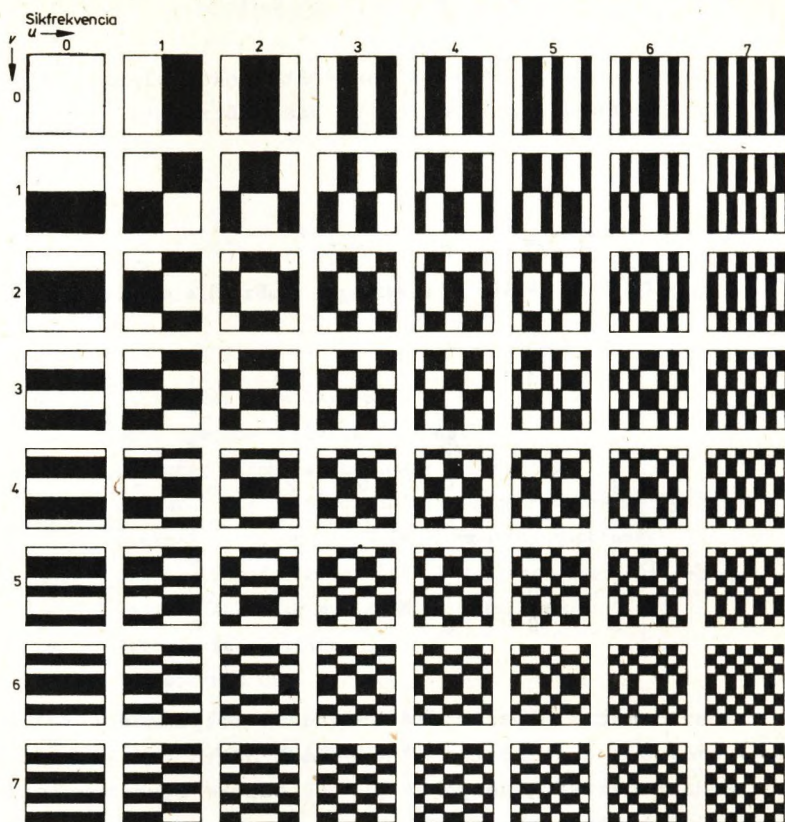
$$F(u, v) = \frac{1}{\sqrt{KL}} \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} f(k, l) (-1)^{c(k, l, u, v)},$$

$$c(k, l, u, v) = \sum_{i=0}^{m-1} (u_{m-i} - u_{m-i-1}) k_i + \sum_{i=0}^{n-1} (v_{n-i} - v_{n-i-1}) l_i$$

képlet adja, ahol u_i, v_i, k_i , illetve l_i az u, v, k , illetve l bináris számokban a 2^i helyiértékű bitet jelöli. A Walsh-bázisfüggvényeket $K = 16$ esetre egydimenziósan az F4-9/a ábrán, $K = 8$ esetre kétdimenziósan az F4-10. ábrán mutatjuk be.



F4-9. ábra. Egydimenziós Walsh- (a), illetve Haar- (b) bázisfüggvények



F4-10. ábra. Kétdimenziós Walsh-bázisfüggvények (fehér $\rightarrow +1$;
 fekete $\rightarrow -1$; vö. F4-9/a ábra)

F4.3.4 HAAR-TRANSZFORMÁCIÓ

A Haar-transzformáció a Haar-mátrixokon alapul. Példaképpen megadjuk az \underline{I}_4 és az \underline{I}_8 Haar-mátrixot:

$$\underline{I}_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} ;$$

$$\underline{I}_8 = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$

A transzformáció olyan mintavételezési folyamatnak felel meg, amelyben a felbontás finomsága 2 hatványai szerint nő. A Haar-bázisfüggvényeket $K = 16$ esetre egydimenziósan az F4-9/b ábrán adtuk meg.

További részletek, illetve transzformációk [F-1]-ben találhatóak.

IRODALOMJEGYZÉK

A gyakrabban előforduló forrásokra az alábbi rövidítéseket használjuk:

AI	Artificial Intelligence
BSTJ	Bell Systems Technical Journal
CGIP	Computer Graphics and Image Processing
CVGIP	Computer Vision Graphics and Image Processing
JACM	Journal of Association for Computing Machinery
IEEE	Institution for Electrical and Electronic Engineering
PAMI	Pattern Analysis and Machine Intelligence
PR	Pattern Recognition
Proc.	Proceedings of ...
SMC	Systems Man and Cybernetics
TC	Transactions on Computers

ÁLTALÁNOS, ÖSSZEFOGLALÓ MŰVEK

- A-1. ÁLLÓ G. - FÜGLEIN J. - HEGEDŰS Gy. Cs. - SZABÓ J.:
Bevezetés a számítógépes képfeldolgozásha. BME Mérnöki
Továbbképző Intézet, R.sz. 5260. Budapest, 1985.
- A-2. ANDERBERG, M.R.: Cluster analysis for applications.
Academic Press, New York, 1973.
- A-3. ANDREWS, H.C.: Computer Techniques in image processing.
Academic Press, New York, 1970.

