

Berke József - Hegedüs Gy. Csaba - Kelemen Dezső - Szabó József

# Digitális képfeldolgozás és alkalmazásai

**ANALÍZIS**

13	12	25	36	88
33	25	44	12	18
47	45	88	77	41
66	36	69	58	48
25	47	13	59	87

**ADAT**

**KÉP**



**SZINTÉZIS**

# HV-PCI6, COMET, NEBULUS

*PCI Buszos  
Video-digitalizáló Kártyacsalád*

## *HV-PCI6*

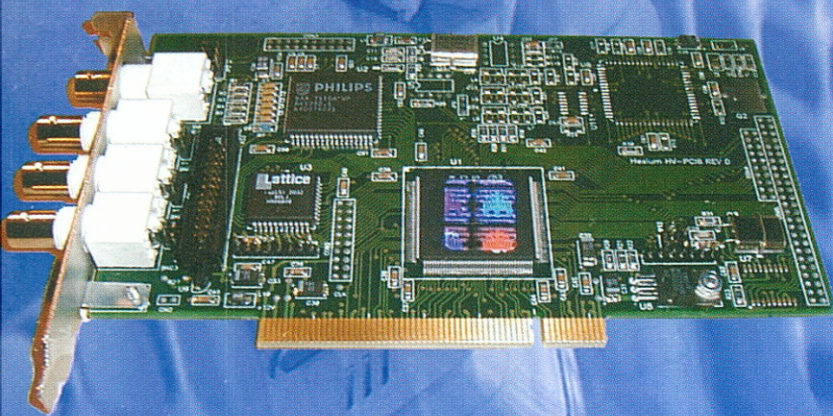
*Kompozit bemenet (CVBS) 6  
SHVS bemenet (Y/C) 3  
RGB bemenet 7  
Video kimenet (CVBS, Y/C) 1  
Valós idejű tömörítés -  
Képfeldolgozás HW támogatással -  
Hardverkulcs opcionális*

## *COMET (2000. júniustól)*

*Kompozit bemenet (CVBS) 6  
SHVS bemenet (Y/C) 3  
RGB bemenet 7  
Video kimenet (CVBS, Y/C) 1  
Valós idejű tömörítés igen  
Képfeldolgozás HW támogatással igen  
Hardverkulcs opcionális*

## *NEBULUS*

*Kompozit bemenet (CVBS) 6  
SHVS bemenet (Y/C) 3  
RGB bemenet 7  
Video kimenet (CVBS) 1  
Valós idejű tömörítés -  
Képfeldolgozás HW támogatással -  
Hardverkulcs opcionális*



### *A kártyacsalád közös jellemzői*

*Fejlesztői támogatás C, C++, C++ Builder, Delphi, Visual Basic környezetben; Skálázható felbontás 768x576-ig; 8 bit gray scale, 15/16/24 bit RGB, 16 bit YUV; PAL, SECAM, NTSC bemeneti jel; 25 ill. 30 kép/másodperc; Windows NT, Windows 95/98; Twain32 adatforrás; Video for Windows adatforrás; Kikapcsolható AGC a bemeneten (6 dB); Minden csatorna 8 bites; 24 hónapos hardver garancia*

*Oktatási intézményeknek 10% kedvezmény!!!*

Berke József - Hegedűs Gy. Csaba - Kelemen Dezső -  
Szabó József

# **Digitális képfeldolgozás és alkalmazásai**

Második, változatlan kiadás

Berke József - Hegedűs Gy. Csaba - Kelemen Dezső -  
Szabó József

# Digitális képfeldolgozás

## és

# alkalmazásai

Nyitott rendszerű képzés oktatási segédlete

Tankönyv CD melléklettel

VE Georgikon Mezőgazdaságtudományi Kar, Keszthely  
PICTRON Kft., Budapest

*A könyvet írták:*

BERKE JÓZSEF, DR. HEGEDŰS GY. CSABA,  
KELEMEN DEZSÓ, SZABÓ JÓZSEF

*A könyvet szerkesztette:*

SZABÓ JÓZSEF pictron@hungary.net

*A CD lemezt szerkesztette:*

BERKE JÓZSEF berke@georgikon.hu

*Lektorálta:*

DR. SZELEZSÁN JÁNOS,  
A MATEMATIKAI TUDOMÁNYOK KANDIDÁTUSA

© Berke József, dr. Hegedűs Gy. Csaba,  
Kelemen Dezső, Szabó József

**ISBN: 963 00 5744 1**

*Kiadó:*

Veszprémi Egyetem Georgikon Mezőgazdaságtudományi Kar, PICTRON Számítás- és  
Videotechnikai Kft., Szeiler és Tsa. Bt., KvarK Számítástechnikai Bt.

*Felelős kiadó:* Szeiler Gábor

*Borító, tördelés:*

Szeiler és Tsa. Bt., Keszthely

*Nyomta:*

K&V Nyomda, Zalaegerszeg

# Előszó

A digitális képfeldolgozás az utóbbi évtizedekben robbanásszerű fejlődésen ment át, és mindennapi életünk részévé vált, még akkor is, ha az emberek többségének fogalma sincs az ezzel kapcsolatos alapismeretekről. Elég a TV-re gondolnunk, amit mindenki nézhet a digitális képfeldolgozással kapcsolatos ismeretek hiányában is, holott a televízió ma már nem tudna működni ezen a technikán alapuló berendezések nélkül. Megtalálhatók a digitális képfeldolgozás eszközei az orvosdiagnosztikai műszerekben, az úrkutatás eszköztárában, a nyomdatechnikában, a fototechnikában, az ipari folyamatirányításban éppúgy, mint az elektronikus játékautomatákban. Bátran mondhatjuk tehát, hogy az élet minden területén használjuk ezeket az eszközöket, ezért egyre több szakembernek van, és lesz szüksége a számítógépes képfeldolgozással kapcsolatos ismeretekre. Elsősorban azoknak szántuk ezt a könyvet, akik munkájuk során kapcsolatba kerülnek vagy fognak kerülni a számítógépes képfeldolgozással, és ezért többet szeretnének tudni erről a területről.

A könyv első kiadásának egyik fő motivációja az volt, hogy a Gábor Dénes Főiskolán a digitális képfeldolgozás tantárgy 1996. évben beinduló oktatásához megfelelő tankönyv álljon rendelkezésre. Ismereteink szerint a könyvet és a CD-t más egyetemeken, főiskolákon és a felsőfokú szakképzésben is sikerrel alkalmazták.

Az olvasó a 2002. évi harmadik kiadás változatlan utánnomását tartja a kezében, ennek készítésekor felhasználtuk az elmúlt évek – több mint hatezer hallgatót érintő - oktatási tapasztalatait. Ezen túlmenően, a korábbi CD lemez helyett egy tartalmában és felépítésében korszerűsített CD melléklettel adjuk ki a könyvet.

A könyv bevezető jellegű, minimális matematikai alapismeretet tételez fel, és önálló tanulásra alkalmas, így bízunk benne, hogy nemcsak szervezett oktatás keretében, hanem egyéni érdeklődés által vezérelt ismeretszerzés céljából is sokan fogják a könyvet és a mellékelt CD lemezt használni.

**A tananyag frissítése, valamint egyéb hasznos információk** az alábbi Internet címen található: <http://www.georgikon.hu/digkep.htm>.

A tananyaghoz közvetlenül kapcsolódnak az alábbi Internet címek: <http://www.pictron.hu>, <http://www.mtr.hu>, <http://www.kvarkmed.hu>.

Miután legeredményesebben a könyv és a CD együttes használatával tanulhatunk, ezért a Bevezetésben röviden tájékoztatjuk az olvasót arról mit, hol találhat meg.

A szerzők ezúton fejezik ki köszönetüket azoknak a kollégáknak, akik a könyv és a CD lemez elkészítésében közreműködtek.

Köszönet illeti szakmai lektorunkat Dr. Szelezsán Jánost, hogy észrevételeivel és tanácsaival hozzájárult könyvünk színvonalának emeléséhez.

Budapest, 2001. december

*A szerzők*

## Bevezetés

### Hogyan használjuk a könyvet?

Az érdemi ismeretek tárgyalása előtt a könyv és a mellékletét képező CD lemez használatához szeretnénk néhány gyakorlati tanácsot adni.

A CD nagy tárolókapacitása lehetővé tette, hogy a TULIPP és FFT gyakorló programok mellett, képpel, videókkal illusztrált alkalmazási példákat is elhelyezzünk a lemezen, sőt azok, akik szívesen használják a számítógépet hosszabb szövegek olvasására – átdolgozott és bővített interaktív formában - a teljes tananyagot megtalálhatják a CD-n.

A könyv 1. fejezetében áttekintjük az emberi látással kapcsolatos alapfogalmakat, ezt követően a 2. fejezetben a képfeldolgozás speciális eszközeiről adunk áttekintést. A képfeldolgozás elméleti témaköreivel 3.-8. fejezetekben foglalkozunk. Érdeemes az elméleti részek olvasása közben a CD megfelelő fejezeteibe is bepillantani, mivel a szemléltető képeket ott helyeztük el. Az elméleti ismereteket tartalmazó fejezetekben gyakorlatokat talál az olvasó, amelyeket a CD-n lévő TULIPP programmal hajthat végre. A program kezelését a következő pontban ismertetjük. Javasoljuk a gyakorlatokat közvetlenül az egyes fejezetek elolvasása után elvégezni.

A 9.-17. fejezetekben rövid áttekintést adunk a különböző alkalmazási területekről. Ezekben a fejezetekben igyekeztünk az egyes alkalmazási területek jellegzetességeit megfogalmazni. Konkrét alkalmazások részletesebb leírását a CD-n helyeztük el képekkel, néhány esetben videó részlettel illusztrálva azokat.

Az egyes fejezetek után ellenőrző kérdéseket talál az olvasó. Az ezekre adott válaszokkal ellenőrizhető, hogy a legfontosabb ismereteket milyen mélységben sikerült elsajátítani.

Az ajánlott irodalom részben néhány könyv címet adtunk meg a további részletek után érdeklődőknek.

### Hogyan használjuk a CD-t?

#### Figyelem!

*A CD használatához kérjük, olvassa el a gyökérben található readme.txt állományt !*

## A TULIPP gyakorló program kezelése

A képfeldolgozás ismeretek akkor válnak igazán élővé, ha a témával ismerkedő a gyakorlatban is ellenőrizheti, megfigyelheti az egyes eljárások hatását a képre, saját szemével győződik meg arról, mi történik, ha megváltoztatja egy-egy paraméter értékét. Az ideális állapot az lenne, ha a könyv minden olvasójának rendelkezésére állna egy jól kiépített képfeldolgozó rendszer, megfelelő képfeldolgozó programokkal. Biztos lesznek ilyenek is a könyv olvasói között, de azokra gondolva, akiknek csak egy „átlagosnak” mondható PC áll rendelkezésére, a PICTRON Kft által kifejlesztett TULIP (TUtorial for Image Processing) oktató és gyakorló program egyszerűsített változatát helyeztük el a CD-n, amely TIFF és PCX formátumú képekkel képes az egyes gyakorlatokat végrehajtani. A program segítségével az ismertetett eljárások többségét az olvasó is kipróbálhatja.

A program használatához olyan PC konfiguráció szükséges, amelyen WINDOWS 95/98 fut, és a VGA kártya minimum 256 színes, 800x600-as felbontású.

A TULIPP program önálló telepítéséhez először hozzon létre egy TULIPP könyvtárat a merevlemezen, és másolja be a könyvhöz csatolt CD lemezen lévő TULIPP könyvtár tartalmát, a CD-n található könyvtárstruktúrával teljesen azonos módon, ide. Ezt követően már csak az install.exe programot kell elindítani a WINDOWS-ból, a program a szokásos módon létrehozza a TULIPP ikont.

A gyakorló program az interaktív keretrendszerből is közvetlenül elérhető. ***Mindenképp javasoljuk, hogy olvassa el a CD lemezen található readme.txt állományt!***

Legjobb, ha rögtön elindítjuk a programot, ekkor a címlap eltűnése után a végrehajtható gyakorlatok felsorolása jelenik meg. Mielőtt bármelyiket kiválasztanánk, kattintsunk a Help gombra, ugyanis a program Help-je ezen a szinten tartalmazza a kezeléssel kapcsolatos összes tudnivalót. Fontos tudni, hogy a Help a teljes TULIPP verzióra vonatkozik, a mellékelt TULIPP program azonban egy szűkített változat, így nem minden funkció működik a Help-ben leírtak szerint. Ezek: a digitalizálás, valamint a tananyag hipertextes, multimédiás feldolgozása. Ez utóbbit hivatott pótolni a CD-n található tananyag és a könyv. Annak érdekében, hogy a gyakorlatok valós alkalmazásokon keresztül legyenek végrehajthatók, néhány képet is csatoltunk a programhoz. Miután átfutottunk a program kezelését tartalmazó Help-en, indítsuk el bármelyik funkciót. Ha ezen a szinten kattintunk a Help-re akkor az adott funkcióra vonatkozó specifikus információt kapunk, mégpedig három szinten; a funkció elméleti háttéréről, a kezelésről, továbbá javaslatokat kaphatunk az elvégzendő gyakorlatokra.

Ha bele akarunk vágni egy funkció végrehajtásába, akkor képeket a Képforrás-on belül a Tesztkép menüpont alatt találunk. A Tesztképen belül van néhány generált



tesztkép, de a Képx megnevezések valódi képeket jelölnek. A Kép1, Kép9 és Kép10 színes kép a többi monokróm szürke árnyalatú (fekete/fehér). A színes képek is feldolgozhatók, mint fekete/fehér képek, és fordítva, de ez utóbbinak nincs sok értelme, hiszen a három színösszetevő azonos lesz, tehát valójában továbbra is monokróm képpel fogunk dolgozni.

Gyakorlatokat csak a képfeldolgozás témaköreit tárgyaló 3.-8. fejezetekhez adtunk. A gyakorlatoknál egyszerűen felsoroljuk a kiválasztott funkciót, utána a kiválasztott menüpontokat és paramétereiket. A gyakorlatok eredményének az értékelését az olvasóra bízjuk, legfeljebb csak ráirányítjuk a figyelmét a vizsgálandó jelenségre. Nem kívántuk senkinek a kísérletező kedvét szegni, de úgy gondoltuk, hogy célszerű néhány olyan feldolgozási sort összeállítani, amelyek a legfontosabb ismeretek alátámasztására szolgálnak, egyben segítik a program kezelésének elsajátítását is. Felhívjuk a figyelmet arra, hogy a gyakorlatok leírásánál a funkció kiválasztástól kezdve az összes menüpontot felsoroljuk annak érdekében, hogy ne csak sorrendben lehessen elvégezni a gyakorlatokat. Természetesen, ha egymást követő gyakorlatokat végzünk, akkor elegendő addig a pontig visszamenni, ahol az első eltérő menüpont vagy paraméter szerepel. Fontos tudni, hogy az input képmezőben kijelölhetünk téglalap alakú részleteket. Ha egy funkciót különböző paraméterekkel, különböző részletekre hajtunk végre, akkor az output képmezőben egyszerre látjuk az eredményeket, így könnyebb összehasonlítani azokat. A fenti általános megjegyzések után javasoljuk, hogy minél több gyakorlatot végezzen az olvasó, ehhez használja az alábbiakban felsoroltakon kívül a Funkció Help-nél található javaslatokat, de bárki maga is bátran kísérletezhet a programmal.

## Az FFTDEMO gyakorló program kezelése

A program a könyvhöz mellékelt CD lemezen található TULIPP program kiegészítéseként készült és az /fft alkönyvtárban található.

A tankönyv 4.5 és 4.6 fejezeteiben szó esik a frekvencia tartományban történő képjavításról, de annak szemléltetésére alkalmas funkciót a TULIPP nem tartalmaz. Ezt a hiányt pótolja az FFTDEMO program. A program használata előtt ajánlatos a tankönyv fenti fejezeteit áttanulmányozni. Az /fft könyvtárban helyeztük el a programhoz mellékelt dokumentációt is.

A program egyetlen fájl, közvetlenül indítható. Kezelése egyetlen menüin keresztül történik. Indítás után 5 ablak jelenik meg, ezeket a Windowsban megszokott módon elhelyezhetjük, átméretezhetjük.

## A digitális képfeldolgozás fogalma

A **számítógépes képfeldolgozás** az alkalmazott matematika, elektronika és számítástechnika viszonylag új, rohamosan fejlődő területe.

A számítógépes képfeldolgozás célja és lényege, hogy a környezetünkből származó vizuális információt a számítógép segítségével feldolgozzuk és kiértékeljük. Bár a környezetünkben lévő tárgyak háromdimenziós objektumok, és egyes alkalmazásoknál ezek egymáshoz való elhelyezkedését is figyelembe kell venni (pl. robotok vezérlése), azonban az esetek jelentős részében elegendő az objektumok **kétdimenziós** (azaz síkbeli) leképezésének feldolgozása. Részben ezért, részben terjedelmi okokból a továbbiakban többnyire csak kétdimenziós képek feldolgozásával foglalkozunk. A kétdimenziós képeket először a számítógép számára feldolgozható formába kell alakítani, azaz **digitalizálni** kell. A kétdimenziós digitális képen a lényegi információt hordozó képelemeket értékeseknek (objektumoknak) tekintjük, melyeket a kép többi részétől - a háttértől (ha van ilyen) - el kell választani. A lényeges információt nem az egyes képpontok, hanem bizonyos csoportjaik hordozzák. Ezekhez olyan jellemző tulajdonságokat (sajátságokat: *features*) rendelhetünk, melyekkel a képpontok önmagukban nem rendelkeznek. Az objektumokat e sajátságaik szerint osztályokba sorolhatjuk. Az objektumok kapcsolatait, egymáshoz való viszonyukat a kép struktúrája fejezi ki. Ennek elemzése olyan, magasabb szintű képleíráshoz vezet, amelyet az objektumok önmagukban nem hordoznak.

A számítógép és kép kapcsolatának két fő ága a **számítógépes grafika**, valamint a **digitális képfeldolgozás**. Ezek célkitűzéseikben és módszereikben lényegesen eltérnek egymástól, bár azonos, vagy igen hasonló eljárásokat is használnak, és a képleírásban is hasonló elveket követnek. Mindkét ágon körvonalazni lehet a következő három szintet:

1. A fizikai szinten (*signal level*) a képet képpontok halmazaként kezeljük. A képi információt az egyes képpontokhoz rendelt számok (világosság-, illetve színkódok) hordozzák. A feldolgozás során a bemenő képből kimenő képet állítunk elő.
2. Az elemzési szinten (*evaluation level*) a sajátságaikkal leírható objektumok állnak a feldolgozás középpontjában. Az elemzés célja az, hogy bemenő képből képleírást (grafika esetén pedig: leírásból kimenő képet) készítsünk.
3. Az értelmezési szinthez (*semantic level*) tartoznak a képleírás és a kép felismerése közötti folyamatok. A végső cél: a kép automatikus megértése, azaz a **látás automatizálása**.

Az egyes szintek feladatait a következő csoportokba sorolhatjuk:

1. A fizikai szinthez tartozó módszereket és eljárásokat **képátalakításoknak** nevezzük. (A képfeldolgozás korai szakaszát szinte kizárólag ezek jelentették.) Két fő területe:
  - A **képkorrekciók** célja egyrészt bizonyos hibák kijavítása (erre szolgál például a geometriai korrekció, valamint az intenzitáskorrekció), másrészt a **kép helyreállítása**, illetve a lényeges képi **információtartalom kiemelése**.
  - A **szegmentálás** az értékes képpontok háttértől való elválasztására szolgál.
2. Az elemzési szinten végzett feldolgozást képosztályozásnak nevezzük. A statisztikus alakfelismerésnek, illetve a textúraelemzésnek az objektumok sajátosságai, az egyre fontosabb szerepet kapó szintaktikus alakfelismerésnek pedig az objektumok kapcsolatrendszere jelentik a bemenő adatokat.
3. A legmagasabb szint a képfelismerés (scene analysis). Ez a - legkevésbé kiforrott - terület kapcsolatot teremt az elemzett kép, és a tudásbázis, illetve modellbázis között.

## Történeti áttekintés

A **számítógépes képfeldolgozás** gyakorlatilag egyidős a számítástechnika megjelenésével. A képi (*vizuális*) információ számítógépes feldolgozásának mintegy 50 éves múltja van.

Bár a számítógépet megelőző időszakban az **optikai képfeldolgozás** (hagyományos fotó-kidolgozás, analóg képelemzés) igen magas színvonalat ért el, hamar nyilvánvalóvá vált, hogy bizonyos felhasználási területeken (pl.: rádiócsillagászati, atomfizikai vagy radar-felvételek kiértékelése) a pontossági igények, valamint a képanyag nagy mennyisége miatt **gyökeresen új módszerekre van szükség**. Ez a felismerés elősegítette a terület komoly kutatását és a fejlesztési kiadások folyamatos növelését.

## Ötvenes évek

Természetesen a kezdeti időszakot az jellemezte, hogy elsősorban a korábbi módszerek számítógépes kiváltására, pontosítására helyezték a súlyt. Kezdettől fogva nyilvánvaló volt azonban, hogy a számítógéppel gyökeresen új alkalmazási lehetőségek is elérhetővé válnak.

A számítógépes képfeldolgozás egyik alapvető feltétele a **képdigitalizálás** eszközeinek kifejlesztése, elméletének kidolgozása volt. Bár a mintavételezés elvi kérdéseire már a 30-as évek elején születtek megoldások, az információelméleti megalapozást

Shannon csak 1949-ben végezte el. Az első A/D átalakító (analóg - digitál átalakító) eszköz megjelenésére pedig kb. az ötvenes évek elejéig kellett várni.

Az ötvenes évek második felére a számítógépek (műveleti sebességben, tárcapacitásban, vagy pl. a folyamatos üzemképesség időtartamában) elérték azt a fejlettségi fokot, amely lehetővé tette a nagy műveletigényű képfeldolgozási feladatok megoldását.

A folyamatos üzemképesség kezdetben igen fontos jellemzője volt a számítástechnikai eszközöknek. Ez abból következett, hogy akkoriban az összes elektronikai eszközhöz hasonlóan a számítógépek is elektroncsöves felépítésűek voltak, melyek átlagos élettartama - még az ipari kivitelű csövek esetében sem haladta meg a 20.000 üzemórát. Így érthető, hogy egy 10-30.000 logikai kaput (azaz csövet) tartalmazó gép életében már az egy órás folyamatos üzem is nagy eseménynek számított. Megjegyzendő még az is, hogy gépeket ma már elképzelhetetlenül nagy kiszolgáló apparátus vette körül a 10...50 kilowattos teljesítmény-felvétel és a hűtési igény miatt.

## Hatvanas évek

A 60-as évek végéig elsősorban a **konkrét feladatok megoldása** volt a cél, s általánosításra, elméleti megalapozásra nem túlzottan törekedtek. Rendszerint az analóg jelfeldolgozási **módszereket adaptálták**, vagy **heurisztikus eljárásokat** fejlesztettek ki. Az eljárások alkalmazhatóságának feltételeit, érvényességi körüket ritkán vizsgálták. Nagy lökést adott a képfeldolgozási fejlesztéseknek az űrkutatás beindulása.

A hatvanas évek végére (1969) megalkották a CCD (Charge Coupled Device) érzékelőt (George Smith és Willard Boyle).

## Hetvenes évek

A 70-es évtizedet - a megelőzővel szemben - a **matematikai megalapozottság igénye** jellemezte. Számos feladatra dolgoztak ki egzakt, vagy matematikailag precíz közelítő megoldásokat. Számos rendszerező munka látott napvilágot.

Az évtized második felére megjelentek az első **univerzális látórendszerek**, melyek kialakulásához a hardver nagyarányú fejlődése és elterjedése is hozzájárult. Nyilvánvalóvá vált, hogy a **látás automatizálásához** elengedhetetlen az elméleti alapok kellő kiépítése. A képfeldolgozás egyre több alkalmazást nyert a mindennapi polgári életben is, bár még meglehetősen költségesnek bizonyult. Az alkalmazási területeknek megfelelően kialakultak a képfeldolgozás **részterületei**.

## Nyolcvanas évek

A nyolcvanas évek elején a kisszámítógépek alkalmazása volt jellemző a képfeldolgozási feladatokban, mert ebben a környezetben az interaktivitás, és a képfeldolgozáshoz használt speciális eszközök (megjelenítő, digitalizáló) illesztése is könnyebben megoldható volt. A képfeldolgozási feladatok jelentős része a számítógépekhez illesztett speciális eszközök segítségével került megvalósításra. Ez az alkalmazási területek bővülését is magával hozta.

## Kilencvenes évek

Az utolsó, (napjainkig terjedő) időszak legfőbb jellemzője a személyi számítógépes képfeldolgozás tömeges megjelenése. A képpel összefüggő alkalmazási területek (pl.: DTP (=desktop publishing); orvosi képek feldolgozása; ipari folyamatszabályozás, minőség-ellenőrzés, optikai karakterfelismerés.; stb.), hihetetlen mértékben kiszélesedtek, a szó szoros értelmében a mindennapi élet részévé váltak. Az Internet és multimédia megjelenése a képek használatát a számítógépek szabványos részévé tette, a képdigitalizáláshoz és megjelenítéshez szükséges eszközök ára magáncélú felhasználásra is elérhetővé vált.

Mindezek ellenére, a számítógépes képfeldolgozás **egységes és általános elmélete** még nem született meg, nagyon sok a kutatni-fejleszteni való. Jelentős eredmények születtek viszont a képformátumok és felhasználói felületek szabványosítása területén.

## Közeljövő

Természetesen nagyon nehéz megjósolni, hogy felgyorsult világunkban - akárcsak néhány év múlva is - milyen technikai eszközöket fejlesztenek majd ki, s a sebességiteljesítőképeségi korlátok mennyire fognak kitolódni. Egyelőre a fejlődés **üteme** is növekszik, másrészt a váratlan, új felfedezések is alapjaiban határozhatják meg a jövő technikáját.

Ami joggal elvárható: hamarosan megjelennek a **„látórobotok”**, azaz a jelenlegi alapszintű képfeldolgozásból kifejlődik az integrált, bonyolult feladatokat intelligensen megoldó **gépi látás**.

## Ellenőrző kérdések

1. Mit nevezünk számítógépes képfeldolgozásnak?
2. Melyek a képfeldolgozás fő ágai?
3. A feldolgozások során milyen szinteket különböztetünk meg?
4. Mi tartozik a fizikai szinthez?
5. Mi a történeti előzménye a számítógépes képfeldolgozásnak?
6. Mi jellemző az ötvenes évek képfeldolgozására?
7. Mikor történik jelentős matematikai alapozás?
8. Van-e általános elmélete a számítógépes képfeldolgozásnak?
9. Mi jellemző a kilencvenes évek képfeldolgozására?
10. Mi várható a jövőben?

# 1. AZ EMBERI LÁTÁS

## 1.1 Bevezetés

Az ember és környezete közötti kapcsolatban a **LÁTÁS** a legfontosabb információ-szerzési mód, rendkívüli tömörsége és összetettsége miatt. A világosság, szín, alak, térbeliség illetve mozgás érzékelése egyaránt része annak az összetett fiziológiai folyamatnak, melyet látásnak nevezünk.

A számítógépes képfeldolgozás sokszor az emberi látást próbálja modellezni, más esetekben pedig az emberi látással kapcsolatos ismereteket felhasználja, beépíti a feldolgozás folyamatába, ezért feltétlenül szükségesnek tartottuk az emberi látással kapcsolatos ismeretek áttekintését.

## 1.2. A látás alapjai

A látás azon alapul, hogy az igen rövid hullámhosszúságú, 380 - 780 nm ( $1 \text{ nm} = 10^{-9} \text{ m}$ ) tartományba eső (azaz kb.  $7.89 \cdot 10^{14}$  -  $3.85 \cdot 10^{14}$  Hz frekvenciájú) **elektromágneses sugárzás** energiája a szemünkben fényérzetet kelt. (Ezt a sugárzást a továbbiakban **fénynek** nevezzük.) Az elektromágneses sugárzással kapcsolatosan további információkat talál az olvasó a távérzékeléssel kapcsolatos 9. fejezetben.

A szem fényérzékelési mechanizmusának lényege: a sugárzási energia hatására a szem recehártyájának (*retinájának*) idegszálvégződéseibe lépnek; fizikai-kémiai folyamatok indulnak be, melyek az agy megfelelő központjaihoz idegingerület formájában továbbítódnak.

Két - egymástól kismértékben különböző - képet érzékelünk, melyek agyunkban térbeli kép érzetét keltik.

A látást (fizikai-optikai oldalról megközelítve) a fény, valamint fényelnyelő-fénytörő-fényvisszaverő testek kölcsönhatásai teszik lehetővé.

A szükséges fényt valamilyen **elsődleges fényforrás** biztosítja. Jellemzői: spektrális energia-eloszlása, irányfüggése.

A látótér objektumai - mint **másodlagos fényforrások** - visszaverik a rájuk eső fényt. A visszavert fény **valamennyi** jellemzője megváltozik; ez a változás hordozza a látáshoz szükséges információt.

A visszavert fény a szembe jut. Az alapvető optikai összefüggések (pl.: **Snellius-Descartes törvény**) a szem működésére nézve is érvényesek. A szemnek, mint leképező rendszernek legfontosabb optikai alkotóeleme a **szemlencse**.

Szemünk alakja automatikusan követi a képalkotás igényeit; ezt a tudat alatti szabályozást nevezhetjük **autófókusz**nak is.

A leképezés eredményeként a retina belső (gömb-) felületén létrejön a fordított állású kép.

A **képérzet** igen összetett fiziológiai folyamat, melyben a szem, ingerület-közvetítő idegpályákon, illetve az agy látóközpontjain kívül igen nagy szerepet kap a többi érzékszerv és az ún. **vizuális emlékezet** is. Ez utóbbi teszi lehetővé a látványból hiányzó képrészek pótlását, az alakzatok **felismerését**.

A látás az összes érzékelt információnak több mint felét, egyes kutatók szerint akár 90 %-át is szolgáltatja. A különböző érzetek általában keverednek egymással, mivel több érzékszerv együttes működésének eredményei. Pl.: ha egy almát és egy barackot nézünk, akkor a látvány agyunkban ízük, illatuk, simaságuk illetve hamvasságuk érzetével is összekapcsolódik.

A **képérzet** többé-kevésbé a látás nélkül is kialakulhat, pl. alvás során, erős képzeletőreijű embereknél pedig gyakorlatilag bármikor. Ez is azt mutatja, hogy a látás összetett folyamatában a szem csupán az egyik közreműködő szerv.

Ugyanerre utal az is, hogy a látás mennyire összefügg a **tanulással**: szemünk elvileg a megszületéstől kezdve alkalmas lenne a látásban játszott szerepének betöltésére, az újszülött eleinte azonban csak homályos foltokat észlel, melyek a vizuális memória töltődésével, a szem mozgatásának tudatosulásával válnak képéretté.

### 1.3. A látás biológiája

A szem **pupilla**, **szemlencse**, **üveges test**, és **retina** együtteséből álló érzékszerv. A szemlencse izmokkal változtatható fókusztávolságú, kétszeresen domború lencse: a fényrekeszen (*pupilla*) keresztül belépő fénysugarakból az éles képet a **retinán** állítja elő. Az átlátszó **üveges test** homogén törésmutatójú közeg: szerepe elsősorban élettani. A retina tartalmazza a fényérzékelő **receptorokat**, azaz a fényesség érzékeléséért felelős **pálcikákat** (*rodes*) (számuk kb. 120 millió) és a színérzékelésért felelős **csapocskákat** (*cones*) (melyek száma kb. 6 millió). Az ingerületek továbbítása az agy **látóközpontjai** felé többszörös áttételen keresztül, az igen nagy hibátűréssel rendelkező ún. impulzuskódmodulációval (**PCM**) történik.

Nem az összes receptor jele jut el az idegpályákon az agyba; mivel a kb. 126 millió receptorra csak kb. 1 millió idegszál jut.

Ez a redukció egy (egyelőre csak részben értett) retina-szintű **előfeldolgozást** takar: elsősorban az élekre, gyors változásokra vonatkozó adatok jutnak a központi idegrendszerbe.