- A-4. DUDA, R.O. - HART, P.E.: Pattern classification and scene analysis. Wiley-Interscience, New York, 1973.
- A-5. FU, K.S.: Syntactic methods in pattern recognition. Academic Press, New York, 1974.
- A-6. FU, K.S. (ed): Digital pattern recognition. Springer-Verlag, Berlin, 1976.
- A-7. FU, K.S. - YU, T.S.: Statistical pattern classification using contextual information. Research Studies Press, London, 1980.
- A-8. FUKUNAGA, K.: Introduction to statistical pattern recognition. Academic Press, New York, 1972.
- A-9. GONZALEZ, R.G. - WINTZ, P.: Digital image processing. Addison-Wesley, New York, 1977.
- A-10. JAROSLAWSKI, L.P.: Einführung in die digitale Bildverarbeitung VEB Deutscher Verlag der Wissenschaften, Berlin, 1985.
- A-11. NEWMAN, W.M. - SPROULL, R.P.: Interaktiv számítógépes grafika. Műszaki Könyvkiadó, Budapest, 1985.
- A-12. PAVLIDIS, J.K.: Algorithms for graphics and image processing. Computer Science Press, Rockville (USA), 1982.
- A-13. PRATT, W.K.: Digital image processing. Wiley-Interscience, New York, 1978.
- A-14. ROSENFELD, A. - KAK, A.C.: Digital picture processing. Academic Press, New York, 1976.
- A-15. SERRA, J.: Image analysis and mathematical morphology. Academic Press, New York, 1982.
- A-16. YOUNG, T.Y. - CALVERT, T.W.: Classification, estimation and pattern recognition. Elsevier, New York, 1974,
- A-17. Ярославский, Л.П.: Введение в цифровую обработку изображений. Советское Радио, Москва, 1979.

HIVATKOZÁSOK

- 1-1. ÁLLÓ G.: Mikrogépes vezérlésű színes képfeldolgozó berendezés. *Információ Elektronika* 1982/2, 64-68.
- 1-2. ÁLLÓ G.: A látás automatizálása 1. *Információ Elektronika* 1984/5, 285-294.
- 1-3. ÁLLÓ G.: A látás automatizálása 2. *Információ Elektronika* 1985/1, 42-52.
- 1-4. ÁLLÓ, G. - STASZNY, G. - FERŐ, L.: Colour Picture Converter for image processing. *Microprocessing and Microprogramming* 19/1 (1987), 27-35.
- 1-5. BINFORD, T.O.: Survey of model-based image analysis systems. *International Journal of Robotic Research* 1/1 (1982) 18-64.
- 1-6. BRACEWELL, R.N.: Two-dimensional aerial smoothing in radio-astronomy. *Australian Journal of Physics* 9 (1956) 297-314.
- 1-7. BRACKEN, P.A. et al.: Remote sensing software systems. In: Colwell, R.N. (ed): *Manual of Remote Sensing*; American Society of Photogrammetry. The Sheridan Press, New York, 1983, 807-839.
- 1-8. BÜTTNER Gy.: Képiró-olvasó (Colormation) berendezés alkalmazási lehetőségei a távérzékelt adatok feldolgozásában. *Geodézia és Kartográfia*, 1981/3, 193-199.
- 1-9. Colour display HT680X-CD. *Hiradástechnika Szövetkezet dokumentáció*, Budapest, 1983.
- 1-10. DAVIDSON, M.L.: Perturbation approach to spatial brightness interaction in human vision. *Journal of Optical Society of America* 58/9 (1968) 1300-1308.
- 1-11. DÉNES J.: Megvalósult újszerű ember-gép kapcsolatok az SZKI-ban. *Információ Elektronika* 1983/5, 257-262.
- 1-12. DOMOKOS Gy.-né: *Távérzékeltetés a műszaki gyakorlatban*. Műszaki Könyvkiadó, Budapest, 1984.
- 1-13. ENDRÓDI B. - SZENES Zs.: Berendezés televíziós képinformáció tárolására számítógépes hozzáférésű memória felhasználásával. SZKI Szolgálati találmány, Nr. 180133. Budapest, 1980.