A leképezés időben gyorsan változó folyamat:

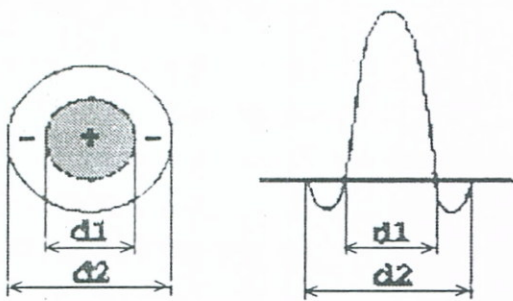


- Agyi parancsra az izmok a szemet (vagy akár az egész testet) úgy mozgatják, hogy a kívánt látvány létrejöhessen. (A szem igen gyorsan képes -a fej elmozdulása nélkül- egyik pontról a másikra „átállni”; így történik agyi parancsra a látvány felépítéséhez szükséges nagymennyiségű adat begyűjtése.)
- A szemlencse fókusz-távolságának és a pupilla nyílásának automatikus változtatása folyamatosan biztosítja az éles képalkotást és a megfelelő fény mennyiséget.

Megjegyzendő még, hogy a szem igen kis mértékben folyamatos rezgő mozgást is végez, mintegy letapogatást biztosítva minden egyes receptor számára. Kísérletek bebizonyították, hogy ez a rezgő mozgás a képalkotás elengedhetetlen része. Ugyanakkor ezzel a rezgő mozgással válik lehetségessé az, hogy látásunk jobb geometriai felbontóképességű, mint amit pusztán a receptorok sűrűsége és a szem optikai tulajdonságai alapján várni lehetne.

Egy rendkívül egyszerűsített modell szerint a receptor-jelek redukciójának lokális jel-erősítés, illetve jel-tiltás az alapja (ez az ún. **lateral inhibition**).

Eszerint: egy kis receptor-folt eredő jelét a folt közepén levő ingerületek **erősítik**,



1. ábra

míg az attól bizonyos távolságra levők **gyengítik**. Az 1. ábra ezt az összefüggést szemlélteti. A „ $d_1$ ” körön belüli jelek erősítenek, míg az azon kívüli, de a „ $d_2$ ” körön belüliek gyengítenek. A jelenség úgy is felfogható, mintha a retina szintjén 2D (kétdimenziós) szűrés történe, mely az él-detektálást, valamint a felületi mintázatok (**textúra**) érzékelését segíti elő.

## 1.4. A látás sajátosságai

A receptorok érzékenysége nemcsak a fény erősségétől, hanem annak hullámhosszától is függ. A látás során a fény hullámhossz szerinti összetételét is érzékeljük: ez **színérzetben** nyilvánul meg.

A csapocskák három csoportba sorolhatók (**P**, **D** illetve **T** típus), melyeket érzékenységük hullámhosszfüggése különböztet meg egymástól.

Az egyes csoportok érzékenységét leíró függvények durván haranggörbével közelíthetők, melyeknek maximumhelye:

- P** típus: 580 nm (vörös),
- D** típus: 540 nm (zöld),
- T** típus: 440 nm (kék).

Megjegyezzük, hogy a férfiaknak kb. 8%-a, a nőknek pedig kb. 1%-a **színvak**, vagy **színtévesztő**: ez a csapocskák teljes, vagy részleges hiányára vezethető vissza.

A retinán a receptorok eloszlása nem egyenletes: a csapocskák elsősorban a látómező közepén (**látógödör**: *fovea*) fordulnak elő, a pálcikák pedig a szélek felé. Ez magyarázza pl. azt, hogy **színlátásunk** a látótér széle felé gyengébb, míg a fényerősség-változásokat a szélek felé érzékeljük jobban.

Az idegpályák kilépési helyén az ún. vakfolton nincsenek **receptorok**, s így az idebeeső fény nem vesz részt az érzékelésben. Az, hogy mégis összefüggő képet látunk, agyunk és vizuális emlékezetünk működésének eredménye.

A receptorok érzékelése és a sugárzás erőssége közti összefüggés logaritmikus, ami azt jelenti, hogy kétszeres fényerősség csak **ln(2)**-szeres fényérzet-erősödést eredményez. Ez a tulajdonság (valamint a **pupilla** tágulása-szűkülése) látásunknak rendkívül nagy dinamikát kölcsönöz: a csillagfényes éjszakától (kb. 10 lux) a tűző napfényig (kb. 100.000 lux) a legszélsőségesebb fényviszonyok közt is látunk.

A látás érzékenysége (finomsága, részletgazdagsága) többféle jellemző együttesétől függ:

- a **geometriai felbontás** azt határozza meg, hogy adott távolságban egymáshoz milyen közel levő pontokat tudunk különálló pontokként észlelni. (Értéke zöld fény esetén kb. 1/2 szögperc; ennél rövidebb illetve hosszabb hullámhosszú fény esetén egyaránt rosszabb.)
- az **intenzitás-felbontás** a fényerősség-változás érzékelését,
- a **színfelbontás** pedig a színárnyalatok elkülönítési képességét adja meg. (Ezekre a jellemzőkre a **színlátás** tárgyalásánál még visszatérünk.)

Külön kell szólni látásunk **időbeli felbontásáról**. Ez azt fejezi ki, hogy mennyi ideig kell egy látványnak tartania ahhoz, hogy különállónak - az előtte és utána következő látványtól különbözőnek - érzékeljük.

Tapasztalataink szerint a kb. 1/15 másodpercnél rövidebb időre „bevillanó” képeket nem tudjuk egymástól elkülönítve érzékelni, pontosabban: **tudatunk szintjén** összefolynak egymással. (A film és a televízió ezt a felbontási korlátot használja ki a mozgás érzetének keltésére: a másodpercenkénti 16-25 állókép (16-25 Hz közötti) megjelenítése számunkra összefüggő, viszonylag villogásmentes mozgóképként jelentkeznek.)

Érdekes, hogy a **tudatalatti képérzékelés** időbeli felbontása a tudatosnál sokkal jobb (60-65 Hz közötti). Ezt régebben - egészen az ilyenfajta reklámtevékenység betiltásáig - ki is használták: pozitív hatású képsorba egy-egy terméket ábrázoló kocka bevágásával a nézőben kimutatható képzettársítást hoztak létre.

Valószínűleg a jobb tudatalatti időbeli felbontásnak köszönhető az is, hogy a hagyományos film és TV nézése, vagy pl. a villogó neonlámpa melletti munkavégzés fá-



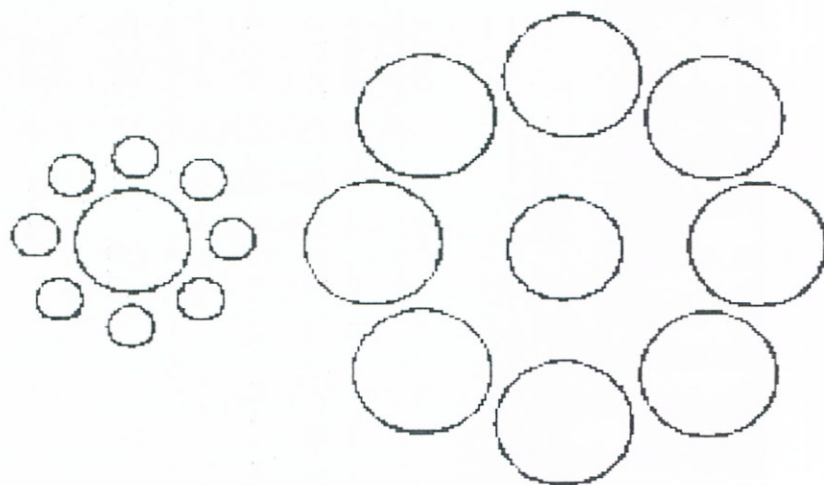
rasztó. Ezért kelt jobb benyomást, pl. a **videodigitalizáló kártya** képe a VGA monitoron, mint ugyanaz a kép a hagyományos RGB monitoron.

Érdekes sajátossága látásunknak a felületek fényesség- (illetve szín-) érzetének viszonylagos állandósága.

Ha pl. egy szürke papírlapot gyengén megvilágított helyiségből napfényre viszünk, továbbra is szürkének látjuk, jóllehet, a papírról visszavert fény mennyisége több nagyságrenddel nagyobb, mint az előbbi helyiségben a fehér papírról visszavert fény mennyisége.

Nyilvánvaló, hogy ezt az érzetet egyedül az érzékelt fény mennyisége alapján nem lehet megmagyarázni. (Természetesen szélsőséges fényviszonyok közt a jelenség nem lép fel.)

Mind a fényesség, mind pedig a szín érzékelése **környezetfüggő**, ami azt jelenti, hogy adott képrészlet által keltett érzetet a szomszédos képelemek fényessége és színe is befolyásolja.



2. ábra

A 2. ábra azt bizonyítja, hogy a fényesség és a szín érzékeléséhez hasonlóan az **alakzatok, méretek** érzékelése is erősen **környezetfüggő**. A középső kör mindkét ábrán **azonos méretű**, eltérő méretűnek csak a szomszédos körök miatt látjuk őket.

## 1.5. Színlátás

A színlátást a csapocskák eltérő spektrális érzékenysége teszi lehetővé. A szem lényegében **integrátorként** működik. Ha egyidejűleg különböző hullámhosszúságú és energiájú (azaz: **spektrális energia-eloszlású**) fénysugarak érik, akkor a csapocskák ingerválaszát a következő összefüggés adja meg:

$$i_j = \int_{380}^{780} I(\lambda) e_j(\lambda) d\lambda$$

$$j \in \{P, D, T\}$$

ahol  $I(\lambda)$  a beeső fény spektrális energia-eloszlása,  $e_P$ ,  $e_D$ ,  $e_T$  pedig a P, D illetve T típus csapocskák érzékenységi függvénye.

**Ha két fényforrás spektrális energia-eloszlása különböző ugyan, de a fenti integrálok azonos értéket adnak, akkor fényüket nem tudjuk megkülönböztetni.**

Az agyunkban kialakuló színérzetnek három jellemző sajátossága van.

1. **színezet** (*hue*) - melyet a hétköznapi életben pontatlanul **színek** mondunk - a fény hullámhosszától függ. Szemünk kb. 200 féle színezetet tud megkülönböztetni.
2. **telítettség** (*saturation*) attól függ, hogy mekkora a **fehér** fény összetevője a többi összetevőhöz képest. (A **monokromatikus** színek nem tartalmaznak fehér összetevőt, azaz 100 %-os telítettségűek, míg pl. a rózsaszín néhány % telítettségű vörös szín.) Szemünk egy adott színezetben kb. 20 telítettségi fokozatot tud megkülönböztetni.
3. **világosság** (*intensity*) az egységnyi térszögben szemünkbe érkező fényenergia mennyiségétől függ. Szemünk világosság szerinti felbontóképessége erősen hullámhosszfüggő: átlagosan mintegy 500 fokozatot tudunk megkülönböztetni.

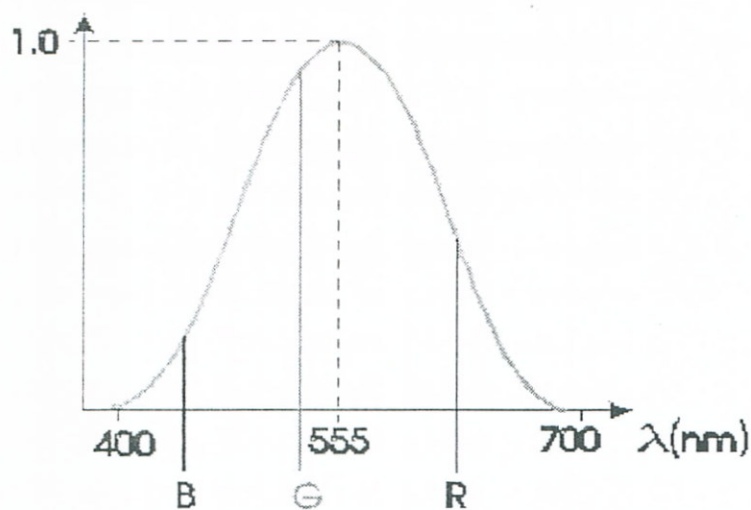
A képfeldolgozásban gyakran használt - a fentiekhez hasonló - színteret ad a színezet (**Hue**), telítettség (**Saturation**) és a fényerő (**Brightness**) - **HSB**.

Amint azt a színérzékelés alapjainál említettük, szemünk nem tudja külön érzékelni az egyidejűleg beérkező, különböző hullámhosszúságú és energiájú fényimpulzusokat, csak ezek **eredőjét**. (Érdekességként megemlíjtük, hogy hallásunk ettől eltérően működik, s pl. kis gyakorlattal egy zenekari akkordban minden hangszer hangját külön-külön is érzékelni tudjuk.)

Adott színérzetet különböző összetevők keverésével - többféle úton - is elérhetünk. A színérzetünk tehát **nem objektív** érzet.

A színkeverés két alapvető módja:

1. **additív színkeverés**: a keverékszín monokróm fényforrások fényének **egy-másra vetítésével** áll elő. Így működik pl. a TV képcső. Az additív színkeverés alapszínei a vörös, zöld és kék (**RGB**, az angol megnevezés első betűiből)
2. **szubtraktív színkeverés**: széles sávban sugárzó fényforrás fényéből „**kivonunk**” egyes hullámhosszakokat (vagy tartományokat); a keverékszín az lesz, ami megmarad. Ez az alapelve pl. a színes nyomtatásnak. A szubtraktív színkeverés alapszínei a sárga (Y), magenta (M), cián (C) (az angol megnevezés első betűiből) és a fekete (Key) - **CMYK**.



3. ábra

A világosság érzete erősen hullámhosszfüggő, azaz azonos energiájú, különböző hullámhosszú fénysugarakat különböző világosságúnak látunk. Az alábbi ábra nagyszámú mérésen alapuló ún. **láthatósági függvényt** mutat. Mint látható, szemünk a **zöld** fényre a legérzékenyebb.

Az, hogy egy adott tárgyat milyen színűnek érzékelünk, egy sor „külső” tényezőtől is függ. **Például:** a búzavirág a ráeső napsugárzás teljes spektrumából a kék és az égszínkék sugarakat veri vissza a legnagyobb mértékben, ugyanakkor elnyeli a piros és sárga sugarakat - ezért **kéknek látjuk**. Ha azonban piros fényel világítjuk meg, feketének, illetve sötétpirosnak fog látszani. Ugyanígy pl. az emberi arc színe gyertyafényben rendellenesen sápadt, sárga lesz.

Általában a tárgyak a rájuk eső fénysugárzás teljes spektrumát módosítják, felületük fényvisszaverő-képességének hullámhossztól való függése, az ún. **spektrális reflektancia-függvény** szerint. A tárgyak felületéről szemünkbe jutó fény spektrális összetétele a fényforrás spektrális összetételét kifejező függvény, valamint a tárgy spektrális reflektancia-függvényének szorzataként alakul ki.

Agyunk ezt az összefüggést jól kompenzálja, s így a tárgyakról alkotott színérzetünk szélsőséges fényviszonyok közt is viszonylag állandó.

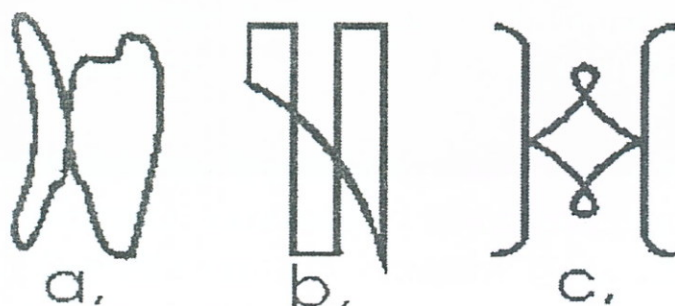
Agyunk ezt az összefüggést jól kompenzálja, s így a tárgyakról alkotott színérzetünk szélsőséges fényviszonyok közt is viszonylag állandó.

## 1.6. Alakzatlátás

Az agyunkban lejátszódó vizuális információfeldolgozási folyamatok első lépése az agykéregben található. Ezért lát „csillagokat” az, kit fejbe vágunk. Agyunk közel egyharmada foglalkozik a látás információinak a feldolgozásával. Először a **mozgás** feldolgozása történik, majd az **alakzatra** és a **térre** vonatkozó információkat nyerjük ki, végül a **szín** információ kinyerése történik. A folyamat végén az agy újra egységesíti a szeparált információt, azonban hogy ezt hogyan végzi az rejtély.

A tárgyakat nem képpontok halmazaként, hanem **térbeli alakzatokként** érzékeljük. Agyunk lényegkiemelő tevékenységének köszönhetően az egyes fényingerek - a vizuális emlékezetből való kiegészítésekkel - **foltok, élek** rendszerévé alakulnak.

Agyunk **alakfelismerő** képessége is rendkívüli: a tárgyakat - méretüktől, állásuktól, színüktől, stb. függetlenül - akár kis részleteikből is szinte hihetetlen biztonsággal



4. ábra

ismerjük fel. (A felismerés sebessége jelenlegi legjobb „látó” eszközeinknél kb. 2-3 nagyságrenddel nagyobb.)

Kísérletek bizonyítják, hogy „**képi adatbázissal**” rendelkezünk, mely mintákat (*prototípusokat*) tárol. A felismerés **mintaillesztéssel** történik. A folyamat visszacsatolást is tartalmaz: a fel nem ismert részletekről „erősebb odanézéssel” próbálunk több információt szerezni.

Az **alakzatfelismerés** során fokozatosan haladunk a mind bonyolultabb részletek felé úgy, hogy az alapvető jellemzők (**szín, alak, mélység, mozgás**) tekintetében a hasonló alakzatokat összekapcsoljuk.

#### A főbb alakzat-csoportosítási szabályok:

- **közelség és hasonlóság:** minél közelebb vannak egymáshoz és minél azonosabb méretűek, annál inkább vonjuk egy csoportba őket;
- **zártság:** az egy görbével körülzárt alakzatokat könnyebben egyesítjük; Pl.: a 4./a ábrán nem önmagát metsző görbét, hanem inkább két, érintkező zárt területet látunk;
- **sima folytonosság:** olyan csoportokat képezünk, melyeknek a lehető legsimább a kontúrvonaluk. Pl.: a 4./b ábrán nem három zárt területet, hanem négyszöghullámot és egy sima görbét látunk.
- **szimmetria és ismerősség:** lehetőleg szimmetrikus csoportokat képezünk és lehetőleg ismert elemekből. Pl.: a 4./c ábráról el tudjuk képzelni, hogy egy **W** és egy **M** betűből áll, de nem látjuk.

A **mikroszerkezet** (*textúra*) **tudatos** felismerésének szabályszerűségei hasonlóak az alakzatok felismeréséhez.

A **nem tudatos** felismerési folyamatban a textúra különböző rendű **statisztikái** (**együttes valószínűségi eloszlásai**) játszanak szerepet.

Egy textúra n-edrendű statisztikája arról tájékoztat, hogy ha véletlenszerűen rádobunk minden lehetséges méretű és alakú n-oldalú sokszöget, akkor milyen valószínűséggel esik minden csúcs egy meghatározott világosságú illetve színű képpontra.

Kísérletileg bebizonyították, hogy az olyan textúrákat, melyeknek első- és másodrendű statisztikájuk megegyezik és csak a harmad- vagy magasabbrendű statisztikájuk különböző, nem tudatos megfigyelés esetén általában **azonosnak látjuk**.

Eddig háromféle texton-típust sikerült találni, ezek:

- a **szín**;
- adott irányú és szélességű, illetve hossz/szélesség arányú **vonaldarabok**;
- ezen vonaldarabok **végpontjainak száma**.

## Ellenőrző kérdések

1. Milyen hullámhosszúságú elektromágneses sugárzás érzékelésére alkalmas az emberi szem?
2. Hogyan történik a képérzet kialakulása?
3. Mi az agy szerepe a képérzet kialakulásában?
4. Mi a szerepe a látásban a szem egyes részeinek?
5. A szem érzékelő receptorai közül melyeknek van szerepe a fényérzékelésben, és melyeknek a színlátásban?
6. Mit jelent a lateral inhibition?
7. Milyen jellemzőktől függ a látás érzékenysége?
8. Mit jelent a látás időbeli felbontása?
9. Mit jelent a látás környezetfüggősége?
10. Mi a színérzet három jellemző sajátossága?
11. Mi a különbség az additív és szubtraktív színkeverés között?
12. Mi jellemzi az emberi alakzatlátást?
13. Melyek a főbb alakzat-csoportosítási szabályok az emberi látásban ?
14. Hol történik a vizuális információ feldolgozásának első lépése az agyunkban?
15. Milyen texton típusokat ismer?
16. Mi az RGB, CMYK és HSB jelentései?

## 2. A KÉPFELDOLGOZÁS ESZKÖZEI

### 2.1 Bevezetés

A (számítógépes) képfeldolgozás a szokásos - egyéb célokra általánosan elterjedt - számítástechnikai eszközökön túlmenően **speciális eszközök** használatát is igényli. Ebben a fejezetben rövid ismertetést adunk ezekről az eszközökről, de nem kívánunk a technikai részletekben elveszni, célunk inkább a működési elvek megismertetése. Hallatlanul gyors ezen a területen a fejlődés, ezért talán ez könyvünk leggyorsabban elavuló része. Tekintettel arra, hogy a távérzékelési alkalmazásokkal kicsit részletesebben foglalkozunk, az ezen a területen használt eszközökre a 9. fejezetben térünk ki.

### 2.2. Képbevitel

• **A képbeviteli eszközök feladata:** valamely objektumról, térrészről azok képének megfelelő, számítógép által kezelhető formátumú adatok előállítás. Az adatok többnyire **képpontok**, melyek **raszteres szervezésben** - pl. sorról-sorra - tartalmazzák a fényesség-, illetve színesség- értékeket.

A képbeviteli eszközök **képérzékelők** és **jelátalakítók** (illesztők, kódolók) együttesét jelentik. Az első csoportba tartoznak pl. a TV kamerák; a másodikba pl. a képdigitalizáló kártyák. Egyes készülékek az érzékelőket és a jelátalakítókat együtt tartalmazzák, pl. ilyenek a szabványos adatvonalra dolgozó szkennerek vagy a digitális kamerák.

A következőkben a főbb képbeviteli eszközöket soroljuk fel; a teljesség igénye nélkül, elsősorban működésük alapjainak ismertetése érdekében.

A képbevitel első lépése: **a kép érzékelése**. Ez - eltekintve a képfeldolgozás egyes speciális felhasználási területeitől, melyek más kölcsönhatásokon alapulnak - a látható tartományba eső elektromágneses hullámok, azaz a **fény** érzékelését teszi szükségessé.

A képérzékelés lényege röviden:

- az ún. **elsődleges fényforrás** által megvilágított objektumok a beeső fény egy részét visszaverik (azaz ún. **másodlagos fényforrássá** válnak). A visszavert fény információt szállít helyükről, alakjukról, felületük különböző jellemzőiről;
- a visszavert fény egy részét optikai eszközökkel (pl.: objektívvel) leképezve éles kép jön létre, melyet
- a **képérzékelő** eszköz érzékel, s pontról-pontra letapogat.



A **képzékelő eszközöket** többféleképpen lehet csoportosítani:

**1. Csoportosítás az eszköz feladata alapján:**

**Síkban** elhelyezkedő objektumok (fényképek, iratok, átvilágítható filmek) mintázatának, képének bevitele.

Néhány alapvető eszköz-típus:

**Kézi (handy) szkennер** kis szélességben, egyidejűleg egy teljes fekete-fehér, illetve színes sort olvas be; a sorra merőleges irányban kézzel kell egyenletes sebességgel húzni. Beépített fényforrást tartalmaz, mely a készülékkel együtt mozog. Fekete-fehér és színes változatuk egyaránt létezik. Felhasználási köre: kisebb írott dokumentumok (számlák, csekkek, aláírások, stb.) beolvasása.

**Síkágyas (flat-bed) szkennер** nagyobb (tipikusan **A4**, illetve **A3** méretű) fekete-fehér, illetve színes dokumentumok beolvasására alkalmas. Többnyire a visszavert fény érzékelésével működik, de létezik az átmenő fény felhasználásán alapuló típus is. Beépített fényforrást tartalmaz, a papír - vagy egyes típusoknál a fényforrás és az érzékelő sor - mozgatása automatikusan történik. Felhasználási köre: dokumentumok beolvasása.

**Dob-szkennер** nagyméretű fekete-fehér, illetve színes dokumentumok, filmek igen jó minőségű beolvasására alkalmas. Többnyire átmenő és visszavert fénnel egyaránt működik. A képet dobfelületre kell rögzíteni, amelyet a beolvasáshoz rendkívül nagy fordulatszámmal megforgatnak. Eközben változtatható finomságú előtolást is alkalmaznak, így a pontszerű fényforrás és az ugyancsak pontszerű érzékelő a képet spirális pályán tapogatja le. Felbontása és visszaállási pontossága 10  $\mu\text{m}$  nagyságrendű. Felhasználási köre: professzionális képfeldolgozás (DTP, úrfelvételek digitalizálása).

**Grafikus vagy digitalizáló tábla** koordinátarendszerrel ellátott elektronikus rajztábla. Általában soros porton csatlakozik a számítógéphez. A készülékhez tartozik egy elektronikus ceruza vagy szátkereszttel ellátott korong. Működési elve: A táblában vékony elektromos huzalokból álló háló található. A metszéspontok egy raszteres felületet alkotnak. A raszteres felület segítségével, a kapcsolóval ellátott ceruza vagy szátkereszt helyzete megadható az általa keltett elektromos impulzus futásideje alapján. A raszterpontok közötti érzékelés esetén interpolációval állapítják meg a ceruza vagy szátkereszt pontos helyét. A pontok illetve a pontokat összekötő vonalak segítségével előállított, tetszőleges rajz közvetlenül a számítógébe jut. Tervrajzok, térképek és egyéb vonalas dokumentumok digitalizálására alkalmazzák.

**Térben elhelyezkedő objektumok** képének bevitele.

A legalapvetőbb eszköz-típus:

**TV kamera** objektívvel: a képérzékelő síkjára leképezett fekete-fehér, vagy színes képet érzékelő és pl. videojellé alakító eszköz.

Különböző kiegészítő eszközökkel (pl.: érzékenység és élesség-állító automatika, zoom, illetve: képmagnetofon) ellátott változatai a mindennapi életbe is bevonultak. Állóképek készítésére alkalmas a **digitális kamera**. Jelentőségük és elterjedésük napjainkban rendkívüli méreteket öltött. Felépítésüket tekintve a leképezés hasonló a hagyományos fényképezőgépekhez, azaz lencserendszer (általában autofókuszos) vetíti a képet a CCD érzékelőre. Majd megtörténik a képpontok kiolvasása, digitalizálása. A kamerát általában saját állandó memóriával ellátott mikroprocesszor vezérli. A processzor elvégzi a kép korrekcióját (leképező rendszer okozta hiba javítása, automatikus fehéregyensúly állítás, gammaérték állítás, stb.), tömörítését és tárolását. Az elkészített képek egy LCD képernyőn közvetlenül megtekinthetők, módosíthatók esetleg törölhetők. A kép mentésével a felvételi körülmények is (dátum, expozíció, vakuhasználat, stb.) tárolásra kerülnek. A fényképezőgép közvetlenül csatlakoztatható szabvány felületen számítógéphez és/vagy TV készülékhez és/vagy nyomtatóhoz. A kamerák felbontása általában 1-2 Megapixel körüli. Az érzékelő és a beépített szolgáltatások erősen ár-függőek. A 15. fejezetben a témával részletesen is foglalkozunk.

A **3D szkennerek** az eddig ismertetett szkennerekkel ellentétben - amelyek csak kétdimenziós digitális felvételezésre alkalmasak - a háromdimenziós felületek digitális rögzítésére használható.

A jól megvilágított, forgatható asztalra helyezett tárgyról CCD kamerával - amely előzőleg az asztalhoz kalibrált - felvételeket készítünk, miközben a tárgy körbefordul. A kamera képe egy komputerhez csatlakozik, amely elvégzi a képek mintavételezését (3. fejezet), rekonstrukcióját és közvetlenül előállítja a textúrákat is tartalmazó háromdimenziós digitális felvételt.

## 2. Csoportosítás az érzékelés elve alapján:

**Hagyományos, elektroncsöves képérzékelővel** működnek pl. a **vidikon**, **plumbikon** stb. csöveken alapuló kamerák.

Működésük alapelve: az optikai úton fényérzékeny rétegre leképezett képnek megfelelő ún. **töltésképet** megfelelően vezérelt elektronsugárral letapogatják, s a pontról-pontra változó töltés kisülésével arányos analóg jelből (sor- illetve kép-szinkronjel hozzáadásával) **videojelet** állítanak elő.

E kamerák többnyire mágneses elektronsugár-eltérítéssel működnek, s így a

külső mágneses térre is érzékenyek. Képkalkotásuk geometriai jellemzői erősen függenek az eltérítő áramkörök kivitelezésétől, s a tápfeszültség ingadozása, sőt, az alkatrészek öregedése miatt is változhatnak. Napjainkban jelentőségük - a CCD eszközök megjelenése következtében - egyre csökken.

**Töltéscsatolt félvezető érzékelők** (angolszász megnevezése - *Charge Coupled Devices* - alapján: **CCD**) alkalmazásán alapuló kamerák.

Az optikai úton leképezett kép itt is a töltés mennyiségének változását okozza; ebben az esetben azonban az analóg jelek kiolvasása az egyes képelemeknek megfelelő töltéscsomagok elektronikus léptetése útján történik. Egy-egy képelemnek több tranzisztor együttese felel meg, melyek a feltöltés - érzékelés - töltéstovábbítás állapotoknak megfelelően működnek.

**CMOS félvezető eszközök:** CMOS (*Complimentary Metal Oxide Semiconductor*) fényérezelők elve azonos a töltéscsatolt félvezetőkével: az optikai úton leképezett kép itt is a töltés mennyiségének változását okozza. Az elektronikai eszközök integráltságának növekedésével és a CCD érzékelők viszonylag magas árával magyarázható, hogy a CMOS érzékelők fejlesztése napjainkban újra előtérbe került. Neves gyártók (Intel, Texas Instruments, Toshiba, Rockwell, AT&T, Polaroid, stb.) számos ilyen érzékelőt fejlesztettek ki és forgalmaznak. A CMOS érzékelők előnyei közé tartozik, hogy kicsi az energiafogyasztásuk, olcsón előállíthatók és bizonyos jelfeldolgozási feladatok (expozíció korrekció, leképező rendszer torzításainak korrekciója, fehéregyensúly állítás, gamma korrekció, zajszűrés) közvetlenül az érzékelő chipbe épített processzor segítségével digitálisan elvégezhetővé válik. A kép minősége is jó, felveszi a versenyt társaival. Várhatóan a CCD mellett a közeljövő kedvelt és elterjedt érzékelőjévé válik.

### 3. Csoportosítás az érzékelés párhuzamossága alapján:

**Egyetlen érzékelőt** és kiegészítő letapogató mechanizmust tartalmaznak pl. a professzionális forgó-dobos szkennerek, vagy egyes műholdas képkalkotó eszközök. Itt az érzékelő paramétereinek állandósága a teljes képkalkotás folyamatában biztosított, azaz az eltérések kompenzációjára nincs szükség. Érzékelőként pl. fotoelektron sokszorozót alkalmaznak, amely nagy sáv szélességgel, jó érzékenységgel és dinamikával jellemezhető.

**Vonali érzékelőt** tartalmaznak pl. az optikai lapolvasók. Ezeknél egy-egy sor minden egyes pontjához külön érzékelő tartozik. Az egyes érzékelők a jelenlegi technikai szinten százalékos pontossággal azonos érzékenységgűek; tökéletesebb azonosság a működtető programba épített kiegyenlítéssel valósítható meg. A vonali érzékelőt 2D képkalkotáshoz kiegészítő mechanizmussal kell a

vonalra merőleges irányban mozgatni. Az érzékelők mellé megvilágító rendszert is integrálnak.

**2D (azaz kétdimenziós) érzékelőt** tartalmaznak pl. a CCD videokamerák, illetve a hagyományos (elektroncsöves) kamerák. A kép minden egyes pontjához külön érzékelő tartozik. A képalkotás egyidejű, de az egyes pontokhoz tartozó adatok kiolvasása egymás után (általában sorfolytonosan) történik.

#### 4. Csoportosítás az érzékelés színelbontása alapján:

**Fekete-fehér** képérzékelők a látható fény tartományában egyenletes érzékenységgel érzékelik a fényességet, s minden képpontról egyetlen adatot adnak. Ez az adat a látható tartományban érzékelt jelek **integráljaként** fogható fel.

**Színes** képérzékelők a látható fényt több tartományra bontva érzékelik a fényességet, s minden képpontról több adatot szolgáltatnak. Ezek az adatok az egyes tartományokban érzékelt jelek erősségével függenek össze, s általában az adott pontbeli „**piros**”-ságot, „**zöld**”-séget és „**kék**”-séget jellemzik.