- 1-14. Format specifications for LANDSAT MSS System Corrected Computer Compatible Tape. Telespazio, Roma, 1976.
- 1-15. GERÉB Gy.: Pszichológia. Tankönyvkiadó, Budapest, 1976.
- 1-16. van GOOL L. - DEWALE P. - OOSTERLINCK A.: Survey: Texture analysis anno 1983. CVGIP 29 (1985) 336-357.
- 1-17. GULYÁS O.: A folytonos paraméterű folyamatok diszkretizálásának módszerei. In: Tusnády G. - Ziermann M. (szerk.): Idősorok analízise. Műszaki Könyvkiadó, Budapest, 1986, 119-139.
- 1-18. HAUSKA, H.: The Lulea image processing system (LIPS). Proc. Machine Processing of Remotely Sensed Data Symposium, 1980, 40-46.
- 1-19. HECHT, S.: The visual discrimination of intensity and the Weber-Fechner law. Journal of General Psychology 7 (1924) 241.
- 1-20. van der HORST, G.J.C. - de WEERT, C.M. - BOUMAN, M.A.: Transfer of spatial chromaticity - contrast at threshold in the human eye. Journal of Optical Society of America 57/10 (1967) 1260-1266.
- 1-21. HUANG, T.S. et al.: Design considerations in PCM transmission of low resolution still pictures. Proc. IEEE 55 (1967) 331-337.
- 1-22. HUBEL, H.D. - WIESEL, N.T.: Brain mechanisms of vision. Scientific American 241 (1979. Sep.) 130-144.
- 1-23. HUNT, R.W.G.: Colorimetry. IBA Technical Review 22. Light and colour principles (1984) 16-27.
- 1-24. JAMESON, D.: Theoretical issues of colour vision. In: Jameson, D. - Hurvich, L.M.: Visual Psychotropics. Springer Verlag, Berlin, 1972, 381-412.
- 1-25. JULESZ, B.: Textons, the elements of texture perception, and their interactions. Nature 290 (1981. March) 91-97.
- 1-26. KANADE, T.: Survey: Region segmentation - signal vs semantics CGIP 13 (1980) 279-297.
- 1-27. MAX, J.: Quantising for minimum distortion. IRE Transactions on Information Theory 6/1 (1960) 7-12.

- 1-28. MÉRŐ, L. - VAMOS, T.: Medium-level vision. In: Kanal, L.N. - Rosenfeld, A. (eds): Recent progress in Pattern Recognition 1. North-Holland Publishing Company, Amsterdam, 1980, 93-122.
- 1-29. MERTZ, P. - GREY, F.: A theory of scanning and its relation to the characteristics of the transmitted signal in telephotography and television. BSTJ 13 (1934) 464-515.
- 1-30. Minilux super camera, TV11-23A. Híradástechnika Szövetkezet dokumentáció, Budapest, 1983.
- 1-31. Modular Image Processing (MIP) rendszer. Felhasználói kézikönyv, SZKI dokumentáció, MAL-01/84, Budapest, 1984.
- 1-32. MORGAN, J.: Introduction to the Meteosat System. European Space Operations Centre, Darmstadt, 1978.
- 1-33. NICHOLS, D. et al.: Digital hardware. In: [1-7], 841-871.
- 1-34. NUDD, G.R.: Concurrent systems for image analysis. In: King-sun Fu (ed): VLSI for pattern recognition and image processing. Springer-Verlag, Berlin, 1984, 107-130.
- 1-35. PRESTON, K.Jr.: Image processing software, a survey. In: [1-28], 123-148.
- 1-36. Raster Graphics Handbook. Conrac Division, Conrac Corporation, Covina (California), 1980.
- 1-37. REBEL, B. - WILHELMI, W.: Trends in image processing architectures and implementations. Kurzreferate 1. Internationale Fachtagung Automatische Bildverarbeitung, Berlin, 1985, B2/1-11.
- 1-38. REVERS, A.P.: Parallel computer architectures for image processing. CVGIP 25/1 (1984) 68-88.
- 1-39. SEBESTYEN, I.: The videodisc revolution. Working paper, IIASA, Laxenburg, 1981.
- 1-40. SHANNON, C.E. - WEAVER, W.: A kommunikáció matematikai elmélete. OMIKK, Budapest, 1986.
- 1-41. SMITH, B.: Instantaneous companding of quantized signals. BSTJ 36 (1957) 653-709.