A színes képalkotás az eszköz felépítésétől függően történhet színváltással (pl.: olyan szkennerek esetében, melyek színszűrők cserélgetésével dolgoznak), vagy egyidejűleg (pl.: a képet alapszínekre bontó, s három képérzékelőt tartalmazó színes kamerák esetében).

A **jelátalakító-kódoló eszközök** alatt többnyire azokat a hardver eszközöket értjük, melyek a képérzékelő eszköztől analóg, vagy digitális formában érkező adatokat a számítógépbe csatolják. Fontos szerepe miatt ezek közül a **videodigitalizáló** kártyákat érdemes kiemelni. Ezek bemenetükön egy vagy több, szabványos videojelet tudnak fogadni. A jelek digitalizálása a számítógép felől vezérelhető módon történik, a digitális kép általában a kártya saját memóriájába kerül. Ez a kép a számítógép felől olvasható, módosítható. A kártyák többsége megjelenítésre alkalmas kimenőjelet is szolgáltat.

A **videodigitalizáló kártyáknak** fontos jellemzői:

- geometriai, illetve intenzitás felbontásuk;
- időfelbontásuk (ami a lassú digitalizálástól a videojel szabvány által biztosított képváltási gyakoriságig terjedhet; ez utóbbit nevezzük valós idejű [**real-time**] képdigitalizálásnak);
- bemeneteik és kimeneteik száma;
- színességük;
- esetleges képfeldolgozási műveletvégző képességük.

A videodigitalizáló kártyák felépítésében és működésében az elmúlt évtized két generációváltást is hozott.

Az első az 1980-as évek végén következett be, amikorra az addig külön dobozban, több kártyán megvalósított eszközöket felváltották az egyetlen, PC-ben elhelyezhető digitalizáló kártyák. Ezek a kártyák (frame grabberek) közvetlenül a PC sínjére csatlakoznak és a felhasználó a képet tároló memóriát a PC memóriájának kiterjesztéseként érheti el. A legismertebbek az amerikai Imaging Technology és Data Translation, valamint a kanadai Matrox cég termékei. Ezeknek a képkártyáknak a képmemóriája 512\*512-től 2048\*2048 képpontig terjed. Maga a digitalizálás eléri a TV szabványoknak megfelelő 768\*576-os felbontást. A nagy képmemória előnye a több kép egyidejű tárolásában és a más forrásból származó nagyméretű képek (pl. úrfelvételek) kényelmes kezelésében jelentkezett. Ez különösen fontos volt, amikor a PC-k saját memóriája csak néhány száz kbyte volt. Ezeknél a kártyáknál rendszerint beépített hardver tette lehetővé a valós idejű nagyítást (zoom), valamint - elég nagy frissítőtár esetén - a függőleges görgetést (scroll) és a vízszintes úsztatást (pan). Számos esetben néhány bites külön grafikus képsíkot is alkalmaztak, ezért a grafikus információt meg lehetett jeleníteni a kép felett, a képtartalom változtatása nélkül (overlay technika). Ezeknek a kártyáknak közös jellemzője, hogy szabvány videó kimenettel rendelkeztek, így a kép egy második videó monitoron jelenik meg.

A következő váltás a PC-k grafikus kártyáinak paramétereiben bekövetkezett fejlődés eredménye. Amikor a VGA kártyák felbontása elérte a 640\*480 képpontot és a 256 színt, már várható volt, hogy megjelennek a hozzájuk csatlakozó digitalizáló kártyák is, melyekkel megtakarítható lesz a kép megjelenítéséhez korábban nélkülözhetetlen második monitor. Ezt a folyamatot erősítette a multimédia koncepció kialakulása, amelynek egyik sarkalatos pontja a képkezelő eszközöknek - és így az analóg be/kimeneteknek - a beépítése a számítógépbe.

Az első - már valóban használható - VGA kártyához kapcsolódó digitalizálót a német FAST cég 1990-ben valósította meg és dobta piacra Screen Machine néven. A kártya lelke egy speciális A/D áramkör, amely a bejövő színes kompozit jelet digitalizálja és YUV rendszerbe kódolja. A világosságinformációt (Y) 8 bitre bontja, a színinformációt (UV) 4-4 bitre tömöríti, ezáltal egy színes képpont 12 biten tárolható, az RGB rendszerben szükséges 24 bit helyett. Az információveszteség a színekben lévő redundancia miatt a látványképben nem érzékelhető, a képminőség tökéletesen megfelel a multimédia eszközökkel szemben támasztott követelményeknek. Kimeneti oldalon a kártya a VGA monitor számára RGB jelet állít elő, emellett a VGA kártya jelét is tudja fogadni és a monitor felé a két képforrás jelét a felhasználó által programozható módon képes kiadni. A VGA kártyán előállított grafikus képen bárhol lehet ablakot nyit-

ni, amelyben megjeleníthető az éppen folyamatosan digitalizált, a befagyasztott, vagy a korábban letárolt és visszahívott digitális kép.

Csak egy évre volt szükség, hogy az amerikai CHIPS cég egy hasonló eszközzel jelenjék meg a piacon, ráadásul - az áramkör értékesítésében látva nagyobb fantáziát - szabad utat nyitott a távol-keleti gyártóknak az erre alapozott képkártyák előállítására. Ennek köszönhetően a Screen Machine-nal azonos elven működő és egymással teljesen azonos felépítésű kártyák elárasztották az egész világot. Magyarországon a tajvani Aver Media Aver2000, és a szingapúri Creative Technology VideoBlaster digitalizálója valamint a Miro cég DC sorozatú eszközei terjedtek el leginkább. Bár ezek a kártyák kedvező árak révén elősegítik a képfeldolgozási alkalmazások elterjedését és számos esetben valóban jól használhatóak, fel kell hívni a figyelmet arra, hogy a speciális kódolás miatt a programozásuk körülményesebb, emiatt megnő a feldolgozási idő és a digitalizálás minősége is gyengébb, mint a professzionális képdigitalizáló kártyák esetében.

Az elmúlt években, a PC-k sebességének és az alkalmazott memória méretének növekedésével feleslegessé vált a digitalizáló kártyák külön képmemóriája, a PCI buszon keresztül a digitalizált kép közvetlenül a PC memóriájába kerülhet. Ma már a 24 bites színes digitalizálás a jellemző, és sokszor a digitalizáléhoz hozzáépítik a TV tunert, a szabvány videójel kimenetet, és esetenként a különböző képtömörítő áramköröket is. Így ezek a kártyák TV műsorok megjelenítésére, multimédia anyagok és videofilmek lejátszására, szerkesztésére, előállítására is alkalmasak. Jól érzékelhető, hogy a tömeges felhasználási igény határozza meg a fejlesztési irányokat. A klasszikus értelemben vett képfeldolgozási feladatokra szánt digitalizáló kártyák ára egy nagyságrenddel magasabb, mint a multimédia kártyáké, de mivel ma már a képminőségben nincs lényeges különbség, ezért néhány kivételes esettől, speciális alkalmazástól eltekintve nyugodtan alkalmazhatjuk a multimédiás digitalizáló kártyákat képfeldolgozási feladatok megoldásában is.

Napjainkban egy új technológia, a digitális kamerák megjelenésének vagyunk tanúi. Bár a képfeldolgozási ismeretek szempontjából nem jelentenek elvi újdonságot, de várható elterjedésük és technikai érdekességük miatt a 15. fejezetben röviden foglalkozunk a digitális fényképezőgépekkel.

### 2.3. Képtárolás

A számítógépes képfeldolgozás általában jelentős háttértár kapacitást igényel. Ennek szemléltetésére lássunk két példát:

- Egy 300 dpi (dot per inch azaz képpont inch-enként) felbontású síkágyas szkennerral beolvasott **A4** méretű színes (24 bites) dokumentum 25 MByte adatot eredményez.

- Egy Landsat TM űrfelvétel - 7 sávban készül egy időben felvétel ugyanarról a területről - 185x170 km-es területről készít 6 képet 30 m-es és egy képet 120 m-es terepi felbontással. A teljes kép mérete:  $(6 \times 33,3 \text{ Mbyte}) + (1 \times 2,1 \text{ Mbyte}) = 202 \text{ Mbyte}$ .

Még a részletekkel való műveletvégzés is szükségessé teszi, hogy a rendszer nagy tárolókapacitású, gyors hozzáférésű operatív,- és háttértárral rendelkezzen.

A számítógépes képfeldolgozási célokra használt memória három csoportba tartozhat:

1. A feldolgozás (műveletvégzés) idején működő, rövid elérési idejű memória.
  2. Nagy tárolókapacitású, gyors hozzáférésű háttértár, melyet a rendszer on-line hozzáféréssel használ.
  3. Archiválási célra szolgáló, off-line felhasználású háttértároló.
1. A rövid elérési idejű memória maga a számítógépben levő **főtár (RAM)**, vagy a grafikus egységek, - pl. **videódigitalizáló kártyák, 3D gyorsító kártyák vagy a képernyő megjelenítő kártyák** - memóriája. Mivel alapvetően meghatározza a műveletvégzés sebességét, napjainkban kifejlesztésre és bevezetésre került, fontosnak ítélt megoldásokkal részletesen is foglalkozunk:

- Az **AGP (Accelerated Graphics Port)** egy nagysebességű, a PCI buszrendszer átviteli kapacitását jelentősen meghaladó, kimondottan grafikus adatátvitelre kifejlesztett szabvány interfész. Az AGP grafikus kártya nem a PCI buszrendszerre kapcsolódik, hanem egy tőle független AGP buszra, amely közvetlen a Chipset irányítása alatt áll. Saját memóriával rendelkezik (általában 2-32 Mbyte), amely az egyes alkalmazások esetén nem jelent megszorítást, mivel a textúrákkal kapcsolatos számításokat a főtár segítségével is elvégezheti. Az átviteli kapacitás 32 bit esetén maximum 533 Mbyte/s lehet, szemben a PCI 132 Mbyte/s maximális értékével. További összehasonlításokat az alábbi táblázat tartalmaz:

AGP	PCI
Pipeline átvitel	Nem pipeline átvitel
Adatok/címek demultiplexeltek	Adatok/címek multiplexeltek
533 Mbyte/s maximális átvitel 32 biten	132 Mbyte/s maximális átvitel 32 biten
Egyetlen, speciális kivitelű port	Több, PCI szabványú port
Csak memória írás/olvasás engedélyezett	Tetszőleges egység írása/olvasása engedélyezett
2D/3D támogatások (textúrák esetén közvetlen főtár használat. z-buffer, alpha-buffer)	2D/3D támogatások általában külön eszköz segítségével

- Az **UMA** (Unified Memory Architecture) a számítógép különböző eszközeinek egységesített, az egyetlen memóriát (főtárat) a Chipset-en keresztül közvetlenül használó elnevezése. Nincsenek az egyes funkciókat ellátó eszközöknek önálló memóriái, mint amilyen a hagyományos képernyővezérlők, hálózati csatolók, stb. saját memóriája. Költségtakarékos, rendkívül magas átviteli kapacitást - 2-4 Gbyte/s - tesz lehetővé. Megvalósítása azonban egyedi, a hagyományos számítógépes architektúrától eltérő megoldásokat jelent. Ilyen belső struktúrával rendelkezik a Silicon Graphics O2 elnevezésű grafikus munkaállomása.

2. On-line tárolásra rendszerint **fix mágneses adathordozót** (merevlemez: ún. *Winchester*) használnak. Ennek oka egyszerű: a többi alkalmazási terület igényeinek kielégítésére ezek a háttértárak nagy fejlettségi fokot értek el, s az árak is elég alacsony ahhoz, hogy a speciálisan képtárolási célra kifejlesztett eszközöket kiváltsák. Nyilvánvaló, hogy amennyiben a háttértároló kifejezetten csak képek tárolására szolgálna, az eltérő adatábrázolás, a hozzáférés-szervezés stb. növelhetné a hatékonyságát. Ez azonban egyidejűleg növelné az árat is, hiszen lassan megtérülő fejlesztést igényelne. Ehelyett a közbenső (*cash*)-memóriával rendelkező, korszerű merevlemezek használata lett napjainkban az általános on-line adattároló eszköz. Ezek kapacitása jelenleg már 100 Gbyte körüli.

A merevlemezes háttértárak számára egyre nyilvánvalóbb vetélytársat jelentett a közelmúltban megjelent **magneto-optikai tároló (MO)**. Ennél a tárolási módnál a információ mágneses hordozó réteg mágneses polaritását változtatják meg lézeres felírással. Előnyük, hogy hosszú ideig (30 év) megőrzik az információt és írható, olvasható tárolást tesznek lehetővé.

Szinte minden alkalmazásnál archiválási célokra a **CD-ROM** alapú tárolók terjedtek el, miután a felhasználó által írható meghajtók megjelentek a piacon. Ez a felírási mód korábban csak egyszeri felírást tett lehetővé, az adatok felírás után már nem voltak módosíthatók. Ezek az eszközök valamilyen hordozóra felvitt réteg vékony lézersugárral való beírásán/kiolvasásán alapulnak. A hordozó általában korong, amelyet megforgatnak, de létezik kártya kivitelű eszköz is. A közelmúltban jelentek meg az újraírható CD-ROM-ok új perspektívát nyitva ennek a tárolási módnak, ami egyben a korábbi MO típusú tárolóknak jelent versenytársat.

A magneto-optikai és CD-ROM tárolóknak hallatlan előnyt jelent a szokásos merevlemezes tárolókkal szemben az, hogy az adathordozó lemez a meghajtóból kivehető, azaz cserélhető. Mivel viszonylag olcsó, nagy élettartalmú, sérülésekkel szemben ellenálló a média, további széleskörű elterjedése biztosra vehető.



Léteznek egyszer írható, akárhányszor olvasható optikai lemezek (*WORM*) 6-12 Gbyte kapacitású kivitelben is.

A legkorszerűbb, tárolási kapacitásban a CD-ROM-ot jelentősen felülmúló, de vele kompatibilis megoldás a DVD, mellyel a CD mellékleten található alkalmazások között önálló fejezetben is foglalkozunk (lásd DVD).

Napjainkban sorozatgyártásra alkalmas holografikus elven működő írható-olvasható tárolóeszközöket is kifejlesztettek. Ezek kapacitása és elérési sebessége közel 1.5-2 nagyságrenddel nagyobb a DVD-nél.

3. Off-line adattárolásra napjainkban az ún. jukeboxok használata terjedt el.

Létezik olyan típus, mely 30-100 db optikai lemezt tartalmaz, s ezek váltása programvezérléssel történik. (Az elnevezés a jól ismert zenegépekre utal, ahol számos hanglemez közül lehet választani.) Ezek általában igen drága, nem hétköznapi felhasználásra készített eszközök; elsősorban központi adatbázisokban van létjogosultságuk.

A korai rendszerekben használatos hagyományos mágnesszalagos adattárolás méret-, adathozzáférés- és élettartam-problémák miatt egyre inkább a háttérbe szorulnak. Jelenleg ezek digitális változata (DAT) terjedt el, melyek kapacitása napjainkban meghaladja a 10 Gbyte-ot.

A számos egyéb off-line tárolási megoldás is létezik:

- Cserélhető lemezes Winchester.
- Analóg mágneses képrögzítés.

## 2.4. Képfeldolgozás

A **képfeldolgozás** olyan műveletek együttes megnevezése, melyek a digitális képet bemenő adatokként használják, kimenő adatokként pedig képet, illetve azzal összefüggő adatokat szolgáltatnak.

E **műveletek elvégzésének** két oldala van:

- *hardver*: műveletvégző processzor illetve processzorok együttese, és
- *szoftver*: működtető programok együttese.

Természetesen e két oldal olykor nehezen választható el egymástól, pl.: az ún. **célhardverek** (azaz meghatározott feladat megoldására készített, nem programozható eszközök) esetében.

A következőkben röviden ismertetjük a képfeldolgozásban használatos **processzorok**, **szoftverek** és **célhardverek** legfontosabb jellemzőit.

## Processzorok

Ismeretes, hogy a különféle számítógép-architektúrák az optimális felhasználási területet illetően is jelentősen eltérnek egymástól. A műveletek jellegétől (egyszerűbb, vagy összetettebb; többnyire logikai vagy inkább aritmetikai; adatról-adatra változó, vagy sok adaton változatlanul elvégezhető, stb.), az adatok jellegétől (pl. szóhossz), valamint az adatsomagok szerkezetétől függ, hogy milyen processzorral valósítható meg optimálisan egy adott feladat.

A képfeldolgozás leggyakrabban egyszerű műveleteket alkalmaz, melyeket nagyszámú, ugyanakkor kis szóhosszúságú adaton (képponton) kell elvégezni. Ilyen műveletek pl.: szűrések, geometriai- illetve intenzitáskorrekció, képek közötti aritmetikai, valamint logikai műveletek, képkódolás. Természetesen számításigényes feladat (pl. alakzatfelismerés) és nagy szóhosszúságon végzendő művelet (pl.: illesztőpontos transzformáció paramétereinek kiszámítása) is előfordul a számítógépes képfeldolgozásban, de nem ez a jellemző.

Az előbbieket alapján elmondható, hogy az egyszerűbb, de több processzort tartalmazó rendszerek általában jobban illeszkednek a képfeldolgozási feladatokhoz, mint a hagyományos, Neumann-elvű (soros) gépek.

A megvalósított párhuzamos rendszerek három csoportba sorolhatók:

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

A **futószalag szervezésű gépekben** a képi adatok az egymást követő részműveleteket végrehajtó, párhuzamosan működő processzorokból álló „futószalagon” haladnak végig. Az ilyen rendszerek optimális kihasználásához a megoldandó feladatot közel azonos idő alatt elvégezhető részfeladatokra bontják. A rendszer inicializálása és az első adat végigfutását követően minden újabb adat feldolgozása a leghosszabb részműveletnek megfelelő idő alatt történik meg. Tipikus futószalag architektúra alakítható ki pl. **transzputerek**, valamint **DSP-k** (*digital signal processor*) alkalmazásával.

A **párhuzamos homogén rendszerek** SIMD (*single instruction stream - multiple data stream*) struktúrát valósítanak meg. Általában egy vezérgép (*master*) dönti el, hogy milyen utasítást kell az összes alárendelt (*slave*) feldolgozó egységnek végrehajtania. A feldolgozó egységek általában saját tárral is rendelkeznek, a fejlettebb rendszerekben egymás között is tudnak adatokat cserélni. Az ilyen felépítésű **CLIP-IV**, **DAP** rendszerek rendszerint bitműveleteket végeznek, melyek használata a szokásostól igen eltérő gondolkodásmódot igényel. Említést kell tenni a **sejtautomatákról**, melyekről kevesen tudják, hogy megteremtésük szintén Neumann János nevéhez fűződik. Ezek nagy számú, egyszerű automatából állnak, melyek egy-egy rendszerben meghatározott szomszédsági viszonyban állnak egymással, s így lokális képi művele-

tek végzésére nagyon alkalmasak. Napjainkban már **VLSI** sejttelepeket is készítenek, valamint PC-be helyezhető sejtautomata bővítőkártya is létezik.

A **speciális inhomogén rendszerek** igen változatos, egymástól is jelentősen eltérő architektúrákat valósítanak meg. Közös jellemzőjük, hogy az egyes speciális funkciókat végrehajtó egységek egy-egy képfeldolgozási művelet közvetlen hardver implementációi, azaz a fentebb említett **célhardverek** körébe sorolhatók. 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 pl. a **PICAP** és a **ZMOB** rendszer. Napjaink korszerű mikroprocesszorai tartalmaznak MIMD típusú műveletvégző egysége(ke)t.

### Grafikus operációs rendszerek

A grafikus programok eleinte önálló grafikai felülettel rendelkeztek. Az egyes operációs rendszerek nem vagy minimális grafikai szabványtámogatást nyújtottak. Talán a Macintosh (Mac OS) és a Silicon Graphics (SGI IRIX) operációs rendszerei bizonyos értelemben kivételek, mivel a teljes operációs rendszer fejlesztése és a gépek hardveres kialakítása is alapvető képfeldolgozási feladatok hatékony támogatására koncentrált. A hardverhez és operációs rendszerhez nem igazított, önálló grafikus szoftverek állapota hamarosan megszűnt, mivel igényként lépett fel az egyes nyelvterületekre történő adaptálás és a gyorsan fejlődő, jó grafikus eszközök közvetlen szoftveres használata. Ezen két szemponthoz csatlakozott egy harmadik fontos elem is. A mikroprocesszorok száma és fajtája kezdett szélesedni, egyre több processzor jelent meg a piacon. Ez alapvetően megszabta nemcsak a grafikus programok, hanem a hardver és az operációs rendszerek fajtáit is.

Mindezek együttesen hozták létre azon igényt, hogy egységes hardveres és szoftveres felület alakuljon ki. Ez a grafikus keretrendszer. Egy ilyen keretrendszernek az alábbi szempontokat kell kielégíteni:

- egységes csatlakozási felület belső hardver fejlesztők számára,
- egységes csatlakozási felület hardver perifériákat fejlesztők számára,
- egységes csatlakozási felület szoftver fejlesztők számára,
- egységes felhasználói felület,
- hardver és szoftver közötti kapcsolódási pontok egységesítése.

Napjaink multitasking operációs rendszerei ezeket a szempontokat rendkívül hatékonyan valósítják meg. A fenti szempontok megvalósításának mélysége viszont jelentősen eltér. A Windows dinamikusan csatlakozó könyvtárrendszerét a Windows 95/98, Windows NT operációs rendszereknél kiegészítették számos olyan felhasználói funkcióval, amelyek nem tisztán grafikus funkciók (API - egységes programozói felü-

let, GUI - grafikus felhasználói felület, NTFS file kezelés, Plug-and-Play, bővített Internet használat, egységesített médiakezelés, stb.). A Macintosh OS operációs rendszere ennél is többet, közvetlen hardveres támogatást nyújt alapvető képfeldolgozási feladatok elvégzésére. Talán egyféle válasz, hogy az Intel MMX, PII, PIII processzorokkal szerelt PC-k megfelelően felkészített szoftver esetén közvetlen processzoros grafikus (multimédia alapú) gyorsítást érhetnek el. Kétségtelenül a legjobb megoldásokat a Silicon Graphics 64 bites IRIX (6.3, 6.4) operációs rendszere nyújtja.

Az alábbi táblázat olyan jellemzőket tartalmaz, amelyek elsősorban képfeldolgozási, grafikai és animációs operációs rendszer szintű támogatások, összehasonlítva több operációs rendszer által nyújtott lehetőséggel.

<b>Funkció</b>	<b>SGI IRIX 6.4</b>	<b>MAC System 9</b>	<b>WINDOWS98</b>
Operációs rendszer felépítése	64 bit, multi-user, multitasking	32 bit, monouser, multitasking	32 bit, monouser, multitasking
Operációs rendszer védelme	réteges, stabil	réteges, stabil	részben réteges, részben stabil
Többprocesszoros környezet támogatása	Igen	Nem	Nem
GUI – grafikus felhasználói felület	Igen	Igen	Igen
Plug-and-Play bővítési lehetőség támogatása	Igen	Igen	Igen
Grafikus elemek alapján felépülő fontkészlet (TTF)	Igen	Igen	Igen
Multimédia támogatás	Igen	Igen	Igen
Egységesített médiakezelés	Igen	Igen	Igen
Háromdimenziós megjelenítés támogatása	Igen	Igen	Igen
Közvetlen Internet támogatás	Igen	Igen	Igen
OpenGL grafikus környezet	Igen, teljesen hardveres támogatású	Igen, részben hardveres támogatású	Igen, főleg szoftveres támogatású
Egységesített memóriakezelés - UMA	Igen	Nem	Nem
POSIX többprocesszoros fejlesztői támogatás	Igen	Nem	Nem
Raszteres fejlesztői támogatás	Igen, ImageVision Library	Igen	Nem

Kétségtelen tény, hogy napjainkban kialakult egy szabványokon nyugvó grafikus operációs rendszer etalon. A megvalósítás és az elért teljesítmény azonban rendkívüli módon függ az alkalmazott hardverelemektől. Gondoljunk csak arra, hogy számos cég alaplapon készít és támogat a szabványhoz illeszkedő vizuális feldolgozásokat, míg mások kénytelenek szoftveresen megvalósítani bizonyos feladatokat. Természetesen a két rendszer között jelentős ár és teljesítménybeli különbség mutatkozik.

### Szoftverek

A képfeldolgozás speciális feladatainak megoldásához, valamint speciális eszközeinek vezérléséhez speciális **programokra** van szükség. Ezek többnyire önálló rendszerek, amelyek egy-egy nagyobb feladatosztály megoldására alkalmasak. Így léteznek, pl. DTP (*desktop publishing*) rendszerek, távérzékeléssel kapott adatok feldolgozását végző rendszerek, orvosi képelemző rendszerek, ipari anyagvizsgálati képelemző rendszerek, melyek kezelése a felhasználótól szakmai elmélyülést, gyakorlást követel. Léteznek egyszerűbb rendszerek (pl.: rajzolóprogramok képfeldolgozási jellegű funkciókkal, sejt számláló program, stb.), melyek kezelése kevesebb felkészültséget igényel. Sőt az ipari folyamatirányításban használt képfeldolgozó rendszerek kezelőjének többnyire semmilyen képfeldolgozási ismeretre sincs szüksége.

Néhány jellemző, amely az képfeldolgozó programokat megkülönbözteti a szokásos programoktól:

- Korábban a hatékonyság növelése (azaz a végrehajtási idők csökkentése) érdekében a programok alacsony szintű (*assembly*) nyelven készültek. A számítógépek kapacitásának növekedésével és a korszerű fejlesztő eszközök megjelenésével egyre inkább lehetővé válik magasabb szintű nyelvek használata, és csak ritkán kell az *assembly* nyelvet alkalmazni. A C nyelv használata különösen elterjedt a képfeldolgozásban.
- Kifejezetten képek tárolására szabványosított adatstruktúrákkal (pl.: **DIB**: *device independent bitmap*) és file-formátumokkal (pl.: **.TIF**, **.PCX**, **.TGA**, **.IMG**) dolgoznak (lásd 16. fejezet).
- Speciális (képfeldolgozási célú) eszközöket vezérelnek.

### Célhardverek

A képfeldolgozás speciális feladatainak hatékony (gyors) megoldásához speciális eszközök jelentik a legjobb megoldást - amennyiben a rugalmasság, a működés algoritmusának, paramétereinek változtatása nem fontos.

A legkülönfélébb feladatok megoldására készültek célberendezések, melyek a hétköznapi eszközöktől (printerek) kezdve az űrkutatási vagy katonai célú eszközökig terjednek. Néhány kiragadott feladat, melyek megoldására létezik célhardver:

- Színrebonatás
- Ditheringes képnymtatás
- Konvolúciós szűrés
- Képek átlagolása
- RGB <-> HSI adatkonverzió
- CCITT képtömörítés
- JPEG képtömörítés
- MPEG képtömörítés

## 2.5. Képi másolat készítése

Képi másolat készítésének (azaz a kép rögzítésének) egyik legegyszerűbb módja a képernyő **lefényképezése**. A polaroid technika megjelenése különösen meggyorsította ezt az eljárást. A fellépő torzító hatások (hordótorzítás, fényvisszaverődés, stb.) kiküszöbölésére pl. különleges, sík felületű és monokromatikus képernyőt használnak, s az egyes színösszetevőket színszűrők közbeiktatásával fényképezik le.

A képbeviteli eszközöknél ismertetett **dob-szkennerek** egy része úgy készült, hogy megfelelő kiegészítéssel alkalmas legyen a képek nagy pontosságú felrajzolására is. Ezekben a **filmíró** („**levilágító**”) berendezésekben egy jól fókuszált, 10-100 $\mu$ m átmérrőjű fénysugár fényerősségét modulálják a képpontok fényességének megfelelően. A fénysugár a képsoroknak megfelelően pásztázza végig a filmet, s így „**rárajzolja**” a képet, mely előhívás után válik láthatóvá.

A korai képfeldolgozó rendszerekkel egy időben jelentek meg a különböző színes plotterek, melyek a képet pontonként rajzolták ki. Az egyik módszer - a **görgős rajzolófejes eljárás** - esetében a papírt forgó henger palástjára erősítették, s a görgőkkel a szubtraktív színkeverés szabályainak megfelelően cián, magenta és sárga színű festék csíkokat vittek fel a papírra. A dob egy körülfordulása alatt egy képsor felrajzolása történt meg, s ezalatt az előtolás érdekében a görgőket tartó szán a dob tengelyével párhuzamosan éppen egy sor magasságával mozdult el.

A tintasugaras (*ink-jet*) eljárásnál kis fúvókákból a megfelelő időpillanatokban kilőtt kis festékcseppecskék alkotják a képpontokat, míg az ún. **thermal-transfer** eljárás hirtelen felmelegített festékszalag felületéről „**kirobbanó**” (szublimáló) festékfoltokat használ. Az előbbi eljárás hatékonyabban bánik a festékkal, míg az utóbbinál a papír és a festék-fólia együtt fogy, függetlenül attól, hogy a képen mennyi a színes részlet.

Szintén elterjedt az elektromos fényképezés különböző változatainak alkalmazása. Az általánosan elterjedt **xerox** eljárás lényege (röviden) a következő:

- a fényelektromos tulajdonságokkal rendelkező anyaggal (pl.: szelén, egyéb félvezetők) bevont henger felületét nagyfeszültségű elektródasorral elektromosan feltöltik.
- a töltéssel rendelkező hengeren optikai úton éles képet hoznak létre, mely a felületi töltés egy részének elvesztéséhez vezet. Így negatív ún. **töltéskép** alakul ki (ahol a hengert sok fény éri, ott kevesebb töltés marad). (Megjegyzendő, hogy a képalkotás vékony sávban történik, s a henger forgása valósítja meg az oszlopírányú előtolást.)
- töltéshomogenizálás után a henger porfestéket tartalmazó tartályhoz fordul, melytől feltöltött részein festékszemcséket vonz a felületére, s a töltésképnek megfelelő kép alakul ki.
- a henger a palástjával azonos sebességgel mozgó papír felületéhez nyomódik, s a festék (az elektrosztatikus vonzás következtében) a papírra kerül, az eredeti képnek megfelelően.
- a továbbforgó henger felületét a megmaradt festéktől elektrosztatikusan és mechanikusan megtisztítják, míg a papíron levő festéket hőhatással rögzítik.

A felvázolt eljárás a hagyományos másológépek alapját is képezi, de így működnek a korszerű **lézerprinterek** is.

A képi másolatkészítő eszközökkel kapcsolatban még azt is meg kell említeni, hogy - eltekintve a fototechnikai készülékektől - csak napjainkban kezdenek elterjedni azok az eszközök, melyek a tónusos kép közvetlen visszaadására is alkalmasak. (A fentebb felsorolt készülékek egy-egy pontban vagy tesznek festékpontot, vagy nem, de a festékpont nagysága vagy színessége nem befolyásolható).

A tónusos képek fényesség- illetve színhű visszaadása a felsorolt eszközökön speciális algoritmus - ún. **dithering** - alkalmazását teszi szükségessé. Ennek lényege: ha már egy-egy pontban nem adható vissza a helyes tónus, legalább a közeli pontok alkotta lokális minták együttese közelítse azt. Az eszközök egy része célhardverként beépítve tartalmazza a szükséges adatkonverziós egységet, más eszközök esetében a működtető szoftvernek kell a konverzióról gondoskodnia.

Megemlítendő még az is, hogy a professzionális nyomdatechnikában a helyes tónus és szín visszaadására speciális technikákat alkalmaznak, melyek pl. a festék papírra kerülésének törvényszerűségeit is messzemenően figyelembe veszik.

## 2.6. Képkódolás, képtömörítés

A digitális kép általában igen nagy, ugyanakkor rendkívül **redundáns** adathalmazként fogható fel.

Ez utóbbi annyit jelent, hogy a képek által tartalmazott **lényegi információt** keve-

sebb adattal lehetne jellemezni, mint ahány adatból maga a kép áll. Lényegében ez adja meg a lehetőséget a **kép tömörítésére**, ami tehát nem a kép méretének csökkentését jelenti, hanem ugyanannak a képnek a tárolásához szükséges adatok számának csökkentését.

A felhasználási területek többsége a képi adatok kismértékű változását is megengedi (ha a felhasználás célkitűzéseit ez a változás nem veszélyezteti), s ez a redundancia kihasználásán túlmenően további tömörítésre nyújt lehetőséget. Így pl. abban az esetben, ha a képek szerepe csupán a **látvány** létrehozása (egzakt feldolgozásra nincs szükség), akkor az emberi látás sajátosságainak figyelembevételén alapuló veszteséges, de jelentős látványromlással nem járó képtömörítési formátumok (eljárások) is alkalmazhatók (pl.: **JPEG, MrSID, Fractal**).

A képtömörítés tipikusan az a művelet, amire érdemes **célhardvert** készíteni. (Az algoritmusok a szükséges szabványosítás után már nem változnak; a célhardver egyéb készülékekkel egybeépíthető, így azok eladhatóságát javítja; a köznapi felhasználási igények miatt a célhardver nagy szériában gyártható, azaz kellően olcsó lehet, stb.).

Teljességre való törekvés nélkül megemlítünk néhány alaptípust:

- CCITT G3, illetve G4 fakszimile kódolást/dekódolást megvalósító eszközök. A kódolt kép kétszintes (*bináris*) lehet, bár újabban megjelent olyan kiterjesztés is, mely három kétszintes kép együttes kezelésével színes képekre is alkalmazható. A kódoló/dekódoló célhardver leggyakrabban telefax-készülékekkel van egybeépítve.
- Az amerikai INTEL cég **i750** (*video processor*) VLSI áramkörével 160:1 arányú mozgóképek tömörítés is elérhető, s így pl. szabványos **optikai lemezen** (CD-ROM-on) 72 percnyi mozgó videofelvétel rögzítésére van lehetőség.
- JPEG és MPEG kódoló/dekódoló eszközök. Léteznek valósidejű (*real-time*) változatai is (pl. Pinnacle Studio MP10), melyek videojelet szolgáltató képforrással együttműködve folyamatosan szolgáltatják az 15:1-50:1 arányban tömörített fekete-fehér, illetve színes képeket. Ezek többnyire néhány **VLSI** integrált áramkörön alapulnak, pl.: INMOS **A121**, C-Cube Microsystem **CL-550**, Texas Instruments **TMS-320C25**.

Megjegyzendő, hogy képek kódolása alatt nem csupán tömörítésüket, hanem pl. illetéktelen használat elleni **titkosításukat** is értjük; természetesen erre a kényes felhasználási területre is készültek célberendezések.

Ebben a fejezetben inkább csak utaltunk a képtömörítésre és képkódolásra, mint olyan területre, ahol a célhardverek igen hatékonyan alkalmazhatók. A téma elméleti és gyakorlati kérdéseivel a 8. fejezetben részletesen foglalkozunk.



## 2.7. Képátvitel

A **képátvitel** a számítógépes képfeldolgozásban egyaránt jelentheti egy rendszerben működő eszközök közti **lokális adatszerét**, valamint nagy távolságra történő **információközlést**.

A **lokális adatszerére** az egyébként is kifejlesztett hálózati csatolások általában minden további nélkül használhatók. Ezek átviteli sebessége a Megabyte nagyságrendű képek 1 másodpercen belüli átviteléhez elegendő; a szokásos felhasználási területeken a háttértároló kezelése, illetve a kísérő adatfeldolgozás ennél lényegesen tovább tart.

Alapvetően írott, gépelt dokumentumok telefonvonalon való átvitelére fejlesztették ki a **telefaxot**, mely kétszintes képek kb. 75 dot/inch felbontású átvitelére is alkalmas.

A nagy távolságú **információközlésre** (pl. műholdas képátvitel) is léteznek olyan általános megoldások, melyek természetesen a képek átvitelére is alkalmazhatók. Ugyanakkor a képek tipikusan nagy adatmennyisége általában adattömörítés alkalmazását teszi szükségessé. Pl.: a szokásos telefonvonalakon kb. 9600 baud (bit/sec.) az átviteli sebesség, mely az átvitel frekvenciakorlátozása és az óhatatlanul fellépő zavarok mellett még megengedhető. Így egy 512\*512 képpontból álló tónusos (8 bit/képpont felbontású) kép átvitele közel 5 percet venne igénybe. A gyakorlatban a **Képtömörítés** c. fejezetben leírt módszerek (általában: tömörítő célhardverek) alkalmazásával, sokkal jobb eredményeket lehet elérni.

Napjaink szabvány (veszteséges) képtömörítő eljárásait közvetlenül felhasználva kialakult a videokonferencia rendszer, amely ISDN kapcsolatot feltételezve mozgókép, hang és adat közel valós idejű, kétirányú átvitelét jelenti. Küszöbön áll a képtelefon teljesen valós idejű megjelenése, azaz a digitális képátvitel bevonulása a mindennapi életbe.

### Ellenőrző kérdések

1. Mi a képbeviteli eszközök feladata?
2. Mi az elsődleges és mi a másodlagos fényforrás?
3. Hasonlítsa össze az elektroncsöves és a CCD képérzékelő eszközöket!
4. Milyen technikai paraméterekkel lehet a digitalizáló kártyákat jellemezni?
5. Mit tud a CMOS képérzékelőről?
6. Melyek az AGP előnyei?
7. Mi az UMA?
8. Hol használják a lapolvasókat?
9. Hol célszerű ipari CCD kamerák alkalmazása?

10. Mely eszközben található ún. *dithering* eljárásra szolgáló célhardver és mi a feladata?
11. Ismertesse napjaink grafikus operációs rendszereinek alapvető jellemzőit!
12. Melyek a képfeldolgozó programok főbb általános jellemzői?
13. Melyek a párhuzamos képfeldolgozás főbb típusai?
14. Mi a különbség és mi az azonosság a dobszkenner és a filmíró között?

## 3. DIGITÁLIS KÉPALKOTÁS

### 3.1 Bevezetés

A digitális képfeldolgozás azon alapul, hogy a valós világ objektumainak képe számítógéppel kezelhető adatokká alakítható. Ezt az átalakítást nevezzük **digitális képalkotásnak**. A digitális képalkotás egzakt matematikai módszerekkel is leírható, de itt megelégszünk egy szemléletes, egyszerűbb modellel. Az érdeklődők az Ajánlott irodalom első két tételeként megadott könyvekben megismerkedhetnek a részletesebb matematikai háttérrel.

A digitális képalkotás célja, hogy a háromdimenziós térben lévő objektumokról a számítógép által értelmezhető és feldolgozható adatokat kapjunk. A továbbiakban látni fogjuk, hogy a digitalizálás folyamán információvesztés lép fel, de ezért kárpótol a számítógép által történő feldolgozás által kínált számos olyan lehetőség, amely hagyományos optikai eljárásokkal nem valósítható meg.

A digitalizálás speciális eszközökkel történik, és az eszköztől függően az egyes fázisok megvalósítása különböző módon történhet, esetleg több fázis egyetlen lépésben is megvalósulhat. Ha az egyes lépéseket logikailag szét kívánjuk választani akkor lényegében három lépést különböztethetünk meg, nevezetesen a **leképezést**, a **mintavételezést** és a **kvantálást**. Vizsgáljuk meg kicsit részletesebben a digitális képalkotás egyes lépéseit.

### 3.2. Leképezés

A leképezés szót most nem matematikai értelemben kívánjuk használni, hanem annak a bonyolult folyamatnak a leírására, amelynek során a háromdimenziós térben lévő objektumokról érkező elektromágneses jelekből egy kétváltozós  $f(x,y)$  képfüggvény jön létre. A képfüggvény matematikai értelemben folytonos, nem negatív és korlátos, és a sík azon pontjain van értelmezve ahová a leképezés történt. A TV-technikában szokásos megjelenítési mód miatt a képsíkon az  $x$ -tengelyt balról-jobbra, az  $y$ -tengelyt pedig általában felülről-lefelé értelmezik az origó pedig a feldolgozandó kép bal felső sarkába esik. A kép általában téglalap alakú. Egyszínű (monokróm) kép esetén az  $f(x,y)$  függvény értéke minden  $(x,y)$  pontban arányos a kép fényességével az adott helyen. Ezt az értéket szürkeségi szintnek (gray level), fényességi értéknek vagy világosságkód értéknek szokták nevezni. Ezt az értéket rengeteg körülmény befolyásolja. A háromdimenziós térben lévő tárgyak minden  $(x,y,z)$  pontja a fényforrás fényét (vagy annak egy részét) visszaveri. Egyes esetekben a testnek jelentős saját elektromágneses

sugárzása is lehet. A visszavert és a saját sugárzás természetesen függ az időponttól ( $t$ ). Nem hagyhatjuk figyelmen kívül a megvilágítás szögét és a megfigyelés irányát sem, amit egy vektorral írhatunk le ( $\underline{n}$ ). Közismert, hogy a fény különböző hullámhosszú összetevőkből áll, a fény hullámhossza ( $\lambda$ ) szintén befolyásolja a kialakuló képet. Mivel a leképezés a gyakorlatban nem egy adott frekvencián, hanem egy bizonyos sávban történik, a képfüggvény a leképezés **spektrális tulajdonságaitól** is függ. Ez pl. adott tartományban többé-kevésbé azonos súlyfüggvénnyel (*érzékenységgel*) történő integrálásnak felel meg. Ha a leképezés egyidejűleg több - pl. más-más sáv tartományból származó - kimenőjelet szolgáltat, akkor a képfüggvény vektor lesz. Az egyszerűség kedvéért egyelőre csak monokróm képfüggvénnyel foglalkozunk.

Ha formalizálni akarjuk az elmondottakat, akkor felírhatjuk az  $e(x,y,z,t,\lambda,\underline{n})$  folytonos függvényt, ami a háromdimenziós térben kialakuló fényességfüggvényt reprezentálja. A képalkotó rendszer ezt „képezi le”, ebből állítja elő a kétváltozós  $f(x,y)$  képfüggvényt. Ismételten hangsúlyozni szeretnénk, hogy itt most nem egy egzakt matematikai leképezésről van szó, hanem arról, hogy az adott berendezés - például a TV kamera, vagy a fényképezőgép - a rendkívül bonyolult és teljes pontossággal matematikailag le sem írható jelenségből létrehoz egy már egyszerűbb, és jobban kezelhető jelenséget, ami az  $f(x,y)$  képfüggvénnyel írható le. TV kamera esetén a töltéskép, a fényképezőgép esetén a film illetve az előhívott fotó tekinthető a képfüggvény fizikai megtestesítőjének. Ha nagyon pontosak akarunk lenni, akkor a töltéskép és a fotó sem tekinthető folytonos függvény realizációjának, mivel a töltéskép esetén az elemi részecskék diszkrét fizikai jelenségként is leírhatók, a fotó esetén nagyon kicsi, de jól meghatározható méretű szemcsékből áll össze a kép. Ennek ellenére az említett példák elfogadhatóak, hiszen éppen a töltésképből valamint a fotóból történik a mintavételezés és kvantálás a digitalizálás további lépései során.

### 3.3. Mintavételezés

A digitális képalkotás következő lépése a **mintavételezés** (*sampling*). A mintavételezés az  $(x,y)$  képsík egy-egy kis területéhez (*képeleméhez*) a fényességtől függő számot rendel. Ebben a nagyvonalú meghatározásban két fontos dologra kell odafigyelni, ami meghatározza a létrejövő digitális kép minőségét. Az egyik lényeges pont az „egy-egy kis terület” mérete, a másik a hozzárendelés módja.

Gondolatban végezzük el a mintavételezést úgy, hogy egy fényképre milliméterpapírt borítunk, és a papír rácsozata határozza meg a „kis területeket”. Satírozunk be minden egyes kis területet a terület közepén lévő szürkének megfelelő árnyalatra. Nyilvánvalóan nem mindegy, hogy egy centiméteres vagy egy milliméteres oldalhosszúságú négyzetet választottunk egységnyi területnek; finomabb rácsozat esetén

jobb lesz az eredménykép, ha durvább az osztás akkor a kép minősége is gyengébb lesz. Ha a szürkeségi értékeket számszerűen meg tudjuk határozni, és ezeket a számokat egy mátrixba beírjuk, lényegében elvégeztük a mintavételezést. Ezek a számok nem negatív valós számok lesznek. Jelöljük  $h(k,l)$ -el azt a függvényt, amely minden négyzet középpontjához, a  $(k,l)$  pontokhoz hozzárendeli a mintavételezéssel kapott értéket. Ez még nem a digitális képfüggvény, mert ugyan csak a diszkrét pontokban van értelmezve, de bármely nem negatív valós számértéket felvehet.

Könnyű belátni, hogy jobb minőségű képet kapunk, ha nem a négyzet közepén lévő értéket választjuk a satírozásnál, hanem átlagoljuk a négyzeten belüli értékeket, és az átlagértéket használjuk. Előfordulhat ugyanis, hogy pont a négyzet közepén szemcsehiba vagy a felület sérülése miatt kiugróan eltérő érték van, a képi tartalomtól függetlenül. Ha ezt az értéket választjuk és alkalmazzuk, akkor a hibát kiterjesztjük az egész négyzetre. Ha viszont az átlagértéket választjuk, akkor az adott területen lévő értékek csökkentik, korrigálják a hibás adatot.

Az előbbi szemléletes példán továbblépve elmondhatjuk, hogy a mintavételezés matematikailag pontosan leírható, és mint matematikai művelet lényegében **integrálás**. A mintavételezés során kapott képpontérték egy kis terület fényességértékeit integrálja (összegzi) úgy, hogy a súlyfüggvény a mintavételezés helyétől távolodva rohamosan csökken, azaz a mintavételezés helyétől távol levő értékeket egyre kevésbé veszi figyelembe.

**Ideális mintavételezés** esetén minden képpontérték a **teljes**  $f(x,y)$  képfüggvény alapján alakul ki. A gyakorlatban azonban az a tartomány, ahol a mintavételezés súlyfüggvénye nem zérus, csak a szomszédos mintavételezési helyek súlyfüggvényeinek értékes tartományáig terjed, s legfeljebb kismértékű átfedésben van azokkal.

Megemlítjük, hogy a képfüggvényre és a mintavételezés módjára vonatkozó bizonyos feltételek teljesülése esetén matematikailag bebizonyítható, hogy a mintavételezett képből még visszaállítható az  $f(x,y)$  képfüggvény. Sajnos a gyakorlatban ezek a feltételek többnyire nem biztosíthatóak, például a zajos képet - márpedig jószerivel minden kép többé kevésbé zajos (lásd Képjavítás fejezet) - nem lehet egyértelműen visszaállítani.

Megemlítjük még, hogy színes és multispektrális képek esetén a mintavételezés színösszetevőnként, illetve az egyes spektrumokra külön-külön történik, így a kép egy-egy kis területéhez  $n$  dimenziós vektorokat rendel, ahol a vektor dimenziószáma a spektrális összetevők számával egyezik meg.

## Gyakorlatok a mintavételezéshez

A felbontás paraméter ebben az esetben azt jelenti, hogy mekkora területről veszünk egy-egy mintát. Bár itt már a kiindulási kép is digitális, de a felbontás paraméter változtatása hasonló eredményre vezet, mintha az analóg képfüggvényből indulnánk ki, és a mintavételezés sűrűségét változtatnánk.

### 1. Gyakorlat:

Geometriai felbontás\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép1\_  
Felbontás\_2\*2\_Kilép\_Végrehajt

*Változtassuk a felbontást, és figyeljük meg a kép minőségének változását!  
Ismételjük meg a gyakorlatot a Kép2-vel!*

### 2. Gyakorlat:

Geometriai felbontás\_Indít\_Szín\_RGB\_Képforrás\_File\_Kép1\_  
Felbontás

*Külön-külön állítsuk be az R, G és B összetevő felbontását! Válasszunk eltérő felbontás értékeket az egyes összetevőknek! Vizsgáljuk meg melyik összetevő felbontása befolyásolja leginkább a kép minőségét!*

### 3. Gyakorlat:

Geometriai felbontás\_Indít\_Szín\_YUV\_Képforrás\_File\_Kép1\_  
Felbontás

*Külön-külön állítsuk be az Y, U és V összetevő felbontását! Válasszunk eltérő felbontás értékeket az egyes összetevőknek! Vizsgáljuk meg melyik összetevő felbontása befolyásolja leginkább a kép minőségét!*

## 3.4. Kvantálás

A digitalizálás utolsó lépéseként a (bizonyos tartományban) tetszőleges értékű  $h(k,l)$  képpontértékekhez a megengedett (diszkrét) képpontértékek valamelyikét kell rendelni. Ezt a lépést nevezik **kvantálásnak** (*quantization*), eredménye a **digitális kép**

amit jelöljünk  $q(k,l)$ -el. A  $q(k,l)$  függvény a  $(k,l)$  pontokban értelmezett, ahol

$$0 \leq k < K, \quad \text{és} \quad 0 \leq l < L$$

egész számok,  $K$  a képsorok számát és  $L$  a képsorok hosszát adja meg.

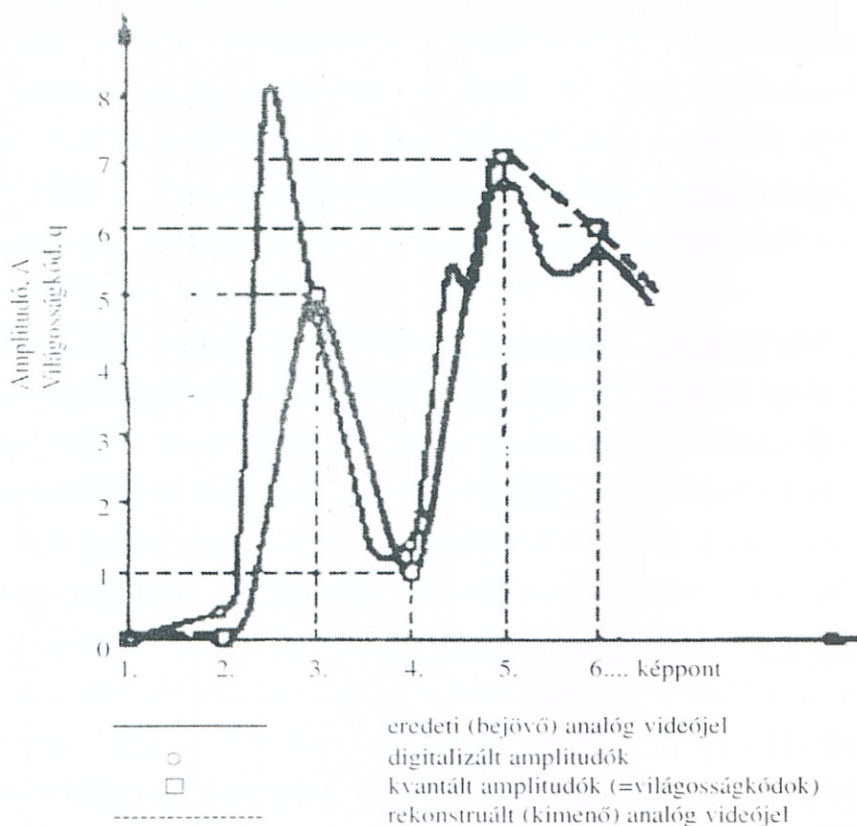
A képpontértékek (más néven: *világosság kódok*) ábrázolása meghatározott számú biten történik, ami a lehetséges  $q$  értékeket a következőképpen szabja meg:

$$0 \leq q \leq 2^b - 1; \text{ ahol } b \text{ az A/D átalakító bitszáma.}$$

(A szokásos esetben  $b=8$ , azaz a képpontértékek a  $\{0, \dots, 255\}$  értékeket vehetik fel.)

Színes kép esetében a kvantálás a mintavételezéshez hasonlóan, összetevőnként, egymástól függetlenül valósul meg.

A mintavételezés és kvantálás eredményét egyetlen képsorra, azaz egydimenziós esetre a következő ábrán szemléltetjük.



5. ábra

egyenletes (vagyis a sűrűségfüggvény állandó), akkor az **egyenlő osztásközű kvantálás** az optimális. Ha az eloszlás nem egyenletes, akkor ennél jobb eredményt kapunk, ha a gyakoribb képpontértékeknél a kvantálási szinteket egymáshoz közelebb választjuk meg. Az ilyen eloszlásfüggő (*tapered*) kvantálót a gyakorlatban úgy valósítják meg, hogy az **A/D átalakító** előtt a képfüggvényt nemlineáris transzformációnak ve-

A kvantálás „jóságát” (a megfeleltetés „hűségét”) az elkövetett négyzetes hibával jellemezzük. Ez egyszerűen a bemenő  $h$  képpontértékek és a hozzárendelt  $q$  kvantált értékek közti eltérések négyzetének összege (integrálja) a teljes értelmezési tartományban.

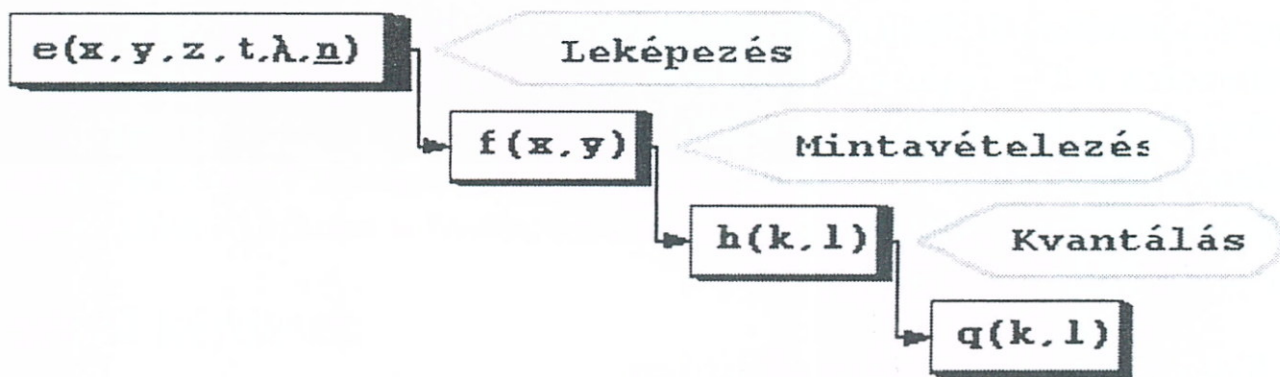
A kvantálást akkor tekintjük **optimálisnak**, ha ez a négyzetes hiba minimális (természetesen rögzített ábrázolási pontosság esetén).

Könnyű belátni, hogy amennyiben a bemenő képpontértékek eloszlása

tik alá, aminek hatására az eredmény sűrűségfüggvénye állandó lesz. Ezek után egyenlő közül kvantálást végeznek, majd a kapott kódokból inverz transzformációval állítják elő a tényleges világosság kódokat.

A digitalizálás során tehát az ún. fényességfüggvényből eljutunk a **digitális kép**-hez, melynek adatai a **képpontok** (ún. pixelek: picture's elements). Sorirányban  $L$ , oszlopirányban pedig  $K$  db képpont alkotja a képet.

Az alábbi ábra a digitalizálás egyes lépéseit mutatja:



6. ábra

A digitális kép tehát számok, (illetve multispektrális képek esetén vektorok) - **képpontok** - halmaza, mely számítógépes feldolgozásra közvetlenül alkalmas. A képszerű használatához újra vissza kell állítani az analóg képet, ami megfelelő eszközök és algoritmusok használatát teszi szükségessé.

## Gyakorlatok a kvantáláshoz

A felbontás paraméter ebben az esetben azt jelenti, hogy hány szintet engedünk meg a kép megjelenítésekor. Bár itt már a kiindulási kép is digitális, de a felbontás paraméter csökkentése hasonló eredményre vezet, mintha a mintavételezett, de még tetőleges fényességi értékeket tartalmazó képfüggvényből indulnánk ki, és a kvantálási szintek számát csökkentenénk.



#### 4. Gyakorlat:

Gradációs felbontás\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép1\_  
Felbontás\_128\_Kilép\_Végrehajt

*Fokozatosan csökkentsük a felbontást, azaz a megengedett kvantálási szintek számát! Figyeljük meg mikor lesz szemmel is érzékelhető a minőségromlás! Ismételjük meg a gyakorlatot Kép2-vel!*

#### 5. Gyakorlat:

Gradációs felbontás\_Indít\_Szín\_RGB\_Képforrás\_File\_Kép1\_  
Felbontás

*Fokozatosan csökkentsük a felbontást, azaz a megengedett kvantálási szintek számát! Figyeljük meg mikor lesz szemmel is érzékelhető a minőségromlás! Ismételjük meg a gyakorlatot Kép9-cel!*

### 3.5. Folytonos kép helyreállítása

A **helyreállítás** (*reconstruction*) során a digitális képből **interpolációval** állítják elő az analóg képjelet. Mint azt már említettük, az eredeti képfüggvény több okból sem állítható vissza **tökéletes pontossággal**.

Az egyik alapvető okot a kvantálás jelenti, ami eleve lehetetlenné teszi az eredeti mintavételezett képpontérték helyreállítását.

Amint azt korábban említettük bizonyos feltételek teljesülése esetén az eredeti képfüggvényt mintavételezés után még tökéletes pontossággal helyre lehetne állítani, azonban a mintavételezéssel kapcsolatos feltétel a gyakorlatban soha nem szokott teljesülni.

A gyakorlati interpolátor (**D/A átalakító**, kiegészítő áramkörökkel) az eredeti analóg képfüggvényt csak korlátozott pontossággal képes helyreállítani. Ez annak köszönhető, hogy az alkalmazhatóság szempontjait is figyelembe véve a matematikailag ideális interpolátor helyett csak gyengén közelítő interpolátorokat alkalmaznak.

Legelterjedtebb megoldásként **lineáris interpolációt** alkalmaznak, illetve ennek kétdimenziós megfelelőjét, az ún. **bilineáris interpolációt**.

### 3.6. A digitális képalkotás és a kép minősége

Az előzőekből látható, hogy a digitális kép minősége, így a további feldolgozás eredményessége szempontjából alapvető fontosságú a digitalizálás. Értelemszerűen a mintavételezés sűrűségének és a kvantálási szinteknek a növelése javítja a digitális kép minőségét, de egyidejűleg a kép méretét is növeli.

A képméret növekedése természetesen hosszabb feldolgozási időt fog eredményezni.

Az egyes alkalmazásoknál tehát a feldolgozási igények alapján kell az optimális digitalizálási paramétereket megválasztani. A nyomdai képfeldolgozásnál például elfogadhatatlan minőségű képet eredményez egy TV kamera és digitalizáló kártya kombináció 512\*512\*8 bites képe. Ez a felbontás ugyanis még a 600 dpi felbontást sem éri el, ha egy inches méretben akarjuk a képet kinyomtatni. Ennél a méretnél ez még elfogadható, de abban a pillanatban, ha nagyítani kezdjük a képet, akkor már a minőség erősen kifogásolható lesz. Ugyanakkor ipari alkalmazásoknál kezelhetetlen és felesleges lenne 4096\*4096\*24 bites színes képek alkalmazása, a hosszú feldolgozási idő miatt.

#### Ellenőrző kérdések

1. Mi a célja a kép digitalizálásának?
2. Milyen tényezők befolyásolják a bemeneti fényességfüggvényt?
3. Milyen lépésekből tevődik össze a digitalizálás?
4. Mit jelent a leképezés a digitalizálás folyamatában?
5. Mit nevezünk mintavételezésnek?
6. Mit jelent az ideális mintavételezés?
7. Mit jelent a kvantálás?
8. Mikor tekintjük a kvantálást optimálisnak?
9. Mekkora helyet foglal el egy 1024\*512 pixelből álló 24 bites digitális kép?
10. Hogyan függ a kép minősége a digitalizálás paramétereitől?

## 4. KÉPJAVÍTÁS

### 4.1 Bevezetés

A képvétel körülményei, a digitalizáló berendezés paraméterei jelentősen befolyásolják a digitális kép minőségét. A valós világ leképezése, esetenként a képnek a feldolgozás helyére történő továbbítása során különböző zavaró hatások érik a képet. Ezeknek a hatásoknak a kiküszöbölésére vagy legalábbis csökkentésére szolgáló módszereket és eljárásokat **képjavitásnak** nevezzük.

### 4.2. A képjavitással kapcsolatos alapfogalmak

A **képjavitási** (*improvement*) módszerek - a megvalósítandó cél szerint - két csoportba sorolhatók:

- **képhelyreállítás** (*restoration*),
- **képfokozás** (*enhancement*).

A **képhelyreállítás** során - mint az elnevezés is jelzi - arra törekszünk, hogy a digitalizált és különböző okok miatt torzult képből előállítsuk azt az ideálisnak nevezhető képet, amelyet a különböző zavaró, torzító hatások nélkül kaptunk volna.

A **képfokozási** (azaz: képminőség-fokozási, image enhancement) eljárások célja, hogy a képet a további kiértékelés vagy feldolgozás szempontjából előnyösebb formába alakítsuk. Ebben az esetben nem feltétlenül az eredetihez leginkább közel álló képet akarjuk megkapni, esetenként éppen ellenkezőleg: a további feldolgozás szempontjából fontos jellegzetességek hangsúlyozásával, fokozásával kapjuk a számunkra legkedvezőbb eredményt.

Megemlítjük, hogy a magyar szóhasználatban sokszor a képjavitási eljárások közé csak az angol terminológiában „image enhancement” névvel jelölt képfokozási eljárásokat szokták sorolni.

A **képjavitási eljárások** két csoportba sorolhatók attól függően, hogy a frekvenciatartományban (pl. a kép **Fourier-transzformáltjával**), illetve az **eredeti képsíkon** (a **világosságkódokkal**) dolgozunk. Mit is jelent ez egy kicsit részletesebben ?

Az előző fejezetben láttuk, hogy az  $f(x,y)$  folytonos képfüggvényből a mintavételezés és kvantálás révén hogyan kapjuk meg a  $q(k,l)$  digitális képet. Kézenfekvő, és a digitális képfeldolgozásban gyakori megoldás, hogy a folytonos függvényekre alkalmazható matematikai eljárásokat, modelleket használunk egy-egy képfeldolgozási eljárás elméleti megalapozásához, majd az így kapott összefüggések digitális (matematikai értelemben diszkrét) közelítését implementáljuk. A frekvenciatartományban vég-

zett képjavítás azon alapul, hogy a függvények bizonyos feltételek esetén egyértelműen és invertálható módon sorbafejthetők. Ez azt jelenti, hogy a bonyolult és analitikusan nem leírható függvények egyszerű és jól kezelhető ún. bázisfüggvényekkel adhatók meg, és a bázisfüggvények súlyát a leírandó függvényre jellemző együtthatók határozzák meg. Nem kívánunk a mélyebb matematikai részletekkel foglalkozni, de miután a képfeldolgozás kapcsán gyakran találkozunk a Fourier transzformáció alkalmazásával (egy függvény Fourier transzformációja a függvény, és egy másik függvény szorzatának integrálja, tehát egy integráltranszformáció), annyit még szükséges elmondani, hogy Fourier transzformáció esetén cosinus és sinus függvényt tartalmazó bázisfüggvényeket alkalmaznak, amelyek különböző frekvenciájú hullámfüggvényeket reprezentálnak. Talán ebből a szemléletes leírásból is érzékelhető, hogy egy képfüggvény nagy homogén foltjainak és a lassú átmenetek leírásában a „lomha”, alacsony frekvenciájú, a gyakori és gyors változások leírásában, a nagy frekvenciájú hullámfüggvények kapnak domináló szerepet. Lényegében a nem kívánt frekvenciát tartalmazó bázisfüggvények együtthatóit nullának választjuk, és így módosíthatjuk a kép tartalmát a céljainknak megfelelően. Ezért hívjuk ezt az eljárást frekvencia tartományban történő képjavításnak. Természetesen a képfeldolgozásban bázisfüggvények diszkrét közelítését alkalmazzuk.