- 1-42. SMITH, V.C. - POKORNY, J.: Spectral sensitivity of the foveal cone photopigments between 400 and 700 nm. Vision Research 15 (1975), 161-171.
- 1-43. SPOT adatok hozzáférési lehetőségei és terjesztése. SPOT szeminárium, Budapest, 1984.
- 1-44. STOCKHAM, T.G.: Image processing in the context of a visual model. Proc. IEEE 60/7 (1972) 828-842.
- 1-45. SZABÓ, J.: Mikrogépes színes display. Mérés és Automatika, 25/5 (1977), 204-206.
- 1-46. TAMURA, H. et al.: Design and implementation of SPIDER - a transportable image processing software package. CVGIP 23 (1983) 273-294.
- 1-47. TSOTSOS, J.K.: Knowledge of the visual process: content, form and use. Proc. 6th International Conference on Pattern Recognition, München, 1982, 654-669.
- 1-48. VIKING rendszer. Programozói kézikönyv. SZKI dokumentáció, MAL-03/85. Budapest, 1985.
- 1-49. VINCENT, V.S.: An overview of the LANDSAT-D project with emphasis on the flight segment. In: [1-18], 2-11.
- 1-50. WICKELGREN, W.: Cognitive Psychology. Prentice-Hall, New York, 1979.
- 1-51. WRIGHT, W.D.: Colour mixture. In: [1-24], 434-454.
- 2-1. ÁLLÓ G. - FERÓ L.: A VIKING képfeldolgozó rendszer. "Neumann János" számítógép-alkalmazási konferencia közleményei, Szolnok, 1986, II./163-171.
- 2-2. BODROGI, H.: Élkemelő eljárások és intenzitáskorrekciók (INT program, [1-31] rendszer része). SZKI dokumentáció, MAL-1/84. Budapest, 1984.
- 2-3. BODROGI, H.: Képsávok közötti műveletek (SAV program, [1-31] rendszer része). SZKI dokumentáció, MAL-35/84. Budapest, 1984.

- 2-4. **BODROGI, H.:** Kompozit kép készítése (MEGKOM program [1-31] rendszer része). SZKI dokumentáció, MAL-40/84. Budapest, 1984.
- 2-5. **COOLEY, J.W. - TUKEY, J.W.:** An algorithm for machine calculation of complex Fourier series. *Mathematics of Computation* 19 (1965) 297-301.
- 2-6. **DARREL, L.W. - STAUFFER, K.C.L.:** A forester's look at the application of image manipulation techniques to multi-temporal LANDSAT data. In: [1-18], 368-373.
- 2-7. **DORST, L.:** A local contrast enhancement filter. In: [1-47], 604-606.
- 2-8. **FREI, W.:** Image enhancement by histogram hyperbolization. *CGIP* 6 (1977) 286-294.
- 2-9. **GEUEN, W. - WIDZGOWSKI, E.:** A simple edge detection algorithm on the basis of visual contour perception. In: [1-47], 843-845.
- 2-10. **KANEKO, T. - MOORE, L.K. - SMART, R.T.:** An interactive color display system for labelling crops. In: [1-18], 408-414.
- 2-11. **KUNDE, A. - MITRA, S.K. - VLADYANATHAN, P.P.:** Application of two-dimensional generalized mean filtering for removal of impulse noise from images. *IEEE Transactions on Acoustics Speech and Signal Processing* 32/3 (1984) 600-609.
- 2-12. **Lebegőpontos speciális processzor (LSP).** Felhasználói kézikönyv. ELGI dokumentáció, Budapest, 1984.
- 2-13. **LEE, J.S.:** Digital image enhancement and noise filtering by use of local statistics. *IEEE PAMI* 2 (1980) 165-168.
- 2-14. **LEE, J.S.:** Refined filtering of image noise using local statistics. *CGIP* 15 (1981) 380-389.
- 2-15. **NAGAO, M. - MATSUYAMA, T.:** Edge preserving smoothing. *CGIP* 9 (1979) 394-407.
- 2-16. **NÁDOR, G.:** FFT, inverz FFT, szűrések (FFTFIL program, [1-31] rendszer része). SZKI dokumentáció, MAL-04/85. Budapest, 1985.

- 2-17. STASZNY G.: Lokális meredekségnövelés. In: [1-48], 149-150.
- 3-1. ANDREWS, H.C. - PATTERSSON, C.L.: Digital interpolation of discrete images. IEEE TC 25/2 (1976) 196-202.
- 3-2. ARCELLI, C. - MASSAROTTI, A.: On the parallel generation of straight lines. CGIP 7 (1978) 67-83.
- 3-3. BRACCINI, C. - MARINO, G.: Fast geometrical manipulations of digital images. CGIP 13 (1980) 127-141.
- 3-4. BRESENHAM, J.E.: Algorithm for computer control of a digital plotter. IBM Systems Journal 4 (1965) 25-30.
- 3-5. EDEN, G.: Two-dimensional fast algorithms applied to memory limited disk operating computers. Computer Programs in Biomedicine 9 (1979) 258-262.
- 3-6. ESCHENBACH, R.F. - OLIVER, B.M.: An efficient coordinate rotation algorithm. IEEE TC 27/12 (1978) 1178-1180.
- 3-7. FREEMAN, H.: Algorithm for generating a digital straight line on a triangular grid. IEEE TC 28/2 (1979) 150-152.
- 3-8. HARALICK, R.: Using perspective transformations in scene analysis. CGIP 13 (1980) 191-221.
- 3-9. HEGEDŰS Gy.Cs.: Háromdimenziós függvények megjelenítése színes raszter-display-n. SZKI dokumentáció, MAL-21/81. Budapest, 1981.
- 3-10. HEGEDŰS, Gy.Cs.: Digitális képek azonosítási pontthalmaz alapján történő geometriai korrekcióját végrehajtó programok kifejlesztése. SZKI dokumentáció, MAL-55/82. Budapest, 1982.
- 3-11. HEGEDŰS, Gy.Cs.: Digitális képek gyors geometriai korrekciói. ELTE egyetemi doktori értekezés, Budapest, 1983.
- 3-12. HEGEDŰS, Gy.Cs.: Digitális képek geometriai korrekciói. Alkalmazott Matematikai Lapok 10/3-4 (1984) 373-388
- 3-13. INGRAM, W.L. - HOOKER, H.L.: The selection of approximation functions for tabulated numerical. Technical Report, X - 64658, G.C. Marshall Space Flight Center, Alabama, 1972.
- 3-14. JORDAN, B.W. et al.: An improved algorithm for the generation of nonparametric curves. IEEE TC 22/2 (1973) 1052-1060.