Számos képjavítási eljárás létezik, amelyek közvetlenül a digitális képpel, a világosságkódokkal végzett számításokkal próbálja a kívánt hatást elérni. Ezek közül talán leggyakrabban alkalmazott a **konvolúció**. A konvolúció a függvénytanban ismert fogalom, itt most megelégszünk konvolúció digitális képekre alkalmazható megfelelőjének ismertetésével. Az eljárás lényege, hogy minden világosságkód értéket helyettesítünk a képpont környezetétől és a céljaink szerint meghatározott konvolúciós szűrőtől függő értékkel. A konvolúciós szűrő egy mátrix, a mátrix mérete meghatározza a képpontnak azt a környezetét, amit a számításnál figyelembe veszünk, a mátrix elemei pedig a szűrő hatását határozzák meg. A konvolúciós szűrés úgy történik, hogy a szűrőmátrixot egy kitüntetett elemével a vizsgált képpontra illeszkedve a képre helyezzük; majd kiszámítjuk a mátrix által letakart képpontok súlyozott átlagát. Minden képpont súlya az őt lefedő együttható. A vizsgált képpontot helyettesítjük az így kiszámított értékkel, majd egy képponttal tovább csúsztatjuk a mátrixot. Nem kötelező, de általában a mátrix sorainak és oszlopainak számát páratlan számúra szoktuk választani, így a mátrix középső elemét illesztve a kiszemelt képpontra a képpont szimmetrikus környezetét vesszük figyelembe az új érték kiszámításánál. A fentiekben megadott eljárást a  $(2m + 1) \cdot (2n + 1)$  méretű téglalap alakú konvolúciós mátrix esetén az

$$r(k,l) = \sum_{i=-m}^{+m} \cdot \sum_{j=-n}^{+n} q(k+i,l+j) t_{ij}$$

képlettel számíthatjuk ki, ahol  $t_{ij}$  a  $\underline{\underline{T}} = (t_{ij})$  konvolúciós mátrix eleme,  $q(k, l)$  az eredeti képet, és  $r(k, l)$  az eredmény képet jelöli.

Ha számítás eredménye nem egész szám, akkor azt a legközelebbi egész számra kerekítjük, hiszen a világosságkód értékek csak egész számok lehetnek.

Természetesen valamennyi képminőséget javító eljárás közös jellemzője az, hogy nem eredményez olyan információt, amely valamilyen formában ne lett volna rajta a kiindulási képen is.

A képjavítási eljárásokat a **javítandó hibák típusa szerint** is szokták csoportosítani. Mivel a legjellegzetesebb három hibatípus a kontrasztosság elszegényedése, zajok hozzákeveredése a képhez és az élek elmosódása, ezért

- **kontrasztfokozó,**
- **zajelnyomó,**
- **élkiemelő**

eljárásokat szokás megkülönböztetni.

Megemlítjük, hogy sok esetben ugyanaz a módszer más paraméterekkel más célra alkalmazható. A **konvolúciós mátrix** együtthatóinak megválasztása határozza meg például, hogy a konvolúció zajelnyomást vagy élkiemelést eredményez.

### 4.3. Képhelyreállítás

A **képhelyreállítás** (*restoration*) célja a rendelkezésre álló, a különböző zavaró hatások miatt torzított képből az eredetit leginkább megközelítő kép előállítás.

Hatékony képhelyreállításhoz ismernünk kell a **zavaró hatásokat**, és azokat **matematika formulákkal**, vagy legalább valamilyen **közelítő matematikai modellel** le kell tudnunk írni. Ha ez sikerült, akkor meg kell határoznunk azt az inverz transzformációt, amely a modell által leírt hatásokat közömbösíti, és ezt az inverz transzformációt kell alkalmaznunk a rendelkezésre álló képre. Nyilván csak **jó modell** esetén várható eredményes helyreállítás.

Figyelembe véve a potenciális zavaró tényezőket - például az optikai rendszer hibáit, az érzékelők nemlineáris voltát, az atmoszférikus hatásokat, a tárgy elmozdulásából eredő homályosságot, a különböző geometriai torzulásokat - érthető, hogy reménytelen teljes körű modell megalkotására vállalkozni. Többnyire erre nincs is szükség, a megoldandó probléma szempontjából legfontosabb zavaró jelenségek kiküszöbölésére alkalmas közelítő modellek kidolgozása is kielégítő eredményre vezet.

A modell kidolgozása általában méréseken alapul, melyeknek az a céljuk, hogy meghatározzák a vizsgált leképező rendszer hatását egy tetszőleges képmezőre.

A képhelyreállítási eljárások alkalmazásának - a modellalkotás nehézségein túlmenően - az a másik problémája, hogy a felállított modell, illetve az ezen alapuló inverz

transzformáció, még egyszerűsített modell esetén is, rendszerint olyan bonyolult összefüggéseket eredményez, amelyek kiszámításához a feladat megoldásához rendelkezésre álló idő nem elegendő. Kivételt képeznek a geometriai torzulások, melyek viszonylag egyszerűen leírhatók, így ezen a területen a képhelyreállítási módszerek alkalmazása többnyire igen eredményes. A geometriai transzformációkkal a következő fejezetben foglalkozunk, a képjavítási eljárások közül a képfokozási (image enhancement) eljárásokat tekintjük át a továbbiakban.

## 4.4. Világosságkód-transzformációk

Az egyik leggyakrabban előforduló képhiba a nem megfelelő megvilágításból, illetve a leképező rendszerben keletkezett fényvesztéséből adódó kontrasztszegénység. Ez szerencsés esetben egyenletesen jelentkezik az egész képen, de mértéke igen gyakran a képen belül is változik.

A világosságkód transzformációk célja ebben az esetben a **kontraszt növelése** a világosságkódok eloszlásának megváltoztatásával.

Előljáróban ismerkedjünk meg a **hisztogram** fogalmával, mivel a képjavításokban - de egyéb képfeldolgozási eljárásokban is - igen fontos szerepet játszik. A **hisztogram** a világosságkódok adott képen belüli eloszlását megadó függvény.

A világosságkódok többnyire egyenlőközűek, ilyenkor a **hisztogramtáblázat** annyi elemű, amennyi a világosságkódok értékkészlete. A világosságkódok értékkészletét

$$Q = \{0, 1, \dots, J\} \text{-vel jelölve}$$

a táblázat  $q$ -ik eleme:

$$x_q = N_q, \quad \text{és} \quad \sum_{q=0}^J N_q = N$$

ahol  $N_q$ : a  $q$  világosságkódú képpontok száma,  $N$ : a teljes digitális kép pontjainak száma.

A hisztogramot általában **lépcsős függvénnyel** vagy **oszlopdigramm**al ábrázoljuk. A vízszintes tengelyen a lehetséges világosságkódokat mérjük fel, a  $\{q, q+1\}$  intervallumhoz tartozó ordináta pedig a  $q$  világosságkódú képpontok **relatív gyakoriságával** arányos:

$$x_q = c \frac{N_q}{N}$$

Ha a hisztogram normalizált ( $c = 1$ ), akkor  $\sum_{q=0}^J x_q = 1$

Ha a világosságkódot egy folytonos valószínűségi változó közelítéseként kezeljük, akkor annak  $P$  valószínűsége, hogy értéke  $q$  és  $q + \Delta q$  közé esik, valamint a  $s(\xi)$  sűrűségfüggvény között fennáll az

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

összefüggés.

Ez éppen  $x_q = c \frac{N}{N^q}$  hányadossal közelíthető, vagyis a kvantálás finomításával  $x_q \rightarrow s(\xi)$ .

A kép általános jellemzőiről sok minden megállapítható a hisztogram alapján, hasznos információt nyújt nemcsak a képjavításhoz, hanem például a különféle szegmentálási eljárásokhoz is.

Ha a világosságkódok nem töltik ki a megengedett világosságkódok tartományát, akkor a kép kontrasztját a világosságkódok tartományának széthúzásával, **azaz skálázással javíthatjuk**.

A skálázást megfelelően választott átviteli függvényekkel valósíthatjuk meg, amelyek a bemenő kép  $q$  és az eredménykép  $r$  világosságkód értékei közötti leképezést definiálják. A gyakorlatban a világosságkód-transzformációk során legtöbbször **lineáris** vagy **szakaszonként lineáris** átviteli függvényt alkalmazunk.

A **kontrasztkiemelés** annak köszönhető, hogy a bemenő világosságkódok egy részhalmazát képezzük le a kimenő világosságkódok egy szélesebb tartományára.

**Kódzsugorításról** akkor beszélünk, ha a bemenő világosságkódok halmaza bővebb mint a kimenő halmaz. Szükség akkor van rá, ha a feldolgozás eredményeként a megengedettnél (ábrázolhatónál), illetve szükségesnél szélesebb tartományba esnek a világosságkódok. Ez az eljárás nyilvánvalóan információvesztéssel jár.

**Inverz megjelenítést** érhetünk el negatív meredekségű lineáris átviteli függvényvel. Speciális esetekben nem monoton, hanem például **fűrészfog** átviteli függvényt is alkalmazhatunk.

A skálázásokat **átszínezésnek** is nevezzük, mivel színes megjelenítés esetén megváltoztatják a képpontok színét. A skálázást (átszínezést) az átviteli függvény alapján kiszámított táblázattal adhatjuk meg, amely minden bemenő világosságkódhoz megadja a kimenő értéket.

A korszerű képfeldolgozó eszközök a kimenő és bemenő értékek közötti kódcserét,

a memóriában tárolt kép felülírása nélkül, a kimenő videó jel előállításakor hardver úton valósítják meg. Ilyenkor természetesen az átszínezés csak a látványképre vonatkozik. Mivel a hozzárendelést megadó **átszínezőtábla** (*LUT: Look-Up Table*) feltöltése, módosítása igen gyorsan megtörténhet, ezért a módosított látványkép valós időben előállítható.

A képpontok küszöbök szerinti átszínezését (átskálázását) **képvágásnak**, pontosabban: **n-1 küszöb** megadása esetén **n szintre vágásnak** nevezzük. Ekkor az átviteli függvény vízszintes szakaszokból áll. Az  $i$ . és  $(i-1)$ . küszöbök közé eső bemeneti világosságokodok mindegyikéhez ugyanazt - az adott szakasz ordináta értékének megfelelő - kimenő világosságkód értéket rendeljük.

Igen gyakori a **két szintre vágás**, amikor a képpontokat - egy küszöb megadásával - **értékes**, illetve **háttérpontoknak** minősítjük. A kétszintes képet bináris képnek is szokás nevezni, mivel képpontjai egy biten ábrázolhatók. (A két szintrevágásnak tipikus felhasználási területe például: **optikai karakterfelismerő** programok számára tónusos képből bemenő adatok létrehozása.)

A vágásokkal gyors és látványos eredményeket lehet elérni. Problémát okoz, azonban ha a kép **zajjal** torzított, ez ugyanis megváltoztatja az egyes képpontok világosságkódját. Ezért egy megadott küszöb a zajmentes képen azonos tartományba tartozó képpontokat a zaj miatt különböző osztályokba, az eltérő osztályokba tartozókat azonosba sorolhatja. Ilyen esetekben a vágás előtt célszerű **zajszűrést** végezni.

A megközelítés módszerében eltérnek az eddigiektől azok az ugyancsak globális módszerek, amelyekben úgy transzformáljuk a világosságkódokat, hogy a kép **hisztogramja előre meghatározott alakú** legyen.

Az eljárást **hisztogramtranszformációnak** (vagy: **közvetlen hisztogramspecifikációnak**) nevezzük, mivel hatására a kimenő kép hisztogramja előre meghatározott függvény (közelítése) lesz.

A hisztogramtranszformáció során olyan átviteli függvényt kell meghatározni, amelynek alkalmazása esetén a kimenő kép hisztogramja valamilyen, számunkra előnyös, előre meghatározott alakot vesz fel.

Jelölje a bemenő hisztogram ordinátáit  $\chi(q)$ , a kimenőét  $\tau(r)$ . Mint ismeretes ezek arányosak a  $q$ -adik, illetve az  $r$ -edik kvantumszinthez tartozó képpontok számával, és

$$\sum_{q=0}^J \chi(q) = \sum_{r=0}^I \tau(r) = 1$$

ahol a bemenő, illetve kimenő világosságkódok maximális értékét jelöljük **J**-vel, illetve **I**-vel. Monoton transzformációt feltételezve a halmozott részhisztogramoknak minden lépésben meg kell egyezniük:



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

E feltételek teljesülése mellett kell elérni, hogy a kimenő hisztogram valamilyen, számunkra előnyös, előre meghatározott alakot vegyen fel. A probléma explicit megoldása általában nehéz; a numerikus iteráció könnyebben vezet célhoz.

A legegyszerűbb esetben azt kívánjuk elérni, hogy a kimenő kép hisztogramja a vízszintes tengellyel párhuzamos egyenes közelítése legyen, azaz:

$$\tau(r) = \frac{1}{I}$$

Mivel ekkor a kimenő kép hisztogramjában az egyes világosságkód intervallumokba közel egyenlő számú képpont esik, ezért ezt az eljárást **hisztogramkiegyenlítésnek** nevezzük. Ekkor tehát a kimenő kép halmozott részhisztogramjait az  $n/I$  összefüggéssel lehet kiszámítani. A kiegyenlítést úgy végezzük, hogy  $n = 1$  -től indulva rendre kiszámítjuk azokat az  $m_n$  értékeket, amelyekre

$$\sum_{q=0}^m \chi(q) \leq \frac{n}{I} < \sum_{q=0}^{m+1} \chi(q)$$

Az így meghatározott  $n$  értékek kijelölik a kimenő hisztogramban a kvantumszinteket, az  $m_n$  index alapján pedig meghatározhatók az adott kvantumszinthez tartozó világosságkódok.

A hisztogram és a sűrűségfüggvény közötti összefüggés felhasználásával a probléma általánosan is megoldható. Gyakran használatos eljárás a **hisztogram hiperbolizáció**. Ebben az esetben azt szeretnénk elérni, hogy a kimeneti kép hisztogramja hiperbolikus legyen. Ennek azért van különös jelentősége mert - mint ismeretes - az emberi szem a világosságkódokat nem lineárisan, hanem logaritmikusan képezi le. Így a hiperbolizált hisztogramú képre a szem által alkalmazott logaritmikus leképezés eredőjeként az agyban kiegyenlített képet kapunk.

## Gyakorlatok a világosságkód-transzformációkhoz

### 6. Gyakorlat:

Hisztogram\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép2\_Számol

*Ismételjük meg a gyakorlatot úgy, hogy először egy kis részletet (20\*20 pontnál kisebb nem lehet) jelölünk ki a háttérben, majd a munkadarabon! Ismételjük meg a gya-*

korlatot a Kép3, Kép4, Kép6 és Kép1 képekkel! Ne felejtsük a Kép1 esetén a színt RGB-re állítani!

## 7. Gyakorlat:

Átszínezés\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép2\_  
Átszínezőtábla\_Identikus\_Végrehajt\_BW

A legutolsó parancs hatására megjelenik az eredeti kép hisztogramja és az átviteli függvény. Ezt követően válasszuk az Átszínezőtábla menüpontot és próbáljuk végig a különböző hisztogramtranszformációk hatását! A BW gombra kattintva megnézhetjük az átviteli függvényeket. A Szintrevágás és Lyuk parancsoknál az átviteli függvény alakját megváltoztathatjuk, ha annak töréspontjait az egérrel áthelyezzük, a Rajzolásnál az átviteli görbét szabadon átrajzolhatjuk.

Ha meg akarjuk vizsgálni valamelyik transzformált kép hisztogramját, akkor a következő módon helyezzük át az output képet az input kép helyére, majd hajtsuk végre újra az Identikus transzformációt:

Képforrás\_Utolsó eredménykép\_ Átszínezőtábla\_Identikus\_Végrehajt\_BW

Ismételjük meg a gyakorlatot Kép4 és Kép6 képekkel!

## 8. Gyakorlat:

Átszínezés\_Indít\_Szín\_RGB\_Képforrás\_File\_Kép1\_Átszínezőtábla\_  
Identikus\_Végrehajt\_R\_G\_B

Végezzük el a lehetséges hisztogram transzformációkat a kiválasztott színes képre, vizsgáljuk meg az egyes összetevőkre az átviteli függvényeket! Ismételjük meg a gyakorlatot a Kép9-cel!

## 4.5. Zajelnyomás

A képet érő zavaró hatások egy része abban nyilvánul meg, hogy a képpontok eredeti világosságkódja - általában véletlenszerűen - megváltozik. A zavaró hatások eredményét (a képre rakódott „zajt”) az érdemi feldolgozás előtt el kell távolítani a képről, vagy legalábbis a lehető legjobban csökkenteni kell. Az erre irányuló eljárásokat nevezzük **zajszűrésnek**, **zajelnyomásnak**, vagy **simításnak**.

A képjavítások általában, így zajelnyomás vagy simítás is történhet a frekvenciatartományban vagy a képtartományban.

### Simítás a frekvenciatartományban

A frekvenciatartományban történő képjavítás azon alapul (amint azt 4.2. pontban a matematikai háttér részletezése nélkül elmondtunk), hogy a képfüggvények sorbafejtethetők, azaz megfelelően választott bázisfüggvényekkel írhatók le. A képfeldolgozásban is gyakran alkalmazott Fourier transzformáció esetén a bázisfüggvények cosinus és sinus tagot tartalmazó komplex kétváltozós függvények. A bázisfüggvényben szereplő tagok periodikus függvények és síkhullámot írnak le a képsíkban. Az egyes bázisfüggvények abban térnek el egymástól, hogy mindegyiknek más a frekvenciája és ebből adódóan a hullámhossza. A folytonos függvényekre érvényes matematikai eljárásokat - amint az a képfeldolgozásban gyakran tesszük - átvisszük a digitális képekre, értelemszerűen az összefüggések diszkrét megfelelőjét alkalmazzuk. A gyakorlatban diszkrét Fourier transzformációt alkalmazunk a képekre ott, ahol a frekvencia értékek csak egész értéket vehetnek fel. Jelölje a megengedett frekvencia párokat

$$(u, v) \text{ ahol } u, v = 0, \pm 1, \pm 2, \dots$$

A mellőzve a részletes képletet elmondjuk, hogy digitális kép esetén a diszkrét Fourier transzformáció során minden  $(u, v)$  frekvencia párhoz egy  $F(u, v)$  komplex számot határozunk meg, amit Fourier együtthatónak nevezünk. Az elnevezés onnan származik, hogy a kép visszaállítása során az így meghatározott számok lesznek a megfelelő  $(u, v)$  frekvenciájú bázisfüggvény együtthatói. Ennek megfelelően ha az  $F(u, v)$  értékeket valamilyen módon megváltoztatjuk, akkor nem az eredeti, hanem a céljainknak megfelelően megváltoztatott képet kapjuk vissza. Ha bizonyos együtthatókat például egyszerűen 0 értékűre változtatunk, akkor a kép visszaállításakor ennek az együtthatónak megfelelő bázisfüggvény nem fog szerepelni. Ebből már látható hogy a frekvenciatartományban történő képjavítás akkor a leghatékonyabb, ha jól meghatározható frekvenciájú periodikus zajt kell a képből kiszűrni, hiszen akkor egyszerűen a zaj frekvenciájának megfelelő együtthatókat kell 0-ra változtatni.

A CD-n található IMAGES/ZAJ alkönyvtárban periodikus zajok szűrésére találunk szemléltető képeket.

Természetesen általános célú zajszűrést is végezhetünk a frekvenciatartományban, hiszen a zaj a világosságkódokban jelentkező véletlenszerű és hirtelen változást jelent. A képen lévő hirtelen változások leírásában a magasabb frekvenciájú, azaz rövidebb hullámhosszú bázisfüggvények kapnak szerepet, ennek megfelelően, ha „kiszűrjük”

ezeket, azaz nem vesszük figyelembe a kép visszaállításakor, akkor a zaj hatása csökken. A zajelnyomás mértéke attól függ, hogy hány magasabb frekvenciájú bázisfüggvényt hagytunk el. Miután az élek is hirtelen átmenetet jelentenek a képen, ezért sajnos a magas frekvenciájú bázisfüggvények elhagyásával az él átmenetek sem lesznek tökéletesen visszaállíthatók, ezért a visszaállított kép homályosabb lesz mint az eredeti.

Gyakorlatilag a frekvenciatartományban úgy történik a szűrés, hogy először előállítjuk a kép diszkrét Fourier transzformáltját, majd egy alkalmasan választott  $w = H(u, v)$  szűrőfüggvénnyel szorozzuk össze:

$$G(u, v) = H(u, v) * F(u, v),$$

ahol  $F(u, v)$  a képfüggvény Fourier transzformáltja,  $H(u, v)$  a szűrőfüggvény és  $G(u, v)$  a módosított Fourier együtthatókat adja meg. Ezek lesznek a kép visszaállításánál a bázisfüggvények együtthatói, tehát ez lényegében a javított kép Fourier transzformáltja. A  $G(u, v)$ -re alkalmazva az inverz Fourier transzformációt megkapjuk a javított képet.

A zajszűrő függvényeket **aluláteresztő szűrők**nek és az eljárást **aluláteresztő szűrésnek** (*lowpass filtering*) is nevezzük tekintettel arra, hogy a magasfrekvenciás összetevőket kiszűrjük, míg az alacsonyabb frekvenciájú komponenseket változatlanul hagyják.

A legegyszerűbb, az „ideális” **aluláteresztő szűrő** átviteli függvénye a következő:

$$H(u, v) = \begin{cases} 1, & \text{ha } r \leq r_0 \\ 0, & \text{különben} \end{cases}$$

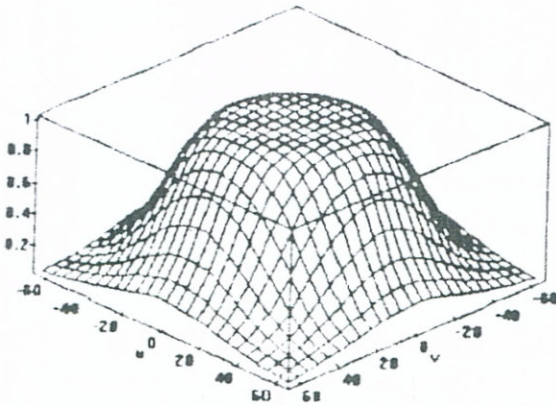
ahol az  $r = \sqrt{u^2 + v^2}$  euklideszi távolságfüggvény megadja az adott frekvenciapár origótól való távolságát, és  $r_0$  egy konstans. (Az  $r$  és  $r_0$  jelentése későbbi szűrőfüggvényekben is változatlan). A képletből látható, hogy az ideális szűrő változatlanul átengedi az  $r_0$  sugarú kör belsejébe eső alacsonyfrekvenciás összetevőket, hiszen ezek változatlanok maradnak, míg a körön kívüli magasabb frekvenciájúakat teljesen kiszűri (0-val szorozzuk a Fourier együtthatókat). Azokat a frekvenciapárokat, ahol a  $H(u, v)$  függvénynek szakadása van vágási frekvenciának nevezzük. A fentiekben megadott szűrőfüggvény az  $(u, v)$  síkra helyezett, a  $w$  tengelyre nézve szimmetrikus, egységnyi magasságú,  $r_0$  sugarú hengerrel ábrázolható, amelynek az alaplapja középpontjával az origóra illeszkedik. A függvény felületét a henger fedőköre és az  $(u, v)$  síknak a függvény értelmezési tartományába eső része alkotja a henger alaplapjának kivételével.

Az ideális szűrő hátránya, hogy a magasfrekvenciás összetevőket teljes mértékben kiszűri, ezzel az éleket is elsimítja, így nagymértékben homályosítja a képet. További hátrány, hogy az elhagyott és megtartott frekvenciák közötti éles vágás miatt zavaró

hatás, az ún. **gyűrődés** jelenik meg a visszaállított képen, ami azt jelenti, hogy a képen periodikusan jelentkező foltok tűnnek fel. Ennek elkerülésére olyan szűrőfüggvényeket szoktak alkalmazni, amelyek fokozatosan módosítják a Fourier együtthatókat.

Az alábbi képlettel megadható **Butterworth szűrő** például kismértékben a magasfrekvenciás összetevőket is átengedi, így kevésbé homályosodik el a kép; és a sima átmenet miatt a gyűrődés sem lép fel:

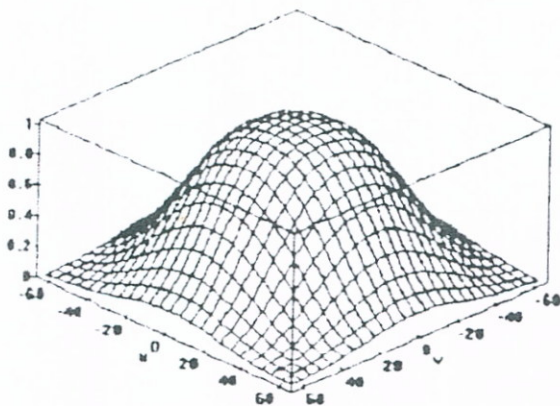
$$H(u, v) = \frac{1}{1 + (r/r_0)^{2n}}$$



7. ábra

Ez a függvény a  $w$  tengelyre körszimmetrikus, sima átmenetet megvalósító (szakadás nélküli), kondér alakú felületet határoz meg, amelynek maximuma az origóban van és a maximum értéke egyhez közelít, míg az origótól távolodva egyre kisebb értékeket vesz fel és 0-hoz közelít. A kitevőben szereplő  $n$  értékének változtatásával befolyásolhatjuk a felület lefutását. A képletből az is kiolvasható, hogy ebben az esetben az  $r_0$  vágási frekvencia azokat a frekvencia párokat adja meg, ahol a

szűrőfüggvény értéke 0.5-re csökken, tehát ennél a küszöbnél nagyobb frekvenciához tartozó bázisfüggvényeket 50%-nál kisebb mértékben fogjuk figyelembe venni az inverz transzformáció során. Miután a képlet alapján nehezebb elképzelni ezt a függvényt mint az ideális szűrőt, ezért szemléltetés céljából a 7. ábrán bemutatjuk a függvény által meghatározott felület drótvázis modelljének perspektivikus képét. Az értelmezési tartomány a -64-től +64-ig terjedő egész számok, és a függvény kirajzolása az  $n=6$ ,  $r_0=50$  paraméterekkel történt. A koordináta rendszer tengelyeit a szokásostól eltérően nem tüntettük fel, hogy a felület alakja jobban látható legyen, de a felületet befoglaló téglatest élein bejelöltük a koordinátákat.



8. ábra

Az alábbi képlettel definiált **exponenciális szűrő** egy a  $w$  tengelyre körszimmetrikus harang alakú felületet ír le.

$$H(u, v) = \exp[-(r/r_0)^n]$$

A meredekebb lecsengése miatt a Butterworth szűrőnél kissé jobban homályosít, de korántsem annyira, mint az „ideális” szűrő, gyűrődés pedig itt sem lép fel. Ennek a

függvénynek a drótvázás modellje a 8. ábrán látható. Az értelmezési tartomány és  $r_0$  értéke ugyanaz mint az előző függvényénél, viszont  $n=3$  értékkel történt a görbe kirajzolása, így a görbék lefutása nem hasonlítható össze a két ábrán. Akkor lenne látható, hogy az exponenciális szűrő lecsengése gyorsabb, ha azonos  $n$  értékek mellett történt volna a kirajzolás.

A **trapéz szűrő** átmenetet képez az „ideális” és a sima szűrők között, ennek  $H(u, v)$  függvénye:

$$H(u, v) = \begin{cases} 1 & \text{ha } r < r_0 \\ \frac{r - r_1}{r_0 - r_1} & \text{ha } r_0 \leq r \leq r_1 \\ 0 & \text{ha } r > r_1 \end{cases}$$

A szűrő onnan kapta a nevét, hogy a megadott függvény a  $w$  tengelyre körszimmetrikus csonka kúpot ír le, aminek minden, a  $w$  tengelyen áthaladó síkkal való metszete egy egyenlőszárú trapéz. A csonka kúp fedőlapjának sugara  $r_0$ , az alaplapjának sugara  $r_1$ . Az  $r_0$  sugarú körön belül változatlanok maradnak a Fourier együtthatók, az  $r_1$  sugarú körön kívül 0-val, míg a két sugár által meghatározott körgyűrűn belül 1 és 0 közötti fokozatosan csökkenő értékkel szorozzuk meg a Fourier együtthatókat.

Részletesebb magyarázattól eltekintve az alábbiakban felsoroljuk azokat a frekvenciatartományban alkalmazható szűrőket, amelyek a gyakorlati zajszűrési feladatoknál periodikus zajok hatékony szűrésére használhatók:

- Inverz szűrő
- Wiener szűrő
- CLS szűrő
- Geometriai átlagoló szűrő
- Sáv-szűrő

## Simítás a képtartományban

A 4.2. pontban már ismertettük a képjavításban nagyon gyakran alkalmazott módszert a **konvolúciós szűrést** és elmondtuk, hogy konvolúciós szűrő együtthatóinak megválasztásával különböző, egymással ellentétes hatást érhetünk el.

Zajelnyomás esetén többnyire olyan szűrőket (konvolúciós mátrixokat) használnak, melynek valamennyi eleme pozitív, és ezek vagy mind azonosak (azaz minden figyelembe vett szomszéd azonos mértékben befolyásolja a végeredményt), vagy a centrumtól távolodva csökken az értékük. Elemeinek összege a gyakorlati alkalmazá-

sok többségében 1, ekkor a homogén foltokat a konvolúció nem változtatja meg. A fígyelembe vett szomszédság a gyakorlatban 3\*3, illetve 5\*5 képpontra terjed.

Az egyszerű programozás érdekében általában négyzetes alakú szűrőket alkalmazunk; ezen belül a zérus elemek-vel hagyhatók ki a szükségtelen elemek.

Például az alábbi 3\*3 -as szűrők zajelnyomó, simító hatásúak.

$$\underline{T}_1 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \underline{T}_2 = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \underline{T}_3 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Számítsuk ki például, hogy a  $\underline{T}_2$  konvolúciós szűrő milyen értékkel fogja helyettesíteni az alábbi képrészlet középső képpontját!

$$\begin{array}{ccccc} 11 & 12 & 12 & 14 & 13 \\ 12 & \boxed{14} & \boxed{12} & \boxed{15} & 12 \\ 13 & \boxed{15} & \boxed{62} & \boxed{61} & 13 \\ 11 & \boxed{12} & \boxed{14} & \boxed{13} & 14 \\ 13 & 15 & 12 & 13 & 12 \end{array}$$

A 4.2 pontban megadott összefüggés felhasználásával kapjuk az új értéket:

$$r(k,l) = \frac{1}{10} (14 + 12 + 15 + 15 + 2 * 62 + 61 + 12 + 14 + 13) = 28$$

Látható tehát, hogy a kiugróan magas 62 érték lényegesen kisebb lesz, ezért a szűrés utáni képen kevésbé lesz feltűnő. Számítsuk most ki hogy a képrészlet bal felső 3\*3 ablakának középső 14 értékű világosságkódja hogyan fog változni. Erre a pontra:

$$r(k,l) = \frac{1}{10} (11 + 12 + 12 + 12 + 2 * 14 + 12 + 13 + 15 + 62) = 18$$

Máris észrevehetjük ezeknek a szűrőknek a hátrányát, nevezetesen a zaj környezetben lévő értékek is megváltoznak, zaj világosságkód értékben jelentkező hatása csökken ugyan, de folttá szélesedik, egyben az élek is elmosódnak. A szűrő méretének növelésével a részletek összemosódása is fokozódik.

Mivel ezek a szűrők a magasfrekvenciás zajokat szűrik ki, ezért mint már említettük aluláteresztő konvolúciós szűrőknek is nevezzük őket. Vegyük észre, hogy  $\underline{T}_1$  lényegében átlagolást definiál, ezért ezeket **átlagoló szűrőknek** nevezzük.

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.

Kedvező tulajdonságai miatt szívesen alkalmazzák zajszűrésre 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 világosságkódjai alkotják. A szűrést úgy valósítjuk meg, hogy minden képpont világosságkódját kicseréljük a ráillesztett szűrő által meghatározott minta mediánjával. Ha például **3\*3** méretű mediánszűrést végzünk, akkor a 9 db érintett képpontot nagyság szerint sorba rendezzük, és az ablak középpontjában lévő világosságkódot a sorba rendezett értékek középső elemével helyettesítjük.

A konvolúciós típusú zajszűrés és a mediánszűrés között lényeges különbség van. Az átlagoló szűrő - amint az előbbi példa is mutatta - „szétkeni” a zajt, azaz csökkenti a kiugró képpontok értékét de a zaj környezetében lévőket megnöveli. Ugyanakkor a medián szűrő gyakorlatilag kiszűri a zajt és a homogén részeket változatlanul hagyja.

A következő példák konkrét képrészleteken mutatják meg a mediánszűrő előnyeit. Vizsgáljuk meg mi történik, ha az előbbi képrészletre alkalmazzuk a medián szűrést. A kerettel megjelölt 3\*3-as szűrő által meghatározott világosságkódokat rendezzük növekvő sorba:

**12 12 13 14 14 15 15 61 62**

A **62**-es világosságkódot a mediánnal, **14**-gyel helyettesítve a közel homogén képrészletből kiszűrtük a kiugróan magas értéként megjelenő világosságkódot. A következő lépésben a szűrőt egy oszloppal jobbra tolva és az eljárást megismételve a **61**-es érték helyére szintén **14** fog kerülni. A medián szűrő az impulzusszerűen jelentkező zajt teljesen elnyomja, ha az impulzus periódusa kisebb, mint a szűrő szélességének fele.

(Az átlagoló szűrő esetén a **62**-es világosságkód helyére **24**, a **61**-es helyére szintén **24** kerül volna, azaz a pontszerűen jelentkező zaj hatását csak csökkenteni képes az átlagoló szűrő, megszüntetni nem. Ugyanakkor a kiugró képpontérték szomszédainak értéke megemelkedik.)

Bizonyos esetekben nem egyetlen képen belül, 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

$$q_i(k, l) = q(k, l) + z_i(k, l)$$

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



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

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 minden képpontban nulla átlagú.

## Gyakorlatok a zajelnyomáshoz

### 9. Gyakorlat:

Zajszűrés\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép5\_  
Szűrő mérete\_3\_Végrehajt

*Az átlagoló szűrés hatását vizsgálhatjuk ezzel a gyakorlattal. Változtassuk a szűrő méretét (maximum 33 és páratlan számot kell megadni), és figyeljük meg az eredményt!*

### 10. Gyakorlat:

Konvolúció\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép4\_  
Szűrő mérete\_5\*5

*Ezt követően jelöljük ki a kép felső részét az egérrel, és utána folytassuk!*

Ált. szűrők\_Átlagoló\_Végrehajt

*Most jelöljük ki a kép alsó részét egérrel majd folytassuk!*

Ált. szűrők\_Súly átlagoló\_Végrehajt

*Figyeljük meg a két részképen a különbséget!*

### 11. Gyakorlat:

Medián\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép5\_Végrehajt

*Ezt követően ismételjük meg a 9. Gyakorlatot 3\*3 szűrő mérettel! Hasonlítsuk össze a két eredményt!*

## 12. Gyakorlat:

Konvolúció\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép5\_  
Szűrő mérete\_3\*3\_Szűrési mód\_Def.szűrővel\_Szűrődef.

*Jelöljük ki a szűrőt úgy, hogy a középső eleme 1 az összes többi 0 legyen, utána:*

Végrehajt\_Képforrás\_Utolsó eredménykép\_Szűrési mód\_Ált. szűrők

*Most jelöljük ki a kép egyik felét az egérrel, és utána folytassuk:*

Átlagoló\_Végrehajt\_\_Képforrás\_Utolsó eredménykép\_Kilép

*Ezzel a feldolgozási sorral elértük, hogy a kiindulási kép helyén egy olyan kép van amelyiknek az egyik fele a zajos kép, a másik fele 3\*3 átlagoló szűrővel van megszűrve. Folytassuk a gyakorlatot:*

Medián\_Indít\_Végrehajt

*Mivel magyarázza, hogy az átlagoló szűrővel már megszűrte képrészen nem igazán javít a medián szűrés?*

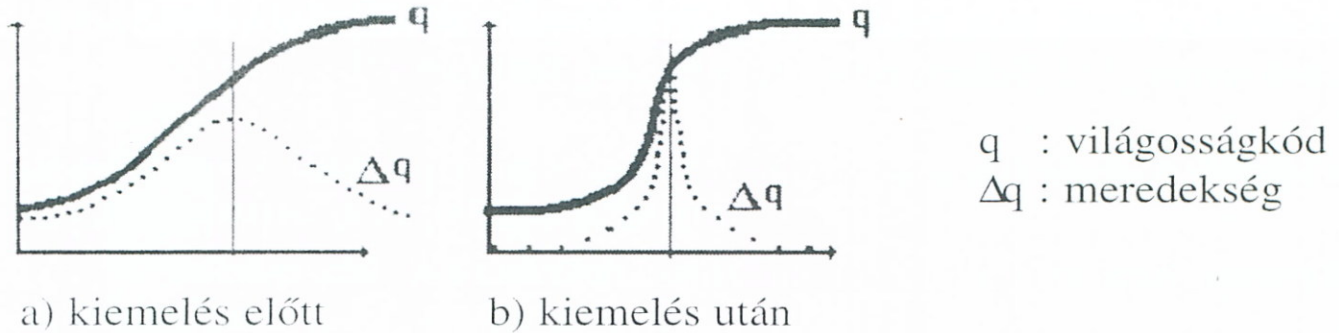
## 4.6. Élkiemelés

A képeken 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 bizonyították, hogy szubjektíven előnyösebb érzetet kelt a túlhangsúlyozott élekkel rendelkező kép, mint a valóság-hű ábrázolás.

Az eljárások célja az egyes képrészletek közötti **átmeneti tartomány szűkítése**, a határátmenetek hangsúlyozása, meredekebbé tétele, az elmosódások korrigálása.

Az élkiemelés hatását szemlélteti a 9. ábra. Az ábrán  $q$  jelöli egy képsor mentén a világosságkód értékeket. Digitális kép esetén  $q$  egész értékeknél van értelmezve és egész értékeket vehet fel. Tehát lényegében diszkrét pontokat kellett volna feltüntetni a rajzon, de a jobb láthatóság érdekében vastag folytonos vonallal jelöltük azt a gör-

bét, amelyre illeszkednek a  $q$  értékek. A meredekséget a  $\Delta q$  pontozott vonal jelöli. Digitális kép esetén, miután a lépésköz egységnyi, a meredekség a szomszédos világosságkód értékek differenciájával adható meg. Az élkiemelés előtti állapotot az (a), az élkiemelés utánit a (b) ábra mutatja.



9. ábra

## Élkiemelés a frekvenciatartományban

Az előbbieken ismertettük a frekvenciatartományban történő képjavítás lényegét és a zajelnyomásnál alkalmazott szűrőket. Az élkiemelésnél lényegében ugyanazt az eljárást kell követni mint a zajszűrésnél, de a szűrők megválasztásához azt kell tudni, hogy az élek (a zajokhoz hasonlóan) a világosságkódokban történő nagyobb változásként jelentkeznek, így a kép Fourier-transzformáltjának magasabb frekvenciájú komponenseit adják. Ezért **felüláteresztő szűrők** alkalmazásával, azaz **felüláteresztő szűréssel** (highpass filtering) tudjuk az élesítést elvégezni. Ezek a szűrők kiszűrik az alacsonyfrekvenciás komponenseket és változatlanul átengedik a magas frekvenciatartományba esőket. Ebből rögtön következik, hogy ha a kiindulási kép zajos, akkor nem célszerű ezt az eljárást alkalmazni, hiszen zaj is megmarad a képen.

A felüláteresztő szűréshez a korábban megismert aluláteresztő szűrőfüggvényeket úgy kell megváltoztatni, hogy most éppen azokat a frekvenciákat engedjék át amelyeket korábban kiszűrtek, és azokat szűrjék ki amelyeket zajelnyomásnál megtartottak.

A **ideális felüláteresztő** szűrőt úgy kapjuk, hogy zajelnyomásnál megismert függvényben a kijelölt tartományokhoz rendelt konstansokat felcseréljük, azaz:

$$H(u, v) = \begin{cases} 0, & \text{ha } r \leq r_0 \\ 1, & \text{különben} \end{cases}$$

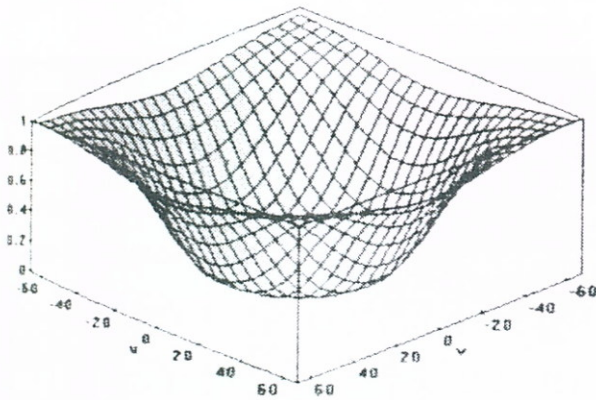
ahol  $r = \sqrt{u^2 + v^2}$  és  $r_0$  egy konstans, a vágási frekvencia. Az  $r$  és  $r_0$  jelentése a további felüláteresztő szűrőknél is változatlan. A függvény most egy, az  $(u, v)$  síkkal pár-

huzamos, egységnyi magasságban elhelyezkedő síklap, amelyen az origó körüli  $r_0$  sugarú kör felett egy lyuk van, és ebben a tartományban az  $(u, v)$  síkra illeszkedő körlap adja meg a függvény által meghatározott felületet.

Az ideális szűrő adott frekvencia alatt **minden** összetevőt levág, azaz teljesen kiszűri a kép alacsonyfrekvenciás összetevőit, és hasonló problémák lépnek fel mint az ideális aluláteresztő szűrő esetén. Ezért a sima átmenetet megvalósító aluláteresztő szűrőknek a felüláteresztő megfelelőit érdemes alkalmazni.

A felüláteresztő **Butterworth-szűrőt** úgy kapjuk meg, hogy a képletben  $r$ -t és  $r_0$ -t felcseréljük:

$$H(u, v) = \frac{1}{1 + (r_0/r)^{2n}}$$

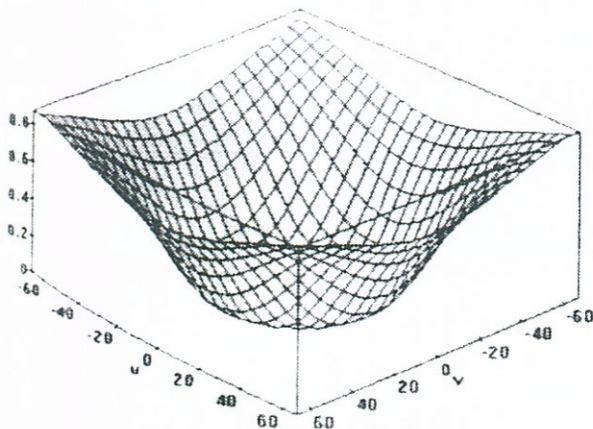


10. ábra

A 10. ábra mutatja a függvény által meghatározott felületet. Látható, hogy az origóhoz közeli frekvenciapárokhoz 0-hoz közeli értékek tartoznak, ami azt jelenti, hogy 0-hoz közeli számmal szorozzuk meg a Fourier együtthatókat, tehát a kép helyreállításában ezeknek a bázisfüggvényeknek a szerepe elenyésző lesz. A magasabb frekvenciapárokat 1-hez közeli értékkel kell szorozni, ezért ezeket a bázisfüggvényeket közel teljes mértékben figyelembe vesszük a kép helyreállításakor. A felület kirajzolása azonos paraméterekkel történt mint a megfelelő aluláteresztő szűrőnél.

A felüláteresztő **exponenciális szűrő képlete** a következő:

$$H(u, v) = \exp[-(r_0/r)^n]$$



11. ábra

Látható, hogy ebben az esetben is egyszerűen csak megcseréltük az aluláteresztő szűrő képletében az  $r$ -t és  $r_0$ -t. A képlet által leírt felületet a 11. ábra mutatja, a paraméterek azonosak mint a megfelelő aluláteresztő szűrőnél.

A felüláteresztő trapéz szűrő képlete:

$$H(u, v) = \begin{cases} 0 & \text{ha } r < r_1 \\ \frac{r - r_1}{r_0 - r_1} & \text{ha } r_1 \leq r \leq r_0 \\ 1 & \text{ha } r > r_0 \end{cases}$$

Ennek a  $H(u, v)$  függvénynek az ábrája egy, az origó körüli  $r_1$  sugarú körlapból, egy erre állított felfelé szélesedő csonkakúp palástjából, és egy erre illeszkedő síklapból áll, amely éppen a csonkakúp felső  $r_0$  sugarú fedő körlapjánál lyukas.

## Élkiemelés a képtartományban

Mint azt a **zajelnyomás** c. pontban láttuk, a zajelnyomás egyik lehetséges módja a képpontok környezetében végzett átlagolás. Az átlagképzés az integrálás diszkrét megfelelője, kézenfekvő tehát, hogy a kép élesítését **differenciálással** próbáljuk elérni. Általánosabban megfogalmazva abból indulunk ki, hogy az élek a képfüggvényen a világosságkódok valamilyen irányban jelentkező lényeges változásaként jelentkeznek. A folytonos  $f(x, y)$  függvény változásainak leírására jól kidolgozott matematikai módszerek állnak rendelkezésre, ezért értelemszerűen ezek diszkrét közelítéseivel érdemes a digitális képek esetén próbálkozni.

Először nézzük az egyszerűbb, egyváltozós esetet. A 9. ábrán egy képsort ábrázoltunk, amelyben egy jól látható emelkedés, él található. Az élgyanús pontokat ott szokták kijelölni, ahol az átmeneti görbe szélsőértéke, vagyis a 9. ábrán szaggatott vonallal rajzolt differencia-görbének maximuma van. Megkeresésükre természetesnek tűnik a differencia-hányadosok elemzése.

Kétváltozós esetben, márpedig a képek kétváltozós függvénnyel írhatók le, az élek tetszőleges irányban helyezkedhetnek el, ezért ha soronként végeznénk az élkiemelést, akkor a soriránnyal párhuzamos éleket nem érzékelnénk, hiszen ideális esetben az él mentén végighaladva nem jelentkezik változás a világosságkódokban. Ezért a gyakorlatban a legalább két irányt figyelő operátorok terjedtek el.

Elöljáróban még elmondjuk, hogy az élkiemelést nagyon sokszor a korábban ismertetett konvolúciós szűrőkkel valósíthatjuk meg, a szűrő együtthatóinak megfelelő megválasztásával.

A továbbiakban néhány ismert élkiemelési eljárást mutatunk be, utalva a kiindulásul használt matematikai fogalmakra.

A **differencia operátor** egy élkiemelési eljárás, amely a matematikában ismert gradiens tulajdonságait használja ki. Az  $f(x, y)$  folytonos függvény gradiense az  $(x, y)$  pontban

$$\underline{G}[f(x, y)] = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Amint az a függvénytanból ismert a gradiens a maximális növekedés irányába mutat, és nagysága a  $\underline{G}$  irányba egységnyi távolságra eső változást adja meg, ami a

$$G = \sqrt{[(\partial f / \partial x)^2 + (\partial f / \partial y)^2]}$$

képlettel számítható. Digitális kép esetén a parciális deriváltakat a differenciákkal helyettesíthetjük, és a

$$G = \left\{ [q(k, l) - q(k+1, l)]^2 + [q(k, l) - q(k, l+1)]^2 \right\}^{1/2}$$

vagy a

$$G \cong |q(k, l) - q(k+1, l)| + |q(k, l) - q(k, l+1)|$$

közelítést alkalmazzuk.

Nézzük meg milyen eredményt ad ez utóbbi összefüggés a következő képrészleten:

34	36	35	39	57
37	36	38	45	56
36	39	40	57	59
37	41	58	60	61
35	42	56	62	61

Számítsuk ki  $G$  értékét az  $(1,1)$  és a  $(3,3)$  koordinátájú képpontra:

$$G[q(1, 1)] = |34 - 36| + |34 - 37| = 5, \quad \text{míg}$$

$$G[q(3, 3)] = |40 - 57| + |40 - 58| = 35.$$

Látható tehát, hogy ahol jelentősen emelkedtek a világosságkód értékek ott a differencia operátor is nagyobb értékeket adott. Ha a kép minden pontjára kiszámoljuk a

gradiens nagyságát valamelyik megadott közelítő képlettel, akkor egy megfelelően választott küszöb értékkel két szintre vághatjuk a képet, és így megkapjuk az élgyanús pontokat. Azért nem mondhatjuk egyértelműen, hogy az élpontokat, mert ahol a képen zaj van ott is nagy lesz a gradiens.

Egy másik közkedvelt élkiemelő eljárás a Laplace operátor alkalmazásán alapszik. Ennél az élkiemelést a **Laplace operátor** közelítéseiből származó megfelelő konvolúciós mátrixok alkalmazásával valósítjuk meg. Az eljárás alapja, hogy a másodrendű parciális differenciálhányadosok az elsőrendű deriváltak szélsőértékhelyeinek meghatározásában játszanak szerepet. Amint láttuk az éleknél a differenciálhányados függvénynek lokális szélsőértéke van. Természetesen digitális képek esetén ebben az esetben is a folytonos függvényeknél alkalmazott képletek diszkrét megfelelőjét, esetleg valamilyen közelítést használunk. A függvénytanban ismert  $\nabla^2$  Laplace operátor az  $f(x, y)$  függvényhez a

$$\nabla^2 \{f(x, y)\} = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$

függvényt rendeli hozzá.

A részletek mellőzésével megadjuk a fenti képletből a differenciálhányadosok differenciahányadosokkal való közelítésével kapott kifejezést, amit a digitális képfeldolgozásban alkalmazunk:

$$\nabla^2 [q(k, l)] = q(k+1, l) + q(k, l+1) + q(k-1, l) + q(k, l-1) - 4q(k, l)$$

Vegyük észre, hogy alábbiakban megadott  $L_1$  -el jelölt konvolúciós szűrő éppen a megadott képletnek felel meg. A másik két Laplace jellegű szűrő a képpont nagyobb környezetét veszi figyelembe a számítás során.

Figyeljük meg, hogy a szűrőelemek összege (a szűrő súlya) mindig 0.

$$\underline{\underline{L_1}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \underline{\underline{L_2}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \underline{\underline{L_3}} = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix}$$

Végezetül ismertetjük az előnyös tulajdonságai miatt kedvelt Sobel operátort, amely az előbbieken megismert differenciáloperátor 3\*3 ablakra (rácspontra) történő kiterjesztése. Amint azt korábban láttuk a differencia operátor a gradiens adaptációja digitális képre. A gradiens nagyságának meghatározásához két egymásra merőleges irányban a szomszédos képpontok különbségét számítottuk ki.

Az egyszerűbb jelölés érdekében jelöljük a  $3 \times 3$  ablakban lévő  $q(k, l)$  képpontokat az alábbiak szerint:

$$\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array}$$

Ezt a jelölést használva definiáljuk a  $3 \times 3$  ablakban számított differenciákat a következő módon:

$$G_k = (g + 2h + i) - (a + 2b + c)$$

$$G_l = (c + 2f + i) - (a + 2d + g)$$

Látható, hogy a  $G_k$  a sorirányú, a  $G_l$  az oszlopirányú különbséget adja meg. A különbségek kiszámításánál a középső értéket tapasztalati megfontolások alapján kétszeres súllyal vesszük figyelembe. A differenciák felhasználásával:

$$G = \sqrt{G_k^2 + G_l^2}, \text{ vagy a}$$

$$G = |G_k| + |G_l|$$

közelítéssel határozzuk meg a gradiens nagyságát. Megfelelően választott küszöb felhasználásával két szintre vágva a képet jelölhetjük ki az élgyanús pontokat.

Vegyük észre, hogy  $G_k$  és  $G_l$  a következő két konvolúciós szűrőből számítható ki:

$$\underline{\underline{G_k}} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \underline{\underline{G_l}} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

A teljesség igénye nélkül hatékony élkiemelő szűrő a:

- Kirsh szűrő
- Robnson szűrő
- Frei-Chen szűrő
- Prewitt szűrő (Sobel-hez hasonló)



## Gyakorlatok az élkiemeléshez

### 13. Gyakorlat:

Konvolúció\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép2\_  
Szűrő mérete\_3\*3\_Szűrési mód\_Ált. szűrők\_Átlagoló\_Végrehajt\_  
Képforrás\_Utolsó eredménykép\_

*Az első két sor az input kép helyére egy átlagoló szűrővel megszürt képet tesz. Ezután definiáljunk egy olyan szűrőt, amelynek középső eleme -1, a tőle jobbra lévő 1, az összes többi 0:*

Szűrési mód\_Def.szűrővel\_Szűrődef.

*Vegyük észre, hogy ezzel a szűrővel végrehajtott konvolúció sor irányban számítja ki a szomszédos képpontok különbségét! Jelöljük ki a kép felső felét, és hajtsuk végre a műveletet, majd a Profil funkcióval vizsgáljuk meg az él átmeneteket az eredeti képen és a differencia képen:*

Végrehajt\_Képforrás\_Utolsó eredménykép\_Kilép\_Profil\_Indít

*Először jelöljük ki egész rövid szakaszt vízszintesen egy él felett, majd a ugyanazon él folytatásának differencia képén! Ismételjük meg a kijelölést a kép teljes szélességében!*

### 14. Gyakorlat:

Laplace operátor\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép1\_  
Végrehajt

*Nézzük meg az eredménykép hisztogramját:*

Képforrás\_Utolsó eredménykép\_Kilép\_Hisztogram\_Indít\_Számol

*A hisztogramból láthatjuk, hogy ez a funkció az eredményt vágással számolja, azaz a negatív értékeket elhagyja. Végezzük el a Laplace operációt a Konvolúció funkcióval:*

Konvolúció\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép1\_  
Szűrő mérete\_3\*3\_Szűrési mód\_Def.szűrővel\_Szűrődef.

*A szűrőmátrix középső eleme legyen -4, a felette, alatta, tőle jobbra és balra lévő elemek értéke legyen 1, a többi 0 (Laplace operátor), majd:*

Végrehajt\_Képforrás\_Utolsó eredménykép\_Kilép\_Hisztogram\_Indít\_  
Számol

*Látható, hogy a hisztogram szimmetrikus, de nem a 0, hanem a 128 értékre. Ennek az a magyarázata, hogy a Konvolúció funkció programban lévő implementációja minden esetben skálázza az eredményt.*

## 15. Gyakorlat:

Sobel\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép3\_Típus\_Él\_Végrehajt

*Változtassuk meg a küszöböt, figyeljük meg mi történik! Változtassuk a Típust az Él képpel együtt változatra! Ismételjük meg a Gyakorlatot a Kép2-vel és a Kép1, Kép9 színes képekkel! Ne felejtsük a Színt RGB-re állítani az utóbbi két esetben!*

## 4.7. Többsávos képek javítása

A többsávos kép ugyanarról a területről (jelenetről) egyidejűleg több frekvenciatartományban készített felvételek együttese. Bár a „többsávos kép” fogalma eredetileg összekapcsolódott az űrfelvételekkel, valójában ide tartozik az összes olyan kép, amely egynél több sávból áll, és az egyes sávok pontról-pontra azonos területre vonatkozó információt tartalmaznak (színes kamerakép, cián-vörös-sárga-fekete összetevőkre bontott nyomdai kép, vagy pl. ugyanarról a jelenetről két különböző időpontban készített és egymáshoz kapcsolt felvételek). Ugyanakkor nem tartoznak ide az olyan felvétel-együttesek, melyek geometriai értelemben véve nem azonosak, azaz pontról-pontra eltérő a jelentésük (pl.: sztereó képpár, mozgó objektumról készült felvételsorozat).

Az előzőekben megismert képjavítási eljárások természetesen alkalmazhatók a többsávos képek egyes összetevőire külön-külön is, számos olyan eljárás is ismeretes azonban, melyek kihasználják a sávok közti kapcsolat által nyújtott többletinformációt.

## Képsávok közti műveletek

Így nevezzük azokat az eljárásokat, amelyekben két vagy több képsáv azonos pozíción levő képpontjai között értelmezett műveletekkel állítjuk elő az eredményképet.

A képsávok közti aritmetikai és logikai műveletek egyrészt lényegkiemelésre (pl.: **átszínezésre**), a sávok közti összefüggések kiemelésére alkalmasak, másrészt bizonyos hibák kiküszöbölésére használhatók. Néhány kiragadott példa:

**Kivonással** kiemelhetők a képek közti változások, eltérések. Ugyancsak kivonással lehet a hasznos képre additív zajként rakódó, és hasznos információt nem hordozó háttér-megvilágítás, beszűrődő fény hatását csökkenteni.

**Osztással** pl. megvilágítási egyenetlenségek kiküszöbölésére van lehetőség. (Az ún. **shading korrekció** esetén a megvilágítási hibákat tartalmazó referenciaképpel való osztás segít a korrekt eredménykép létrehozásában.)

**Összetett műveletek** (pl. lineáris kombináció) lényegkiemelésre használhatók. Pl.: multispektrális és multitemporális úrfelvételeken ismert összetevőjű képrészek kiemelésének az osztályozásnál gyorsabb módja a műveletekkel való összekapcsolás, bár a lényegi eltérések miatt nem mindig alkalmazható.

## Gyakorlatok a többsávós képek transzformációjához

### 16. Gyakorlat:

Kétképes műveletek\_Indít\_Szín\_fekete/fehér\_1.Képforrás\_File\_Kép6\_  
2.Képforrás\_File\_Kép7\_Végrehajtási mód\_Skálázás\_Művelet

*Állítsuk be és végezzük el a különböző műveleteket ! Mi történik, ha változtatjuk a Végrehajtási módot?*

## Ellenőrző kérdések

1. Melyik két nagy csoportba sorolhatók a képjavítási módszerek az eljárás célja tekintetében?
2. Mi a képhelyreállítás (restoration) célja?

3. Mi a képfokozás (enhancement) célja?
4. Hogyan csoportosíthatjuk a képjavítási eljárásokat a javítandó hibák jellege alapján?
5. Mire van szükség képhelyreállítási eljárások kidolgozásához?
6. Milyen tényezők nehezítik a képhelyreállítási eljárások létrehozását?
7. Milyen esetekben használjuk a világosságkód-transzformációkat?
8. Ismertesse a hisztogram fogalmát!
9. Milyen hatásokat érhetünk el a lineáris vagy szakaszosan lineáris átviteli függvényekkel?
10. Ismertesse a hisztogramtranszformációs (közvetlen hisztogram specifikációs) eljárást!
11. Hogyan lehet zajelnyomást megvalósítani Fourier transzformációval?
12. Mit jelent a konvolúciós szűrő digitális kép esetén?
13. Hogyan lehet zajelnyomást megvalósítani konvolúciós szűrők segítségével?
14. Mit jelent a mediánszűrés?
15. Hasonlítsa össze konvolúciós és medián szűrő hatását!
16. Mivel magyarázza, hogy az átlagoló szűrővel már megszűrt képrészen nem igazán javít a medián szűrés?
17. Mi az élkiemelés célja?
18. Hogyan végezhetünk élkiemelést Fourier transzformációval?
19. Milyen konvolúciós jellegű szűrőket lehet alkalmazni élkiemelésre?
20. Mit nevezünk többsávós képnek?
21. Milyen hatások érhetők el képsávok közötti műveletekkel?

## 5. GEOMETRIAI KORREKCIÓ

### 5.1. Bevezetés

A geometriai korrekció az egyik leggyakoribb képhelyreállítási eljárás. Általános esetben egy (vagy több, geometriai szempontból **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 pontossággal teljesüljön.

A **geometriai korrekció**, mint alapvető képfeldolgozási művelet, **célja** szerint lehet

- **képjavítás**. Ekkor geometriai hibákban jelentkező torzítás megszüntetése a feladat, például: a képfelvevő rendszer optikai hibái, a perspektív torzítás, vagy felvétel közbeni elmozdulás miatt.
- **képkorrekció**. Ekkor pl. a képek illesztése - **szerkesztése** -, a vetületi rendszerbe való transzformáció, illetve a helyes méretarányok megvalósítása a feladat.

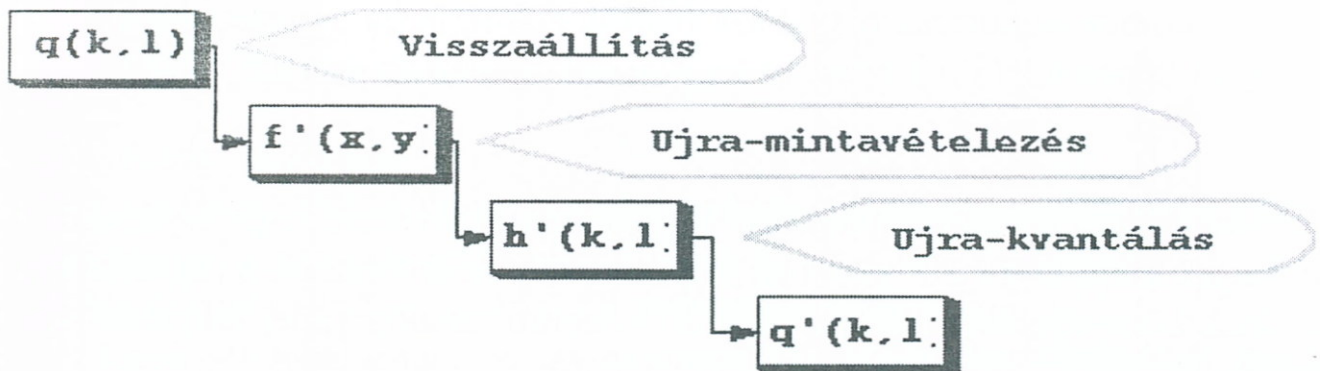
Általában elmondható, hogy a geometriai korrekcióra, mint **képjavításra** a korszerű képalkotó eszközök fejlettsége miatt csak a különösen nagy precizitást igénylő felhasználási területeken van szükség.

A geometriai korrekciók az esetek jelentős részében globálisan (vagyis a teljes képre vonatkozóan) nemlineárisak; emiatt, valamint a nagy adatmennyiség miatt is, **jelentős a számításigényük**. A megvalósításnak a pontossági követelmények teljesítése mellett az elfogadható végrehajtási sebesség is alapfeltétele. Szerencsére bizonyos összefüggések kihasználásával (pl. invertálható korrekciós összefüggések alkalmazása esetén a szomszédos pontok szomszédosak maradnak; lineáris korrekció esetén egyenesek képe egyenes marad, stb.) igen hatékony algoritmusokat lehet kidolgozni.

Az alábbiakban az elméleti alapok és a főbb fogalmak ismertetése mellett röviden kitérünk a megvalósítás algoritmusait érintő kérdésekre is. Vizsgálatainkban a gyakorlatban legnagyobb jelentőségű problémakörre, a kétdimenziós képek megfordítható (*invertálható*) geometriai korrekcióira szorítkozunk.

## 5.2. A geometriai korrekció modellje

A geometriai korrekció modellje a következő ábrával szemléltethető:



12. ábra

Az első lépés a digitalizálás bemenő adatait jelentő  $f(x,y)$  kétváltozós, folytonos képfüggvény minél pontosabb helyreállítása (*rekonstrukciója*). Ezt a függvényt a fenti ábrán  $f'(x,y)$  -nal jelöltük.

Az eredeti **mintavételezés** (*sampling*) inkorrektisége, a bemenő kép **gyengén sávkorlátos** volta, valamint a **kvantálás** (*quantisation*) következtében a gyakorlatban az eredeti képfüggvény és a helyreállított képfüggvény **különböznek egymástól**.

A képfüggvény helyreállításával a geometriai korrekciót visszavezettük a folytonos kép **újrarmintavételezésére**, és **újrakvantálására**.

A korrekció geometriai jellegét az eredeti képalkotásnál alkalmazott **mintavételezés**, valamint az **újrarmintavételezés különbözősége** adja. Az újrarmintavételezés rendjét a **korrekciós összefüggés** határozza meg.

A modell egyes lépései - ugyanúgy, mint a digitális képalkotás esetében - a gyakorlatban nem mindig különülnek el egymástól. Például: a modellben szereplő visszaállítás csak az újrarmintavételezési pontokban valósul meg, ez utóbbi pedig együtt jár az újrakvantálással, s közvetlenül az új világossággódot adja.

A modellből is jól látható, hogy a geometriai korrekció két lényegi részfeladatra bontható:

- koordinátatranszformáció (az újrarmintavételezés rendjének meghatározása); illetve
- visszaállítás és újrarmintavételezés (a világossággódot meghatározása a megadott pontokban).

## 5.3. Koordinátatranszformációk

Tegyük fel, hogy a digitális képalkotásnál alkalmazott mintavételezés a valós, háromdimenziós tér 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 legyen, akkor szükség van az

$$(x, y) \rightarrow (x', y')$$

ún. **koordinátatranszformációra**. A geometriai korrekció végrehajtásakor így határozhatjuk meg, hogy a  $P'$  pontbeli képpontérték előállításához a digitális kép mely pontjának környezetében kell az eredeti(t közelítő) analóg képet helyreállítani, majd újradigitalizálni.

Tekintsük először a **lineáris koordinátatranszformációkat**. Mint ismeretes, adott bázis (**lineárisan független vektorrendszer**) esetén az  $n$ -dimenziós vektorok lineáris transzformációit  $n \times n$ -es mátrixokkal írhatjuk le. A képfeldolgozásban fontos kétdimenziós esetben a lineáris transzformációkat  $2 \times 2$ -es mátrixok alkalmazásával lehet megvalósítani.