- 3-15. KIM, C.E.: On cellular straight line segments. CGIP 18 (1982) 369-381.
- 3-16. LOCEFF, M.: The line. Computer 13/6 (1980) 56-66.
- 3-17. PERNY, D. et al.: Perspective mapping of planning textures. Computer Graphics 16/1 (1982) 72-89.
- 3-18. RIESENFELD, R.F.: Homogeneous coordinates and projective planes in computer graphics. IEEE Computer Graphics & Applications 1/1 (1981) 50-55.
- 3-19. ROSENFELD, A.: Digital straight line segments. IEEE TC 23/12 (1974) 1264-1269.
- 3-20. TÁNCZER, T. - HEGEDŰS, Gy.Cs.: Digital processing of TIROS-N images. Időjárás 84/4 (1980) 325-331.
- 3-21. TÁNCZER, T. - HEGEDŰS, Gy.Cs.: An attempt for objective analysis of TIROS-N images. Advances in Space Research 1/1 (1981) 327-330.
- 3-22. TÓZSA, I. - HEGEDŰS, Gy.Cs.: Budapest a világűrben. Földrajzi Értesítő 31/1 (1982) 121-130.
- 3-23. ULICHNEY, R.A. - TROXEL, D.E.: Scaling binary images with the telescoping template. IEEE PAMI 4/3 (1981) 331-335.
- 3-24. WEIMAN, C.F.R.: Highly parallel digitized geometric transformations without matrix multiplication. Proc. International Conference on Parallel Processing, Detroit, 1976, 1-10.
- 3-25. WU, L.D.: On the chain code of a line. IEEE PAMI 4/3 (1981) 347-353.
- 4-1. ARCELLI, C. - SANNITI DI BAJA, G.: A thinning algorithm based on prominence detection. PR 13/3 (1981) 225-235.
- 4-2. ÁLLÓ G.: Élgyanus pontokat kijelölő program (KERES). In: [1-48], 218-221.
- 4-3. ÁLLÓ G.: Élkorrekciós program (ELKORR). In: [1-48], 198-202.
- 4-4. ÁLLÓ, G. - STASZNY, G.: Two algorithms for image segmentation. Microprocessing and Microprogramming 19/1 (1987) 19-25.

- 4-5. BINFORD, T.O.: Inferring surfaces from images. AI 17/1-3 (1981) 205-244.
- 4-6. BLANZ, W.E. - REINHARDT, E.R.: General approach to image segmentation. In: [1-47], 188-191.
- 4-7. BOUKHAROUBA, S. - REBORADO, J.M. - WENDEL, P.L.: An amplitude segmentation method based on the distribution function of an image. CVGIP 29 (1985) 47-59.
- 4-8. BURT, P.J. - HONG, T.H. - ROSENFELD, A.: Segmentation and estimation of image region properties through cooperative hierarchical computation. IEEE SMC 11/12 (1981) 802-809.
- 4-9. CHIEN, C.H. - AGGARWAL, J.K.: A normalized quadtree representation. CVGIP 26 (1984) 331-346.
- 4-10. DANIELSSON, P.E.: An improved segmentation and coding algorithm for binary and nonbinary images. IBM Journal of Research and Development 26/6 (1982) 698-707.
- 4-11. DAVIS, L.S. - MITICHE, A.: Edge detection in textures. CGIP 12 (1980) 25-39.
- 4-12. DUDA, R.O. - HART, P.E.: Use of the Hough transformation to detect lines and curves in pictures. Communications of ACM 15 (1972) 11-15.
- 4-13. EBERLEIN, R.B.: An iterative gradient edge detection algorithm. CGIP 5 (1976) 245-253.
- 4-14. FERÓ L.: Eróziós programok (ERVEZ, EREX, ERSUB2, ERSUB3, EVLAG). In: [1-48], 133-134, 202-203, 204-206.
- 4-15. FERÓ L.: Foltzámpláló programok (FSVEZ, FSEX). In: [1-48], 139-140.
- 4-16. FÜGLEIN, J.: On edge gradient approximations. Pattern Recognition Letters. 1 (1983) 429-434.
- 4-17. FREI, W. - CHEN, C.C.: Fast boundary detection: a generalization and a new algorithm. IEEE TC 26/10 (1977) 988-998.
- 4-18. HARALICK, R.M. - KELLY, G.I.: Pattern recognition with measurement space and spatial clustering for multiple image. Proc. IEEE 57 (1969) 654-665.