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

$$\begin{bmatrix} x & y \end{bmatrix} \underline{\underline{A}} = \begin{bmatrix} x' & y' \end{bmatrix} \quad \text{ahol} \quad \underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

alakban írható fel, ami a mátrixszorzás szabályai szerint megfelel az alábbi

$$x' = a_{11}x + a_{12}y$$

$$y' = a_{21}x + a_{22}y$$

összefüggéseknek.

### Elemi koordinátatranszformációk

Érdemes röviden összefoglalni az ún. **elemi koordinátatranszformációkat**, melyek a lineáris transzformációt leíró  $\underline{\underline{A}}$  mátrix speciális esetei.

1. Ha  $a_{12} = a_{21} = 0$ , akkor **léptékváltás** (*scaling*)ról, azaz **skálázásról** beszélünk, mégpedig:

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

$0 < |a_{11}| < 1$ , illetve  $0 < |a_{22}| < 1$ , esetén pedig **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  $a_{11} = 0$ , illetve  $a_{22} = 0$ , **vetítésről** (*projekcióról*) beszélünk, amely a kép elfajulását eredményezi.

A két koordináta transzformációja egymástól független.

- Ha  $a_{11} = a_{22} = 1$ , de  $a_{12}$ , illetve  $a_{21}$  nem zérus, akkor a transzformációt **nyírás**-nak nevezzük.
- Az alábbi mátrix  $\Phi$  szögű elforgatást eredményez az origó körül.

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

Megjegyezzük a következőket:

- Az általános,  $2 \times 2$ -es mátrixszal leírt síkbeli lineáris transzformáció egyedül az origót hagyja változatlanul.
- Síkbeli eltolás  $2 \times 2$ -es mátrixszal nem valósítható meg.
- A lineáris transzformációk egymásutánját az egyes transzformációt megadó **mátrixok szorzataként** fejezhetjük ki. Mivel a mátrixszorzás nem kommutatív művelet, a tényezők **sorrendje** lényeges.

Ennek szemléltetésére lássuk a következő példát:

- A  $P(x,y)$  pont  $90^\circ$ -os origó körüli elforgatásának és  $y$  tengelyre való tükrözésének eredménye:

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

- Ugyanennek a pontnak előbb az  $y$  tengelyre való tükrözése, majd az origó körüli  $90^\circ$ -os elforgatása után láthatóan **más eredményre jutunk**:

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

Mint azt láthattuk, az eltolást - ami pedig igazán „elemi” transzformáció - nem lehet mátrixművelettel leírni. Ugyanakkor a transzformációk minél egységesebb leírása vitathatatlan előnyökkel jár. A problémát a **homogén koordinátás** formalizmus bevezetésével oldhatjuk fel, mivel ez a további transzformációkra is kiterjeszti az egységes tárgyalásmódot. Rendeljük a  $P(x, y)$  tetszőleges síkbeli ponthoz a  $(h \cdot x, h \cdot y, h)$  számhármast. Amennyiben  $h$  nem zérus, a számhármast a  $P(x, y)$  síkbeli pont **homogén koordinátáinak** nevezzük. Minden pontnak létezik ún. normalizált homogén koordinátás megfelelője, ahol  $h=1$ .

A homogén koordináták transzformációi is leírhatók mátrixszal, de ezzel a továbbiakban nem foglalkozunk.



## 5.4. Gyakorlati megvalósítás

A geometriai korrekció gyakorlati megvalósítása számos problémát vet fel, melyek a végrehajtás hatékonyságával (**sebességével**), s ugyanakkor a korrekció pontosságával kapcsolatosak. A teljességre való törekvés nélkül felsorolunk néhány olyan megoldást, melyek a színvonalas geometriai korrekciós programokban szerepet kapnak.

### 1. A korrekciós összefüggés definiálása

A korrekció definiálásának két alapvetően különböző útja:

- **függvénykapcsolat** megadása a bemeneti és kimeneti kép képpontjainak koordinátái között, illetve
- **azonosítási pontok** (azaz mindkét képen ugyanazt jelentő pontok) megadása. Ekkor az esetleges fizikai jelentést is magába foglaló függvénykapcsolat megtalálása interpolációs - approximációs feladat.

### 2. A korrekciós összefüggés értelmezése

Belátható, hogy a korrekciós összefüggést **inverz** értelemben célszerű megállapítani, azaz úgy, hogy az összefüggés a kimeneti képpontokhoz tartozó bemeneti pozíciókat szolgáltatassa. Előnyös például, ha nincs szükség utólagos interpolációs képjavításra.

### 3. A korrekció szervezése

A szokásos képtárolási formához igazodva, a kimenőkép koordinátarendszerében **sorfolytonos** adatkezelésre van szükség. Azt a tartományt, amelyre a korrekciós összefüggést alkalmazni kell, célszerű előzetes **maszk**készítéssel behatárolni, így a nyereség a kieső területekkel arányos, míg a maszkkészítés műveletigénye csak a kép határolóvonalainak hosszától függ.

### 4. Felületelemes közelítés

Az általánosság megsértése nélkül feltételezhető, hogy kis tartományban (azaz lokálisan) a geometriai korrekciós összefüggések akkor is lineárisak, ha globálisan (a teljes képre vonatkozóan) nem azok. Mivel a nemlineáris korrekció számításigénye többnyire nagyobb, a korrekciót a lehetőséghez mérten **célszerű visszavezetni** lineáris összefüggések alkalmazására.

Ez a visszavezetés célszerűen úgy történhet, hogy a globálisan nemlineáris transzformációt a kép **diszjunkt** részletekre (ún. **felületelemekre**) bontásával „rakjuk össze” úgy, hogy egy-egy ilyen részlet transzformációja a lehető legegyszerűbb legyen.

A gyakorlati megvalósítás számára igen előnyös megoldás, ha

- a kimeneti képet szabályos rácsozatban elhelyezkedő téglalapokkal fedjük le (minél „simább” a korrekciós összefüggés, annál nagyobbak lehetnek a téglalapok):



A feladat tehát: a  $P'$  pozícióhoz kell a bemenőkép (közeli) képpontjainak függvényében képpontértéket rendelni.

A szakirodalomban a különböző **újrámintavételezési** eljárások közül leggyakrabban a **legközelebbi szomszéd módszer** alkalmazására találhatunk példát. Ez az előbbi ábra jelölése szerint az **A** pozíció képpontértékét adja. Igényes korrekciókhoz nem használják, mivel pl. zavaró mintázatok (**Moiré minták**) megjelenéséhez vezet.

A gyakorlatban az ún. **bilineáris interpoláció** alkalmazása a legelterjedtebb. Ennek oka: egyszerű számíthatósága még elfogadható minőségű eredménnyel párosul. A 14. ábra jelöléseivel az  $f(P)$ -t a  $P'$  paramétereivel a

$$f(P) = \beta [af(C) + (1 - a)f(D)] + (1 - \beta) [af(B) + (1 - a)f(A)]$$

képlettel számítjuk ki. A számításigény jelentősen csökkenthető:

- az  $a, \beta$  valós számoktól függő tagok elkülönítésével, s a velük való szorzás szükségességének homogenitástól függő elbírálásával,
- a szorzás elvégzése helyett a szorzatok közvetlen kiolvasásával az összes lehetséges szorzattal előre feltöltött táblázatból.

A **köbös** vagy „**bicubic**” interpolációs módszer alkalmazása egyike azon módszereknek, amelyek nem csupán a közvetlen szomszédos képpontokat használják fel a képtartalom rekonstrukciójához, hanem tizenkét „szomszédos” pontot. Az eljárás ugyan számításigényes feladat, azonban a korszerű képfeldolgozó programok (Photoshop, PhotoPaint, Paint Shop Pro, stb.) mindegyike tartalmazza, mivel az egyik legjobb képminőség így érhető el.

## Gyakorlatok a Geometriai korrekcióhoz

### 17. Gyakorlat:

Geometriai korrekció\_Indít\_Szín\_fekete/fehér\_Képforrás\_Tesztkép\_Rács\_Transzformáció

*A legördülő menüben felkínált lehetőségeket próbáljuk végig különböző paraméterekkel ! Célszerű az output képet transzformáció előtt törölni, mivel az eredmény nem minden esetben tölti ki teljesen az output képmező, és az egymásra írt eredmények zavaróan hatnak.*

*Ismételjük meg a gyakorlatot a Kép2 és Kép1 képekre! Ez utóbbinál a Színt állítsuk RGB-re!*

**18. Gyakorlat:**

Újramintavételezés\_Indít\_Szín\_fekete/fehér\_Képforrás\_Módszer

*Jelöljük ki egy negyed ablaknyi részletet az input képen, és próbáljuk ki a felkínált három módszert! Most jelöljük ki különböző méretű ablakokat az output képmezőben és újra próbáljuk ki az egyes módszereket! Ismételjük meg a gyakorlatot a Kép9-cel! Ekkor a Színt állítsuk RGB-re!*

**Ellenőrző kérdések**

1. Mi a geometriai korrekció célja?
2. Melyek az ún. azonosítási pontok?
3. Mi a különbség a képalkotáskor történt mintavételezés valamint az ún. újramintavételezés között?
4. Mit jelent a felületelemes közelítés?
5. Miért előnyösebb az inverz értelmű korrekciós összefüggés?
6. Hasonlítsa össze képminőség és számításigény szempontjából a legközelebbi szomszéd módszert, a bilineáris és a köbös interpoláción alapuló újramintavételezést!
7. Szükségszerű-e, hogy a geometriai korrekció együtt járjon a képi információ torzulásával?
8. Soroljon fel olyan feladatokat, ahol szükség van geometriai korrekcióra!
9. Mit tud a bilineáris interpolációról?
10. Mi a köbös interpoláció hátránya?

## 6. SZEGMENTÁLÁS

### 6.1. Bevezetés

A szegmentálás a kép céltudatos részekre bontását jelenti: ezáltal érhető el az, hogy a különböző jelentésű képrészeket egymástól függetlenül tudjuk kezelni. A szegmentálás célja többnyire az objektumok elkülönítése a számunkra érdektelen képrészekről (háttértől). Természetesen cél lehet több objektum egymástól való elkülönítése, vagy például adott objektum részleteinek elkülönítése is. A felsorolt példákban lényegében azonosak:

- azon jellemző tulajdonságok (sajátságok) megfogalmazása, melyek a kép viszonylag kis részletének elemzésével kinyerhetők, s amelyek eléggé jellemző (szignifikáns) adatok ahhoz, hogy az elkülönítés alapját képezzék;
- a kép összes kis részletére (szokásos esetben: **képpontjára**) vonatkozóan ezen sajátságok kinyerése;
- az e sajátságok alapján **hasonló** szomszédos képrészek összevonása, egy folt-hoz sorolása;
- e foltok kezeléséhez a megfelelő segédadatok előállítás.

A szegmentálás eredményeként tehát olyan segédadatok állnak elő, amelyek megmondják, hogy a képtartalmat hogyan kell értelmezni.

A szegmentálás leggyakrabban csak lokális értelmezésű sajátságok kinyerésén alapul, így az eredmény is csak alaptulajdonságokra (fényességre, **texturális** jellemzőkre) vonatkozik. Vannak esetek, amikor ez megfelel (pl.: sejt számlálás), általában azonban a tényleges cél eléréséhez ezután még további elemzésekre van szükség. Ebben az esetben tehát a szegmentálás csak előfeldolgozásnak minősül.

A szegmentálás bizonyos értelemben osztályozásnak is felfogható, hiszen hasonlósági mértéken alapuló hovatartozást vizsgál. A leglényegesebb különbség az **OSZTÁLYOZÁS** c. fejezetben leírtakhoz képest: az egy osztályba sorolásnak egyéb feltételek teljesülése mellett a geometriai értelemben vett **szomszédosság** is feltétele.

(Ez azt eredményezi, hogy azonos sajátságú, de egymással nem összefüggő képrészek különböző objektumokhoz lesznek besorolva. Ezen utólagos feldolgozással lehet segíteni, így a szegmentálás ebben az értelemben az osztályozás előfeldolgozási lépésének is tekinthető.)

## 6.2. A szegmentálás modellje

A szegmentálás folyamatát három részre lehet bontani:

- I. **Definíció.** Az összes olyan paramétert előzetesen meg kell határozni, amelynek a szegmentálás sajáttságkinyerési - döntési tevékenységénél szerepet játszik. Ezek a következők:
  - A. **Sajáttságvektorok.** Rendszerint egy-két, de legfeljebb három-négy sajáttság használatával definiáljuk azt a vektort, amely csak a szegmentálási feladat szempontjából **lényeges tulajdonságokat** tartalmazza. (Ennél több sajáttság általában nehezen tartható kézben, instabil viselkedést eredményez.) A legáltalánosabban használt sajáttság a világosság - illetve színkód, de használható a **lokális hisztogram**, vagy pl. a **gradiens** is. Fontos szempont a könnyű (hatékony) számíthatóság, s - a felesleges számítások elkerülése érdekében - a sajáttságvektor elemeinek minél nagyobb függetlensége is.
  - B. **Távolságfüggvény.** A szegmentálás a sajáttságvektor értéke alapján hozott **döntés**en alapul. (Lényegében azt kell megállapítani, hogy az adott sajáttságvektor elegendően hasonlít-e a szomszédos sajáttságvektorokhoz, s így közös folthoz kell-e besorolni őket.) A döntés alapjául (azaz a hasonlóság eldöntésére) a **szegmentálási távolság** (röviden: *s-távolság*) szolgál. Matematikailag ez a sajáttságvektorok távolsága a sajátstágtérben, melyet valamilyen előre megadott **távolságfüggvény**-nyel (*metrikával*) mérünk. A távolságfüggvény megválasztása erősen befolyásolja a szegmentálás „jóságát”.
  - C. **Döntésfüggvény(ek).** A szegmentálás alapja a távolságfüggvényen alapuló döntés: az egymáshoz elegendően „hasonló” (azaz: *közeli*), s ugyanakkor a képen is közvetlen szomszédságban levő képrészek lesznek egy és ugyanazon foltra sorolva. A hasonlóság megfogalmazása az ún. **döntésfüggvény** segítségével történik. Ez legtöbbször (előre meghatározott, vagy pl. a képtartalomtól függően helyről-helyre változó) küszöbértékekkel való összehasonlítást ír elő.
  - D. **Eredményjellemzők.** A szegmentálás eredménye - amit tehát speciális feltételeken alapuló osztályozásnak is tekinthetünk - olyan adatstruktúra, mely képrészeket **elkülönít**, más képrészeket pedig **összekapcsol**. Előre meg kell határozni, hogy
    - milyen adatszerkezet írja le az egybetartozást, (Néhány példa: maszkkép, melyen egy-egy objektumot különböző értékű kódok azonosítanak; objektumok határolóvonalát leíró görbe, pl. **iránykódok** sorozatával megadva; szimbolikus leírás, pl.: a szegmentált köralakú objektumok középpontjainak és sugarainak adatai.),

- milyen jellemzők (adatok) tartozzanak a szegmentálás eredményéhez (a sajáttságvektorok adatai közül a továbbiak szempontjából melyeknek van jelentősége),
- milyen jellemző *jelölje* a szegmentálási eredményeket (pl.: objektumokhoz rendelt színkódok, esetleg minták megadása).

E. **Tanítók.** A képelemek „egymáshoz sorolásának” feltételeként eddig csak a „hasonlóságot” és a geometriai szomszédosságot említettük. Ebben az esetben az egyes objektumok képelemeinek sajáttságvektoraira nézve csak annyi az előzetes kikötés, hogy azok megfelelő mértékben különüljenek el a szomszédos objektumok hasonló jellemzőitől. Az **OSZTÁLYOZÁS** c. fejezetben ezt a megközelítést **tanító nélküli osztályozásnak** nevezzük.

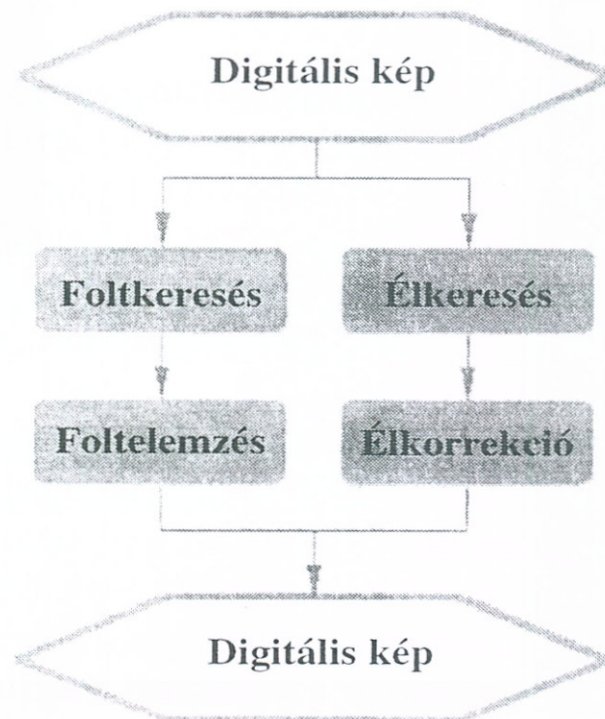
Gyakori azonban az az eset is, amikor elő akarjuk írni, hogy az objektumok milyen tulajdonságú képelemek együtteséből álljanak. Ennek az előírásnak többféle módja lehet; valamennyinek az a lényege, hogy definiáljuk az ún. **minta-sajáttságvektorokat**, melyek a kívánt osztályokat jellemzik. Ebben az esetben minden osztályhoz **legalább** egy mintát, ún. **tanítót** kell rendelni. (Ezt nevezik **tanítóval történő osztályozásnak**.)

II. **Osztályozás.** Ebben a lépésben

- A. kiszámítjuk a képrészekhez (leggyakrabban: magukhoz a képpontokhoz) tartozó sajáttságvektorokat; valamint
- B. a döntésfüggvény alapján elvégezzük a besorolást, két képrészletet akkor sorolunk ugyanahhoz az objektumhoz, ha
  - egymástól elég kicsi az **s-távolságuk**, illetve
  - ha ugyanazon osztály valamelyik tanítójához vannak a legközelebb.

III. **Összefűzés.** Megkeressük az egyes osztályokba sorolt képrészek (képpontok) **összefüggő halmazait**, azaz a foltokat (*objektumokat*) alkotó tartományokat. Az eredményt az előírt módon kódoljuk.

A szegmentálással **foltokhoz**, illetve **élekhez** jutunk, attól függően, hogy a figyelembe vett sajátságok **hasonlósági**, illetve **különbözőségi** jellemzőket mérnek-e. Ideális esetben ezek a kép objektumainak felületeit illetve határoló vonalait jelölik ki; a gyakorlatban azonban a képek ennél sokkal bonyolultabbak, s a szegmentálás eredményének utólagos javítását teszik szükségessé. A szegmentálás egyszerűsített modelljét mutatja a 15. ábra.



15. ábra

### 6.3. Foltkeresés

A szegmentálás egyik alapvető módszere a **foltkeresés**. Az eddigiek értelmében ekkor a geometriai értelemben **szomszédos** és sajátágvektorai szerint **hasonló** képrészekből álló tartományokat keressük meg. Nyilvánvaló, hogy a feladat lényege: a vizsgált sajátágokra nézve **homogén** tartományok megkeresése.

A szegmentálás „jóságának” ebben az esetben az az egyik legfőbb ismérve, hogy milyen mértékben teljesül a homogenitási feltétel. További kíváncsi, hogy a szomszédos tartományok jelentősen **különbözzenek** homogén sajátágaik tekintetében. Egyes esetekben kíváncsi még a foltok határolóvonalainak simasága is.

A szegmentálás a gyakorlatban - a képek zajossága, összetettsége következtében - olyan eredményt szokott szolgáltatni, melyet utólagosan javítani kell. Így pl.: szükség lehet a homogén foltok belsejében levő lyukak eltüntetésére; a határolóvonal zögös kanyarulatainak simítására; de akár foltok egyesítésére, eltüntetésére, illetve részekre vágására is.



## Globális foltkereső eljárások

A globális foltkereső eljárások általában a kép hisztogramjának felhasználásán alapulnak. A vizsgált sajtáság (többsávós képek esetén: sajtáságvektor) **a világosságkódok relatív gyakorisága**. Akkor kapunk jó eredményt, ha kevés, diszjunkt objektum található, közel homogén háttér előtt. Az ilyen kép hisztogramján jellegzetes **csúcsok** (*módusok*) találhatóak, melyeket világosan látható **völgyek** választanak el egymástól.

A hisztogram módusai a különböző típusú objektumoknak, illetve a háttérnek felelnek meg.

Részletesebb elemzés nélkül megemlítjük, hogy a hisztogram alapján történő szegmentáláshoz a **hisztogram partíciónálása** adhatja meg a vágási határokat.

Ennek lényege nem egyszerűen a völgyek megkeresése, hiszen kisebb völgyek véletlenszerűen, fizikai jelentés nélkül is kialakulhatnak. Sokkal inkább arról van szó, hogy meg kell keresni, az adott hisztogram milyen egyedi (egy-egy objektumot jellemző) hisztogramok eredőjeként (**szuperpozíciójaként**) alakulhatott ki.

Általában kézenfekvő az a feltételezés, hogy a homogén fényességű objektumok hisztogramjában a világosságkódok eloszlása közelítőleg **normális**, azaz haranggörbe alakú. Ekkor a partíciónálás e haranggörbe alakú összetevők metszéspontjának meghatározását jelenti.

## Lokális foltkereső eljárások

A **globális** eljárások közös hibája, hogy érzéketlenek a kép helyről-helyre változó jellegére, ami például a megvilágítási egyenetlenségek következtében szinte minden esetben fellép. Ennek következményeként pl. a homogén foltok „kilyukadnak”, részekre szakadnak; a határolóvonalak pedig „zajosak” és kiszélesednek.

A **lokális** szegmentálási eljárások bizonyos értelemben követik a kép változásait, s így a felsorolt hiányosságok kevésbé fordulnak elő.

A lokális módszerek kis - általában  $3 \times 3$ ,  $5 \times 5$  képpontos - környezet figyelembevételén alapulnak. (Az ennél jelentősen nagyobb környezet figyelembevétele esetén **regionális** eljárásról beszélünk.)

A következőkben röviden ismertetünk néhány alapvető lokális szegmentálási eljárást.

## *Összefűzési módszerek*

- Az **egyszerű összefűzési módszerek** a szomszédos képpontokból indulnak ki. Két szomszédos képpontot akkor kapcsolnak össze, ha **s-távolságuk** elegendően kicsi. A módszer előnye az egyszerűsége, s ennek köszönhetően a nagy sebesség; hátránya, hogy rendkívül érzékeny a képhibákra. Elég például egy „téves kapocs” ahhoz, hogy két különböző foltot egyesítsen. Sok esetben az is hátrányos, hogy a különböző objektumok közti **lassú** átmenetet nem veszi észre.
- A **kombinált összefűzési módszerek** a zajérzékenység csökkentése érdekében környezetelemzést is végeznek. Csak akkor kapcsolnak össze két szomszédos pontot, ha **kis környezetük s-távolsága** is elegendően kicsi.

## *Területnövesztés*

Azokat a módszereket, melyek a még nem osztályozott pontokat a már kialakított, de még be nem fejezett tartományokhoz fűznek hozzá, **területnövesztés**es módszereknek nevezzük. Előnyük, hogy lassú átmenetek esetén is működőképesek, és viszonylag érzéketlenek a képhibákra.

## *Szétvágás és egyesítés*

A módszer alapelve a következő:

- vágjuk részekre az aktuális képszegmenst, ha nem eléggé homogén;
- egyesítsük azokat a szomszédos részeket, amelyek egymáshoz eléggé hasonlóak.

A kiinduló szegmens rendszerint a teljes kép, s a vágás/egyesítés számítástechnikai megfontolásokból minden közbenső állapotban négy egyenlő területű részre vágáson alapul. A módszer hátránya, hogy lépcsős határolóvonalakat eredményez, és nagy az előállított fa leírásához szükséges tárkapacitásigény.

## *Hibrid módszerek*

A hibrid módszerekben **globális** és **lokális** eljárásokat egyaránt alkalmaznak a szegmentálás megvalósítására. Ezenkívül azt a lehetőséget is kihasználják, hogy a szegmentálás eredménye **utófeldolgozóssal** (például: **mediánszűréssel**, **többségi elvű szűréssel**) javítható.

## 6.4. Foltelemzés

A geometriai és/vagy topológiai jellemzők szerinti osztályozás a foltelemzés egyik legfontosabb részterülete. Mindkettőnek az a lényege, hogy a szegmentálás eredményeként kialakult foltokat új szempontok szerint (azaz új sajáttságvektor alapján) kell csoportosítani.

Számos lehetőség közül említést érdemel a **mintaillesztés** (*template matching*), mely alakzatok közti **hasonlóság/eltérés** elemzése alapján osztályoz. Talán a legismertebb alkalmazási terület az **optikai karakterfelismerés** (*OCR*), bár ez utóbbi területen a sebesség és a felismerési pontosság növelése érdekében ma már sokkalta összetettebb módszereket alkalmaznak.

A foltelemzés másik fontos területe a **foltszámlálás**. Ekkor nincs szükség újabb sajátságok kiszámítására. Különböző gyors algoritmusok léteznek sorfolytonos, illetve foltkövető szervezésű számlálásra.

## 6.5. Élkeresés

Ha a szegmentálás a különbségi jellemzők alapján történik, **élkeresésről** beszélünk. Bár a megtalált élek általában a foltok (objektumok) határolóvonalai is egyben, az élkeresés a szegmentálás szempontjából jelentés nélküli vonalakat is eredményez. Az összetett élkeresésnek ezért a szükségtelen élek kiszűrése is feladata.

Az olyan folytonos és zárt élvonalakat, melyek egy foltot, illetve lyukat teljesen körülhatárolnak, **kontúrvonalaknak** nevezzük. Az élvonalak élpontokból állnak, melyek egymással 8-as szomszédságban vannak.

Az élpontokat ott keressük, ahol a sajáttságvektorok hirtelen, ugrásszerű változást mutatnak. Ez egyben azt is jelenti, hogy a szomszédos képpontok sajáttságvektorának **s-távolsága** elegendően **nagy**.

A gyakorlatban az élkereső eljárások többnyire lokálisak, s inkább a feleslegesen talált élek utólagos eltüntetésével érik el ugyanazt a hatást, amit a globális élkereső eljárás szolgáltatna. Így az élkeresés általában két fázisból tevődik össze: **élkitűzés**-ből (ennek eredményeként előáll az „élgyanús” pontok halmaza), valamint **élkorrekció**-ból (amikor az élgyanús pontokból csak a valóságos élpontok maradnak meg).

### Élkitűzés

A pontokat két osztályba soroljuk: az élpontok lesznek az „értékes” pontok (amelyeket még egyszer végig kell nézni az élkorrekció érdekében), a többiek pedig a továbbiakban felesleges háttérpontok.

Például a **gradiens módszer** (ld. 4.6. részben) egy pontot akkor minősít „élgyanús” pontnak, ha a környezetében számolt gradiens egy adott küszöbnél nagyobb. A gradiens számítása természetesen valamilyen közelítéssel történik. Szokásos pl. a Sobel-, illetve a Laplace-operátor alkalmazása. Ezeket az eljárásokat már ismertettük a Képjavítás fejezetben.

## Iránykereső módszerek

Az iránykereső módszerek különböző irányokban vizsgálják az átmeneti görbe meredekségét, s 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, melyekkel a vizsgált pont körüli 8 első differenciahányadost számítjuk ki, majd adott küszöbértékhez hasonlítjuk őket.

## Vonalkereső módszerek

Az eddigi módszerek az éleket pontokból rakták össze. A **vonalkereső módszerek** az éleket jól közelítő egyenes szakaszok megkeresésére irányulnak, s ezekből fűzik össze a végleges élképet. E módszerek átmenetet képeznek a lokális és regionális módszerek között, mivel az aktuális képpont vizsgált környezete általában elég nagy (7\*7 ... 11\*11 képpont).

### A Hückel-féle vonalkereső eljárás lényege:

- a képsíkot (közel) kör alakú (átfedő) ablakokkal fedi le;
- ötparaméteres ideális **élfüggvényből** kiindulva egy hibafüggvény minimumát keresi úgy, hogy ha az illeszkedés elég jó és az éldarab nem túl hosszú, az eredményt valódi él közelítéseként fogadja el.

Sajnos ez az eljárás meglehetősen számításigényes, és bizonyos esetekben hibás eredményre is vezethet.

Sokkal használhatóbb a (magyar kutatók által kidolgozott) **Mérő-Vassy-operátor**, amely azon a felismerésen alapul, hogy ha egy ablakban létezik határvonal (azaz az ablakban eltérő világosságkódú foltok találkoznak), akkor ennek iránya és helyzete **egyértelműen meghatározható** a világosságkódok ablakon belüli **eloszlásából**.

A Mérő-Vassy módszer előnye a hatékony megvalósítás lehetősége, valamint a valódi él esetében megfelelő pontossága. Hátránya, hogy **mindig** talál élt az ablakban; akkor is, ha egyáltalán nincs él, s akkor is, ha az élvonal törik, vagy például több él találkozik. Az első esetet egyszerű kijavítani döntési küszöb bevezetésével; a második és harmadik esetet **élkorrekcióval** lehet kijavítani.

## 6.6. Élkorrekció

A tapasztalat azt mutatja, hogy pusztán a képsajátságok elemzésével nem lehet kielégítő szegmentálást eredményező élképhez jutni. Szükség van az élpontok, illetve éldarabok egymáshoz viszonyított helyzetének vizsgálatára, az élkijelölés eredményének korrekciójára.

Az élkorrekció célja a „hamis” élek eltüntetése, a szakadások összekötése, valamint a túlságosan zezugos kontúrvonalak kisimítása.

Számos módszer létezik e részfeladatok megoldására; röviden felsorolunk néhányat.

### Analitikus módszerek

Az analitikus módszerek alapelve az, hogy az éleket a megtalált „élgyanús” pontokra illeszkedő digitális görbék segítségével alakítják ki.

A gyakorlatban - a számításigény csökkentése érdekében - a közelítés **szakaszosan lineáris**, vagyis görbék helyett tört vonalakkal történik. Mivel ezek a módszerek nem alkalmasak a hamis élpontok kiszűrésére, ezeket előzőleg el kell tüntetni.

Egy másik lehetőség a szakaszosan folytonos függvényekkel (*spline*-okkal) történő interpoláció, mely az egyes pontokban elkövetett hibákra viszonylag érzéketlen lesz. Ekkor az élgyanús pontok a görbe tartópontjai.

### Vonalvékonyítás

Az élkorrekció viszonylag egyszerű és elterjedten alkalmazott módja a következő:

- a döntési küszöböt annyira alacsonyra választják, hogy az összefüggő élek biztosan ne szakadjanak el, de ugyanakkor a szomszédos élek még elkülönüljenek egymástól.
- az élképre vonalvékonyító eljárást alkalmaznak, aminek eredményeként a széles „élszalagok” egyetlen pont vastagságú élekké zsugorodnak össze, a folytonosság megtartása mellett.

A vonalvékonyítás után megmaradt pontok alkotják az alakzat ún. **csontvázát** (*skeletonját*).

## Gyakorlatok a szegmentáláshoz

### 19. Gyakorlat:

Átszínezés\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép2\_  
 Átszínezőtábla\_Identikus\_Végrehajt\_BW\_Átszínezőtábla\_Szintrevágás\_  
 Végrehajt\_BW

*Mozgassuk addig az átviteli függvényt, amíg a „munkadarabról” a legtisztább képet kapjuk!*

*Ismételjük meg a gyakorlatot Kép3-mal, és válasszuk külön a sötét objektumokat a háttértől!*

*Ismételjük meg a gyakorlatot Kép6-tal és próbáljuk a folyót elkülöníteni a kép többi részétől!*

### 20. Gyakorlat:

Konvolúció\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép1\_  
 Szűrő mérete\_3\*3\_Szűrési mód\_Def.szűrővel\_Szűrődef.