- 4-19. HARALICK, R.M. - SHANMUGAM, K.: Textural features for image classification. IEEE SMC 3 (1973) 610-621.
- 4-20. HARALICK, R.M.: Statistical and structural approaches to texture. Proc. IEEE 67 (1979) 786-804.
- 4-21. HARALICK, R.M.: Edge and region analysis for digital image data. CGIP 12 (1980) 60-73.
- 4-22. HARALICK, R.M.: Digital step edges from zero crossings of second directional derivatives. IEEE PAMI 6/1 (1984) 58-68.
- 4-23. HARALICK, R.M. - SHAPIRO, L.G.: Image segmentation techniques. CVGIP 29 (1985) 100-132.
- 4-24. HORN, S.K.P.: Understanding image intensity. AI 8 (1977) 201-231.
- 4-25. HUECKEL, M.H.: An operator which locates edges in digital pictures. JACM 18/4 (1971) 113-125.
- 4-26. JAIN, R. - RHEAUME, D.: A two-stage method for fast edge detection. CGIP 14 (1980) 177-181.
- 4-27. KAPUR, J.N. - SAHOO, P.K. - WONG, A.G.C.: A new method for gray-level picture thresholding using the entropy of the histogram. CVGIP 29 (1985) 273-285.
- 4-28. KIRSCH, R.: Computer determination of the constituent structure of biological images. Computers in Biomedical Research 4 (1971) 315-328.
- 4-29. KITTLER, J.: On the accuracy of the Soebel edge detector. Image and Vision Computing 1/11 (1983) 37-42.
- 4-30. KITTLER, J. - ILLINGWORTH, J.: Minimum error thresholding. PR 19/1 (1986) 41-47.
- 4-31. KONRÁD K.: Hisztogram felbontó (HISZTI) program. In: [1-48]. 142-143.
- 4-32. KOVÁCS Gy.-né: PLIMA - Képfeldolgozó programkönyvtár MO8X mikroszámítógépre. Műszaki terv. SZKI dokumentáció, ELL-0176/81. Budapest, 1981, 27-30.
- 4-33. LAURENT, J.P.: La structure de controle dans les systemes experts. RAIRO Technique et science informatique 3/3 (1984) 161-177.

- 4-34. LEVINE, M.D. - SHAHEEN, S.I.: A modular computer vision system for picture segmentation and interpretation. IEEE PAMI 3/5 (1981) 540-556.
- 4-35. MACKWORTH, A.K.: Interpreting pictures of polyhedral scenes. AI 4 (1973) 121-137.
- 4-36. MARR, D. - HILDREDTH, E.: Theory of edge detection. Proc. Royal Society (London) B207 (1980) 187-217.
- 4-37. MÉRŐ L.: Konturkeresés zajos, digitalizált képekben. Kandidátusi értekezés. MTA SZTAKI Tanulmányok 96/1979, Budapest, 1979, 14-24.
- 4-38. NEVATIA, R. - BABU, K.R.: Linear feature extraction and description. CGIP 13 (1980) 257-269.
- 4-39. OHLANDER, R. - PRICE, K. - REDDY, D.R.: Picture segmentation using a recursive region splitting method. CGIP 8 (1978) 313-333.
- 4-40. OTSU, N.: A threshold selection method from gray-level histograms. IEEE SMC 9 (1979) 62-66.
- 4-41. PANDA, D.P. - ROSENFELD, A.: Image segmentation by pixel classification in (gray level-edge value) space. IEEE TC 27 (1978) 875-879.
- 4-42. PIETIKÄINEN, M. - ROSENFELD, A.: Image segmentation by texture using pyramid node linking. IEEE SMC 11/12 (1981) 822-825.
- 4-43. PIETIKÄINEN, M. - ROSENFELD, A.: Edge-based texture measures. :In: [1-47], 298-300.
- 4-44. PONG, T.C. et al.: Experiments in segmentation using a facet model region grower. CVGIP 25/1 (1984) 1-23.
- 4-45. PREPELICZA, G.: Vonalvékonyító algoritmusok. Államvizsga szakdolgozat. ELTE, Budapest, 1985.
- 4-46. PREWITT, J.: Object enhancement and extraction. In: Lipkin, B.S. - Rosenfeld, A. (eds): Picture processing and Psychotropics. Academic Press, New York, 1970, 75-149.

- 4-47. ROBERTS, L.G.: Machine perception of three-dimensional solids. In: Tippet, J.T. et al. (eds): Optical and Electro-Optical Information Processing. MIT Press, Cambridge, 1965, 159-197.
- 4-48. ROBINSON, G.S.: Edge detection by compass gradient masks. CGIP 6/5 (1977) 492-501.
- 4-49. ROSENFELD, A.: Iterative histogram modification. IEEE SMC 8/4 (1978)
- 4-50. ROSENFELD, A.: Iterative methods in image analysis. PR 10 (1978) 181-187.
- 4-51. SCHACHTER, B.J. et al.: An application of relaxation methods to edge reinforcement. IEEE SMC 7/11 (1977) 813-816.
- 4-52. SHAW, G.B.: Local and regional edge detectors: some comparisons. CGIP 9 (1979) 135-149.
- 4-53. SKLANSKY, J. - CHANZIN, R.L. - HANSEN, B.J.: Minimum perimeter polygons of digitized silhouettes. IEEE TC 21 (1972) 445-448.
- 4-54. SKLANSKY, J. - GONZALEZ, V.: Fast polygonal approximation of digitized curves. PR 12 (1980) 327-331.
- 4-55. STASZNY G.: Szines papirképek szinhelyes tárolása. SZKI Intézeti tanulmány, MAL-21/82, Budapest, 1982.
- 4-56. TENNENBAUM, J.M. - BIRROW, H.G.: Experiments in interpretation-guided segmentation. AI 8 (1977) 241-274.
- 4-57. WATANABE, S. et al.: An automated apparatus for cancer prescreening: CYBEST. CGIP 3 (1974) 350-358.
- 4-58. WESZKA, J.S. - NAGEL, R.N. - ROSENFELD, A.: A threshold selection technique. IEEE TC 23 (1974) 1322-1326.
- 5-1. AHUJA, N. - ROSENFELD, A.: Mosaic models for textures. IEEE PAMI 3/1 (1981) 1-11.
- 5-2. BALL, G.H. - HALL, D.J.: A clustering technique for summarizing multivariate data. Behavioral Science 12 (1967) 153-155.

- 5-3. CHETVERIKOV, D.: Textural anisotropy features for texture analysis. Proc. IEEE Computer Society Conference on Pattern Recognition and Image Processing, Dallas, 1981, 583-588.
- 5-4. CHETVERIKOV, D.: Texture imperfections. Pattern Recognition Letters 5 (1987) 45-50.
- 5-5. CSETERIKOV, D.: A digitális texturavizsgálat néhány új módszere. Kandidátusi értekezés. MTA SZTAKI Tanulmányok 204/1988, Budapest, 1988.
- 5-6. DÁVID L. - FÜGLEIN J. - SZABÓ J.: Számítógépes képfeldolgozás a vizsgádzádkodásban. Magyar Vizgádzádkodás 1981/5, 17-22.
- 5-7. DÁVID, L. - SZABÓ, J.: Analysis of the spatial distribution of water quality and pollution sources of a shallow lake by digital image processing: Proc. 1st Thematic Conference on Remote Sensing of Arid and Semi-arid Lands, Cairo, 1981:
- 5-8. DAVIS, L.S. - JOHNS, S.A. - AGGARVAL, J.K.: Texture analysis using generalized co-occurrence matrices. IEEE PAMI 1 (1979) 251-259.
- 5-9. DAVIS, L.S.: Image texture analysis - a survey. University of Texas Technical Report, TR-139 (1980).
- 5-10. FOLEY, D.H. - SAMMON, J.W.: An optimal set of discriminant vectors. IEEE TC 24/3 (1975)
- 5-11. FÜGLEIN J. - KEISZ P. - SZABÓ J.: Számítógép alkalmazása a képkéértékelésben - hazai eredmények. MTESZ IV. Földfotó Szeminárium, Vác-Penc, 1980.
- 5-12. FÜGLEIN, J. - KITTLER, J.: The effect of pixel correlations on class spearability. Pattern Recognition Letters. 1 (1983) 401-407.
- 5-13. FU, K.S.: Syntactic pattern recognition and applications. Prentice-Hall, New York, 1982.
- 5-14. FUKUNAGA, K. - KOONTZ, W.L.G.: Application of the Karhunen-Loève expansion to feature selection and ordering. IEEE TC 19 (1970) 311-318.
- 5-15. GONZALEZ, R.C. - THOMASON, M.G. Syntactic pattern recognition. Addison-Wesley, New York, 1978.

- 5-16. RÉVÉSZ, P. - FRITZ, J.: Az alakfelismerés statisztikus módszerei. MTA MKI, Budapest, 1974.
- 5-17. ROSENFELD, A.: Picture languages, formal models for picture recognition. Academic Press, New York; 1979.
- 5-18. SPÁTH, H.: Cluster analysis algorithms for data reduction and classification of objects. Ellis Horwood Ltd., Chichester, 1980.
- 5-19. SWAIN, P.H.: Pattern recognition: a basis for remote sensing data analysis. LARS, West Lafayette, 1973.
- 5-20. TOU, J.T. - GONZALEZ, R.C.: Pattern recognition principles. Addison-Wesley, New York, 1974.
-
- F-1. AHMED, N. - RAO, K.R.: Orthogonal transforms for digital signal processing. Springer-Verlag, Berlin, 1975.
- F-2. ARATÓ, M.: Fejezetek a matematikai statisztikából számítógépes alkalmazásokkal II. SZÁMKI közlemények 22/1979, Budapest, 1979.
- F-3. BAHVALOV, N.Sz.: A gépi matematika numerikus módszerei. Tankönyvkiadó, Budapest, 1977.
- F-4. BRIGHAM, O.E.: The fast Fourier-transform. Prentice-Hall, New Jersey, 1974.
- F-5. DANIELSSON, P.E.: Incremental curve generation. IEEE TC 19/9 (1969) 783-793.
- F-6. FREEMAN, H.: Computer processing of line-drawing images. Computer Survey 6 (1974) 57-97.
- F-7. GAAFAR, M.: Convexity verification, block-chords and digital straight lines, CGIP 6 (1977) 361-370.
- F-8. KOLMOGOROV, A.N. - FOMIN: Bevezetés a valós függvénytanba. Műszaki Könyvkiadó, Budapest, 1981.
- F-9. MESZÉNA, Gy. - ZIERMANN, M.: Valószínűségelmélet és matematikai statisztika. Közgazdasági és Jogi Kiadó, Budapest, 1981.
- F-10. ROGERS, D.F. - ADAMS, J.A.: Mathematical elements for computer graphics. McGraw-Hill Book Company, New York, 1976.
- F-11. VINCZE, I.: Matematikai statisztika. Tankönyvkiadó, Budapest, 1974.
- F-12. Vlagyimirov, V.Sz.: Bevezetés a parciális differenciálegyenletek elméletébe. Műszaki Könyvkiadó, Budapest, 1979. —

A kiadásért felelős az Akadémiai Kiadó és Nyomda Vállalat főigazgatója
A nyomdai munkálatokat
az Akadémiai Kiadó és Nyomda Vállalat végezte
Felelős vezető: Hazai György
Nyomdai táskaszám: 18804
Felelős szerkesztő: Sente László — Műszaki szerkesztő: Merkly László
A kötéstervező: Bogdán Hajnal munkája
Kiadványszám: 2174
Megjelent 29,75 (A/5) ív + 1,75 (A/5) ív melléklet terjedelemben
HU ISSN 0236—6312

A digitális képfeldolgozás a számítógépek alkalmazásának egyik legkorábbi területe, szélesebb körben azonban csak az igen nagy integráltságú félvezető tárcák és a nagysebességű mikroprocesszorok megjelenése után kezdett elterjedni. Eleinte egy-egy konkrét feladatra születtek megoldások, általánosításokra nem nagyon törekedtek. A 70-es években kezdtek el foglalkozni a képfeldolgozás elméleti megalapozásával és a megoldások érvényességi körének egzakt vizsgálatával. A 80-as évek elején, amikor a japánok elindították az 5. generációs számítógépek kifejlesztésének programját, a „látás automatizálása” külön hangsúlyt kapott. Bár azóta is intenzív kutatómunka folyik, az egységes elmélet ez ideig nem született meg.

Ebben a kötetben a szerzők áttekintik a digitális képfeldolgozás hazánkban is művelt legfontosabb területeit. Az elmélet hiánya megnehezítette a rendszerező munkát, ezért a szerzők kidolgoztak egy modellt, és az anyagot eszerint rendszerezték 4 fő témakörbe (képjavítások, geometriai korrekciók, szegmentálás, alakfelismerés). Ezeken belül a megoldási módszereket vették a csoportosítás alapjául. A könyv megértéséhez szükséges matematikai ismereteket függelékben foglalták össze.

Magyar nyelven ez ideig nem jelent meg hasonló tárgyú mű, és ilyen rendszerező munka az idegen nyelvű irodalomban is ritka. Külön kiemelhető a geometriai korrekciók tárgyalása, amely ebben a formában első ízben jelenik meg.

A könyvet képi információk feldolgozásával foglalkozó szakembereknek ajánljuk. A felsőfokú matematikusok, programozó, programtervező matematikusok és villamosmérnökök képzésében tankönyvként is használható.