*A szűrőmátrix középső eleme legyen -4, a felette, alatta, tőle jobbra és balra lévő elemek értéke legyen 1, a többi 0 (Laplace operátor), majd:*

Végrehajt\_Képforrás\_Utolsó eredménykép\_Kilép\_

*Lényegében a 14. Gyakorlat egy részét ismételtük meg, és kijelöltük az élgyanús pontokat. Ezután próbáljuk kiválasztani csak az élpontokat kétszintre vágással:*

Átszínezés\_Indít\_Átszínezőtábla\_Identikus\_Végrehajt\_BW\_Átszínezőtábla\_  
 Lyuk\_Végrehajt\_BW

*Változtassuk az átviteli függvényen a lyuk szélességét és figyeljük meg mi történik! Ismételjük meg a gyakorlatot a Kép6-tal!*

### 21. Gyakorlat:

Konvolúció\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép6\_  
 Szűrő mérete\_3\*3\_Szűrési mód\_Ált. szűrők\_Irányszűrők\_

*Sorban válasszuk ki a különböző irányokra érzékeny szűrőket, és figyeljük meg a kanyargó folyó mentén hatásukat! Próbáljuk az Átszínezés funkcióval az élpontokat kiválasztani!*

Végrehajt\_ Képforrás\_ Utolsó eredménykép\_ Átszínezés\_ Indít\_ Átszínezőtábla\_Lyuk\_ Végrehajt\_BW

*Változtassuk az lyuk átviteli függvény szélességét és pozícióját! Ismételjük meg a gyakorlatot 5\*5 szűrővel és más képekkel is!*

## **Ellenőrző kérdések**

1. Mi a szegmentálás célja?
2. Mi a sajátság, illetve a sajátságvektor? Mondjon példát általánosan használt sajátságokra!
3. Mi a távolságfüggvény?
4. Mi a döntésfüggvény?
5. Milyen kapcsolat van a távolságfüggvény és a döntésfüggvény között?
6. Mit nevezünk tanítónak?
7. Az osztályozás során mi történik?
8. Melyek a foltkeresési eljárások főbb jellemzői?
9. Miben különböznek a lokális és a globális foltkereső eljárások?
10. Ismertesse az összefűzési módszerek alapelvét!
11. Mi a területnövesztő eljárás lényege?
12. Mi a kontúrvonal?

## 7. OSZTÁLYOZÁS

### 7.1. Bevezetés

Az osztályozás mintegy belépő a kép felismeréséhez, megértéséhez. A számítógépes képfeldolgozás e rendkívül sokrétű területe képpontokon és régiókon alapuló műveleteket, statisztikus és szintaktikus módszereket egyaránt magába foglal. A kép osztályozása - sokfélesége ellenére - lényegileg egyazon feladat megoldását jelenti:

- képpontokat, kisszámú (összetartozó) képpont együttesét, illetve szegmentált alakzatokat kell tulajdonságai (**sajátságai**) alapján **felismerni, kategorizálni**, illetve megadott lehetőségek valamelyikébe **besorolni**, s ezzel
- létrehozni a kép magasabb szintű **leírását**.

Néhány alapeset:

#### alakfelismerés (*pattern recognition*)

a vizsgált **sajátság** (feature)-ok a kép makroszerkezetét, **objektumait** jellemzik, s általában az előzetesen **szegmentált** képen alapulnak. (Megjegyzendő, hogy a talán szerencsésebb **alakzatfelismerés** elnevezés nem terjedt el; így ugyanaz az elnevezés él a **felismerések** sokrétű, köznapi értelemben véve „alakkal” semmilyen összefüggésbe sem hozható alkalmazási területein.)

#### textúraelemzés

a vizsgált **sajátságok** a kép mikroszerkezetét (**textúráját**) írják le. A felismerés érvényessége többnyire a mintázat szempontjából homogén, kisméretű tartományokra terjed ki.

#### képpontértékelemzés

a vizsgált **sajátság** (*feature*) maga a képpontérték(vektor), ami tehát lehet egy **világosságkód**, többsávos kép esetén az egyes spektrális összetevők vektora, különböző időpontban készült felvételek esetében pedig a geometriai értelemben összetartozó adatok vektora. (Az utóbbi két eset a **multispektrális**, valamint **multitemporális** osztályozás.)

### 7.2. Az osztályozás modellje

Módszereit tekintve, mind az alakfelismerésnek, mind pedig a textúraelemzésnek két, lényegesen eltérő útja van (16. ábra):

A **statisztikus (döntésméleti)** osztályozás valószínűségszámítási és matematikai statisztikai módszerekkel dolgozik. Az így előállított jellemzők - ellentétben a szintaktikus alakfelismerés eredményeivel - az objektumok, illetve textúraelemek közt fennálló síkbeli (térbeli) összefüggésekkel nem foglalkoznak.



A **szintaktikus (strukturális)** módszerek ezzel szemben az objektumok, illetve **textúraelemek** közti síkbeli (térbeli) összefüggéseken alapulnak. Módszerei közt fontos szerepet kapnak a matematikai nyelvészeti módszerek.

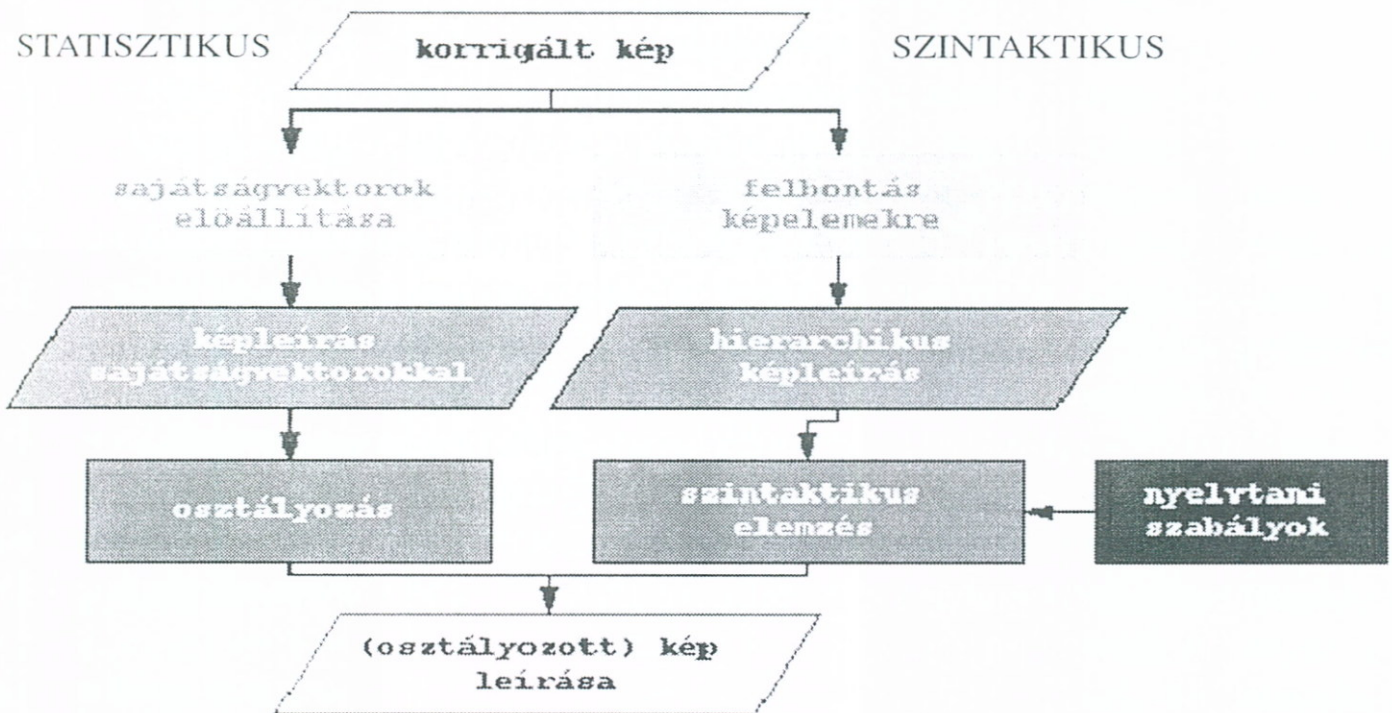
A modell két ágának fő lépései közt nehéz párhuzamot vonni (annyira eltérő műveleteket takarnak), a következő három lépés mégis mindkettő esetében érvényes:

1. A modell szerinti első lépés a jellemző adatok kinyerése valamennyi objektumra, illetve textúraelemre vonatkozóan. (Ez a statisztikus módszerek esetében valószínűségi változóként kezelhető, mérhető jellemzőkkel - **sajátságvektorral** - kapcsolatos, míg a szintaktikus módszerek esetében elemi alakzatok megkeresését jelentheti.) Ez a lépés felfogható **lényegkiemelésnek** is, hiszen a további műveletek számára csak a lényeges adatok maradnak meg.
2. A második lépés az objektumok, illetve textúrák besorolása az előre megadott, vagy menet közben kialakított osztályokba. (A három lépés közül ezt lehetne leginkább osztályozásnak nevezni.) A besorolás kritériuma a szóban forgó sajáttságvektor hasonlósága valamelyik osztályéhoz, illetve különbözősége más osztályokétól.
3. A harmadik lépés az eddigi eredmények alapján az (osztályozott) **képleírás** előállítás.

### 7.3. Statisztikus alakfelismerés

A statisztikus alakfelismerés az említettek szerint matematikai statisztikai és valószínűségszámítási módszerek alkalmazásán alapul. Három alapesetet különböztetünk meg:

- Az osztályokról elegendő adat áll rendelkezésünkre ahhoz, hogy el tudjuk végezni az optimális osztályozást. Ekkor pl. a Bayes-féle döntési módszert alkalmazhatjuk.
- Az osztályokat csak egyes mintaelemeik, a **tanítók** révén ismerjük. Ezért nem biztos, hogy minden paraméterük egyértelműen adott, sőt, a tanítók halmazából álló **tananyagban** átfedések, vagy akár ellentmondásosak is lehetnek. Ekkor pl. a **távolságmérő-módszereket** alkalmazhatjuk .
- Ha az osztályokról **semmit sem tudunk** (sőt: legtöbbször még a számukat sem ismerjük), akkor a **klaszterezést** alkalmazhatjuk. Itt bizonyos értelemben az osztályozás fordítottjáról van szó: adott elemekhez keressük a megfelelő osztályokat.



16. ábra

Megjegyzendő, hogy a **képelemek** (képpontok, objektumok, textúraelemek) a gyakorlatban sohasem függetlenek egymástól, ezért a jó eredmény eléréséhez a fenti módszerek mindegyikét célszerű a környezeti adatokat is figyelembe vevő korrekcióval kiegészíteni.

A statisztikus alakfelismerés alapját képező **sajátságvektor** azon lényeges adatokat tartalmazza, melyeknek hasonlósága, illetve különbözősége a besorolás eldöntéséhez felhasználható. Azt, hogy a saját-sajátságvektor milyen komponenseket tartalmazzon, mindig az adott feladat és a kép szerkezete alapján kell eldönteni.

Néhány gyakran használt (objektumleíró) jellemző:

- ívhossz vagy **kerület**
- **terület**
- **alaktényező** (terület/kerület)
- legnagyobb és legkisebb **átmérő**
- **elnyújtottság** (a legnagyobb és legkisebb átmérő aránya)
- **konvexitás** (az objektumot befoglaló legkisebb területű konvex poligon területének aránya az objektuméhoz);
- **topológiai jellemzők** (pl.: lyukak száma)
- **objektum képpontértékeinek statisztikája** (sávonkénti átlag, szórás)
- **objektum képpontértékeinek területi eloszlása** (például: sávonkénti súlypont)

Minél több jól kiválasztott és független sajátsgot használunk fel az osztályozáshoz, annál biztosabb (stabilabb) eredményhez jutunk. Másrészt, a sajátsgvektor elemeinek számát (azaz a **dimenziószámot**) célszerű minél alacsonyabb értéken tartani, mivel a sajátsgok számával általában erősen nő a besorolás számításigénye.

A **dimenziószámcsökkentő módszerek** sajátsgvektor elemei közti esetleges belső összefüggések minimalizálásával csökkentik a figyelembevett sajátsgok számát, s így a besoroláshoz szükséges számításigényt. (Nem szabad azonban elfelejteni, hogy többnyire e módszerek alkalmazása is jelentős számítási feladat.)

Két fő módszere van:

- **lényegkiemelés** (*feature selection*) az új sajátsgvektor a kevésbé fontos (kevesbé *szignifikáns*) komponensek elhagyásával alakul ki;
- **lényegtömörítés** (*feature extraction*) esetén az új (csökkentett dimenziószámú) vektor komponensei a régeből valamilyen transzformációval jönnek létre. (Nyilvánvaló, hogy ez egyben lényegkiemelést is jelenthet.)

Megemlítjük, hogy a legismertebb módszer, az ún. **Karhunen-Loeve módszer** optimális, mivel az eredményvektor komponensei teljesen **függetlenek** (**korrelálatlanok**) lesznek.

A **döntésfüggvény** azt határozza meg, hogy a sajátsgvektor értékétől függően mi lesz a besorolás eredménye. Alapvető követelmény a kis számításigény, valamint a besorolási biztonság (azaz: pontosság).

Természetes kíváncsi az olyan döntésfüggvény választása, melynek hibája (azaz a hibás döntések száma) **a lehető legkisebb**. Ha ez a feltétel teljesül, **Bayes-döntésről** beszélünk.

**Diszkrimináns függvények szerinti döntésről** akkor beszélünk, ha a sajátsgvektor-teret diszjunkt **régiókra** bontjuk, s a sajátsgvektor komponenseinek értéke alapján összehasonlítások döntenek el a hovatartozást.

Az ún. **távolságmérő-módszerek** alkalmazása esetén a döntések a sajátsgvektorok és a valamilyen módon előre kiválasztott mintavektorok **távolságának mérésén** alapulnak. A gyakorlatban az alábbi távolságok használatosak:

- Euklideszi távolság
- Maximum távolság
- Minimális vagy spektrális távolság
- Abszolút távolság
- Mahalanobis távolság
- Például a **legközelebbi szomszéd módszer** esetén egy objektumot akkor sorolunk az *i*-edik osztályba, ha a mintapontok közül a hozzá legközelebb eső pont az *i*-edik osztályhoz tartozik. E módszer egyszerűsége ellenére rendkívül hatékony, s hibája csak legfeljebb a kétszerese az optimális Bayes-módszerrel elérhetőnek.

- Ugyancsak távolságmérésen alapul a **doboz-módszer** ( box-módszer ). Ekkor a tanítók sávonként megadott intenzitástartományokkal jellemezhető „téglatestek”; a besorolás kritériuma a téglatestek valamelyikének belsejébe esés. Ha a dobozok közt átfedés is van, további kritériumot jelenthet pl. a centrumtól mért távolság szerinti döntés. A módszer rendkívüli előnye a gyors számíthatóság. Hátrányos viszont az, hogy nincs tekintettel a sajátságok tényleges szerkezetére, s egy azoktól idegen, téglatestalakú szerkezetet erőszakol rájuk (az előzőekben ismertetett távolságmérő módszerek mindegyike gyakorlatban is elvégezhető a TULIPP gyakorlóprogram segítségével).

## 7.4. Klaszterezés

A **klaszterezésre** (*clustering*) akkor kerül sor, ha olyan adathalmazt kell osztályoznunk, melyről semmilyen előzetes ismerettel nem rendelkezünk. A feladat végül is az adatok olyan csoportosítása, aminek eredményeként az egymáshoz hasonló adatok egyazon csoportba kerülnek, az egyes csoportok pedig elegendően különböznek egymástól.

A klaszterezés esetén magukat az osztályokat is a sajátságvektorok alapján kell meghatároznunk. Sőt, az is gyakori eset, hogy azon sajátságokat sem ismerjük előre, melyek alkalmazásával az adathalmaz elegendően szétválk (*szeparábilis*).

A feladat úgy is felfogható, hogy

- keressük azokat a sajátságokat, melyek terében az adatok „hisztogramja” jellemző sűrűsödési pontokat mutat; és
- keressük ezeket a sűrűsödési pontokat alkotó pontokat (osztályokat).

Egy adathalmaznak nyilvánvalóan többféle klaszterezése is lehetséges, hiszen az eredmény függ például attól, hogy mik a sajátságvektor elemei; mivel mérjük a sajátságvektorok hasonlóságát; milyen az adattér esetleges kiindulási partíciója. A klaszterezés eredményéről nehéz eldönteni, hogy megfelelő-e, s a nehézséget nem matematikai vagy számítástechnikai probléma okozza. Inkább az a probléma, hogy a klaszterezés eredménye mögött **nincs fizikai jelentés**, s a csoportosítást számos egyéb körülmény is befolyásolja.

A számos klaszterezési eljárás közül a következők a legfontosabbak:

### 1. *Hierarchikus módszerek*

A **hierarchikus klaszterezés** alapelve a következő: minden egyes lépésben vonjuk össze a két, egymáshoz legközelebbi klasztert mindaddig, amíg az előírt számú klaszterig nem jutunk.

Csak kis elemszám esetén használható, mivel a tárolóigény a klaszterek számával négyzetesen arányos.

## 2. Klaszterezés célfüggvény alapján

Az osztályozás **jóságát** valamilyen célfüggvénnyel is mérhetjük. Például azt mondhatjuk, hogy a klaszterezés akkor a legjobb, ha az egyes klaszterekbe sorolt elemek sajáttságvektorai és a klaszterközéppontokba mutató vektorok közötti távolságok négyzetösszege minimális. (Hasonló feltételt szabhatunk az átlagos távolságokra, vagy például a maximális távolságra.) Ekkor a klaszterezés egy szélsőértékprobléma megoldásával azonos.

## 3. Iteratív módszerek

A célfüggvény kiválasztása után az osztályozás tulajdonképpen diszkrét optimalizációs problémává válik: meg kell határozni a képelemek halmazának egy olyan felosztását (*particionálását*), amely a célfüggvény kívánt szélső értékét (maximumát vagy minimumát) szolgáltatja. Kicsi adathalmaz esetén ez egyszerűen megoldható az összes lehetséges partíció ellenőrzésével, nagyobb adathalmazok (pl. képek) esetén azonban ez az út már nem járható. A probléma iteratív módszerekkel hidalható át.

Az iteratív módszerek lényege benne van a nevükben: az eredmény több lépésben (*iterációval*), a célfüggvénnyel kapcsolatos szélsőérték egyre pontosabb és pontosabb teljesítése mellett alakul ki.

Például: Legyen a célfüggvény a klaszterközépponttól vett távolságok négyzetösszegének minimuma, és számítsuk ki minden egyes besorolás után a kialakult klaszterek középpontjait. Ekkor (a gyakorlatban) a klaszterközéppontok egyre stabilabbak lesznek, s a besorolás eredménye is egyre kevésbé változik.

Az iteratív klaszterezés módszerei közül talán az **ISODATA-eljárás** (*Interactive Self-Organizing Data Analysis Techniques*) a legismertebb.

Lényege röviden:

1. Jelöljük ki a kezdeti klaszter-centrumokat. Pl.: minden sajáttséghez számítsunk ki két értéket: az egyikhez az átlagértéket növeljük, a másikhoz pedig csökkentsük a megfelelő szórásértékkel. Képezzünk klaszter-centrumokat az így kapott értékek összes lehetséges kombinációjával. Nyilvánvaló, hogy ezek száma a sajáttságvektor elemszámának négyzete lesz.
2. A döntésfüggvény alapján végezzük el a besorolást, és határozzuk meg az egyes osztályok jellemző adatait.
3. Az egymáshoz **túl közeli** klasztereket vonjuk össze; a **túl kicsiket** szüntessük meg; a **túl lazákat** vágjuk szét. A felszámolt klaszterek elemeit minden esetben soroljuk be azonnal.
4. Számítsuk az új osztályjellemezőket.
5. Ha az osztályjellemezők egy bizonyos küszöbnél nagyobb mértékben módo-

sultak (és egy előre megadott iterációs számot még nem léptünk túl), folytassuk az eljárást a 2. ponttal; egyébként a klaszterezés befejeződött.

## 7.5. Környezet figyelembevétele

Az eddig ismerttetett (statisztikus) módszerek nem használják ki azt a többlet-információt, amit a besorolandó szomszédos adatok közti összefüggés adhat. Pedig - különösen, ha az adatok maguk a képpontok - a képeken az egyes adatok közt erős az összefüggés.

A képi jelleg figyelembevételének néhány példája:

### Előfeldolgozás

Ekkor a sajátságvektorban környezeti információt is figyelembe veszünk, s ezáltal a környezet befolyásoló hatása automatikusan érvényesül. Ilyen „globális” jellemző lehet például a környezet átlagos világosságkódja, vagy annak szórása.

### Előszegmentálás

Ekkor az osztályozandó képrészt előzetesen szegmentáljuk, s a spektrálisan homogénnek tekinthető (s feltehetően egyazon objektumot ábrázoló) foltokat egyszerre soroljuk be valamelyik osztályba.

### Utófeldolgozás

Az osztályozás után vesszük figyelembe a környezetet. Például, az ún. **többségi elvű szűrés**nél a pontok közeli környezetében megszámloljuk a különböző osztályba sorolt képpontokat, s ha ezek közül valamelyik relatív gyakorisága meghalad egy küszöbértéket, a vizsgált pontot átsoroljuk ehhez az osztályhoz.

## 7.6. Szintaktikus alakfelismerés

A statisztikus alakfelismerési módszerek egyes esetekben csak nehézkesen alkalmazhatók, nem nyújtanak megfelelő eredményt. Nehézséget okoz például az **összetett alakzatok** kezelése, mert ezek szerkezete nem írható le sajátságvektorral. Általában is elmondható, hogy a statisztikus módszerek csak a nem túlságosan bonyolult szerkezetű, azaz alacsony dimenziószámú sajátságvektorral jellemezhető objektumok esetén alkalmazhatók.

A **szintaktikus alakfelismerési** módszerek más oldalról közelítik meg a besorolási problémát. Alapelvük lényege:

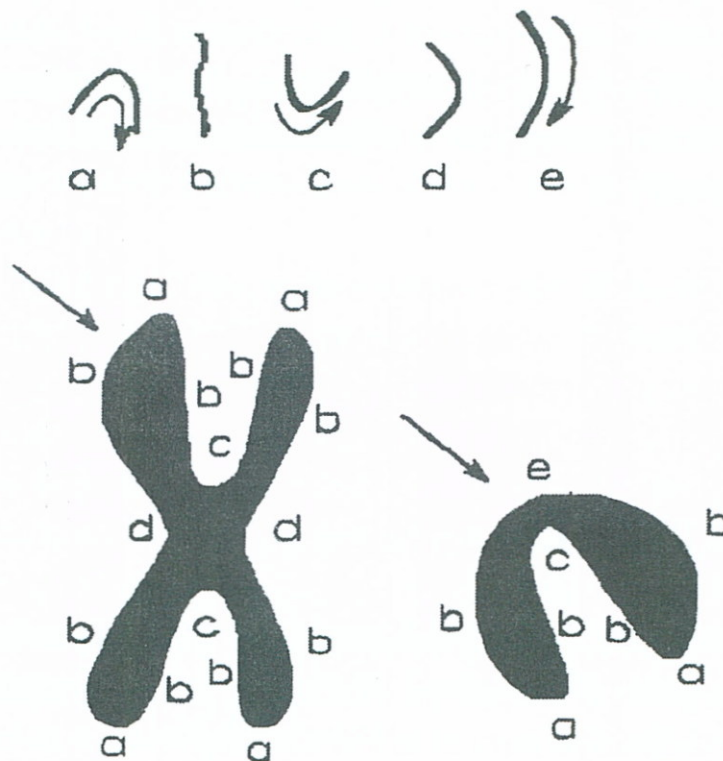
- bontsuk fel a vizsgált objektumot olyan egységekre, amelyek egyszerűbb szerkezetűek; ezeket is bontsuk tovább még egyszerűbb alkotórészekre; egészen addig, míg olyan elemi objektumokig nem jutunk, melyeket a képen nehézség nélkül lehet azonosítani. Ezeket az azonosítható objektumokat **képelemeknek** nevezzük.

- egyúttal állapítsuk meg az adott objektum esetében a képelemek közti kapcsolat rendszerét, s írjuk le ezzel az objektumot.
- az osztályozás ezek után abból áll, hogy megnézzük, az adott objektum **leírása** illeszkedik-e a megadott osztályok valamelyikének leírásához.

Az objektumok szintaktikus leírásának példaként gyakran használják a **kromoszómákat**, melyek - a 17. ábrán látható módon - az elemi alkotórészek szekvenciájaként adhatók meg. Ha ezen alakzatok felismerése a cél, akkor például:

- megkeressük az objektumokat körüljáró görbéken az **a, b, c, d** alkotóelemeket (az ábrán ezek mintái felül láthatók), majd pedig
- ezek **egymásutánisága** alapján osztályozunk. A két alsó ábra így a következő adatokkal írható le:

abcbabdbabcbabdb, és  
ebabcbab



17. ábra

## 7.7. Textúraelemzés

A textúraelemzés módszereinek áttekintését megnehezíti az a körülmény, hogy a textúrának nincs általánosan elfogadott, precíz definíciója.

A lényegét kifejezi a következő meghatározás: a **textúra** statisztikusan ismétlődő (azonos vagy hasonló) **vonalt-** illetve **területelem**ekből felépülő, szabályos vagy véletlenszerű (*nonfiguratív*) **mintá**. Lényeges, hogy a textúraelemek elegendően kicsik ahhoz az alakzathoz képest, amelyet alkotnak.

A textúra kis mintái önmagukban többnyire nem észlelhetők, csak az objektumról alkotott benyomásainkat határozzák meg alapvetően. A látványt tehát nem egyedi tulajdonságaik, hanem **struktúrájuk, együttes előfordulásuk, rendezettségük** határozza meg. A textúra elemzésének objektumai a **textúraelemek**, a sajátságvektorok pedig a mikrostruktúrát jellemzik.

**Homogénnek** nevezzük az olyan textúrákat, melyeket kis részletek statisztikai ismétlődése jellemez. Ezt nevezhetjük „önhasonlóságnak” is, ami azt fejezi ki, hogy a kép „eléggé nagy” részképei egymáshoz hasonlóak, helyzetüktől, alakjuktól és méretüktől függetlenül. Ennek ellentettje az **inhomogén textúra**.

A **textúrális sajáttságvektorelemek** például a lokális statisztikai jellemzőkből képezhetők.

Képezhetünk sajáttságvektort egyszerűen a **hisztogramból**, az **átlag, szórás, medián, maximum/minimum**, stb. értékek felhasználásával. Másféle jellemzőket adnak a képpontértékek geometriai elhelyezkedésével összefüggő adatok. Ilyenek pl. a **futási hosszak (futamhosszak)**, melyek az azonos világosságkódú, szomszédos képpontok számát adják meg valamilyen irányban.

A leggyakrabban használt „sajáttságforrás” az **együttelőfordulási mátrix** (*co-occurrence matrix*), mely azt fejezi ki, hogy egy adott **világosságkódú** képponttól adott irányban és adott távolságra milyen relatív gyakorisággal találunk **ugyanolyan világosságkódú** a képen.

A textúrák **kvantitatív** jellemzésére három tulajdonságot szokás értékelni:

- a finomság (vagy durvaság) a textúraelemek átlagos méretével függ össze;
- a szabályosság a textúraelemek síkbeli elhelyezkedésével függ össze, míg
- az anizotrópia azt vizsgálja, léteznek-e a textúraelemek elhelyezkedésében ki-tüntetett irányok.

## Gyakorlatok az osztályozáshoz

Ennél a funkciónál a program minden képet színes, azaz 3 sávós képként kezel. A Kép1, Kép9 és Kép10 képek valamint a CD - IMAGES/MUHOLD; IMAGES/LEGIFELV



könyvtárában található képek valóban 24 bites színes képek, ezekkel érdemes ezeket a gyakorlatokat elvégezni. A képek színmélységéről kapunk információt, amennyiben bármely gyakorlat során a Képforrás\_File\_Info funkciót használjuk. Az osztályozás funkció számos paraméterezési lehetőséget tartalmaz, érdemes a Funkció Help-et elolvasni. Itt most csak a legegyszerűbb paraméterezésre teszünk javaslatot.

## 22. Gyakorlat

Osztályozás\_Indít\_Képforrás\_File\_Kép9\_Módszer\_Tanítóval-box\_

*Jelöljük ki egymás után négy pontot, és ezzel négy mintaterületet az input képen. Legyen az első minta a folyó, a második a bal felső sarokban látható vöröses mezőgazdasági tábla, a harmadik a mellette lévő élénkzöld tábla, a negyedik a barnás színnű erdős terület! Ezután:*

Végrehajt

*Vizsgáljuk meg az eredményt! Jelöljük ki több mintát!*

## 23. Gyakorlat

Osztályozás\_Indít\_Képforrás\_File\_Kép9\_Módszer\_Tanítóval-legközelebbi szomszéd\_

*Jelöljük ki egymás után négy pontot, és ezzel négy mintaterületet az input képen. Legyen az első minta a folyó, a második a bal felső sarokban látható vöröses mezőgazdasági tábla, a harmadik a mellette lévő élénkzöld tábla, a negyedik a barnás színnű erdős terület! Ezután:*

Végrehajt

*Vizsgáljuk meg az eredményt! Jelöljük ki több mintát! Hasonlítsuk össze az eredményeket a box módszerrel!*

*Változtassuk a távolságfüggvényt és nézzük meg hogyan változik az osztályozott kép! Mentsük le a különböző metrikával osztályozott képeket, majd kétképes műveletek segítségével magyarázzuk meg a különbségeket!*

## 24. Gyakorlat

*Ha előzőleg is osztályozást végeztünk akkor először lépünk ki, és indítsuk előről a funkciót. Ellenkező esetben az előző feldolgozás osztályait kiindulási osztályközéppontoknak fogja értelmezni a program.*

Osztályozás\_Indít\_File\_Tesztkép\_Kép9\_Módszer\_Tanító nélkül\_

*Ha előről indítottuk a programot, akkor a kezdő klaszterek száma 0. Kattintsunk az Új centrum gombra többször egymás után, és figyeljük meg hogyan alakul ki a klaszterezett kép!*

*Vizsgáljuk meg az eredményt! Hasonlítsuk össze az eredményt a másik két módszerrel!*

## 25. Gyakorlat

*Végezzük el az előző három gyakorlatot egy légifelvételen (IMAGES/LEGIFELV/44.tif) és egy eredeti felbontású űrfelvételen is (IMAGES/MUHOLD/SPOT/Sp1994ms.tif).*

*Milyen különbségeket találunk az eredeti gyakorlatokhoz képest?*

## Ellenőrző kérdések

1. Mit jelent a digitális kép osztályozása?
2. Mi az alakfelismerés?
3. Mi a textúraelemzés?
4. Mi a képpontértékelemzés?
5. Mit nevezünk multispektrális osztályozásnak?
6. Mit nevezünk multitemporális osztályozásnak?
7. Melyik osztályozási modell foglalkozik textúraelemek közti síkbeli összefüggésekkel?
8. Melyik osztályozási modell alkalmaz valószínűségszámítási módszereket?
9. Mi a lényegkiemelés?
10. Mi a jellemző a Bayes-féle döntési módszerre?
11. Mi a jellemző a távolságmérő módszerekre?
12. Milyen metrikákat ismer?
13. Mi a jellemző a klaszterezésre?

14. Milyen osztályozási módszert alkalmazna, ha az osztályok teljesen ismeretlenek?
15. Mivel csökkentik a figyelembe vett sajátosságok számát?
16. Melyik lényegtömörítési módszer eredményei teljesen függetlenek?
17. Mit nevezünk döntésfüggvénynek?
18. Milyen döntésről beszélünk, ha a hiba minimális?
19. Mi a hátránya és az előnye a Box-módszernek?
20. Mekkora a legközelebbi szomszédok módszer hibája?
21. Milyen módszert választana ismeretlen műholdfelvétel osztályozására?
22. Mikor használható a hierarchikus klaszterezés?
23. Mikor nem használható a hierarchikus klaszterezés?
24. Melyik esetben vezethető vissza a klaszterezés egy szélsőértékprobléma megoldására?
25. Mi lehet célfüggvény?
26. Mi lehet jellemző az ISODATA-eljárásra?
27. Melyik példa veszi figyelembe a környezetet osztályozáskor?
28. Mi a szintaktikus alakfelismerés lényege?
29. Mit nevezünk textúrának?
30. Mikor homogén egy textúra?
31. Mit nevezünk co-occurrence mátrixnak?

## 8. KÉPKÓDOLÁS ÉS TÖMÖRÍTÉS

### 8.1. Bevezetés

A (számítógépes) képfeldolgozás lényegében nagy adathalmazon (képen) alapuló **adattfeldolgozás**. Mivel az adatok tárolása, **átvitele** a képek esetében is egyike az alapfeladatoknak, az **adattömörítés** (itt: **képtömörítés**) fontos, külön tárgyalást érdemlő részprobléma. Ebbe a csoportba soroltuk az általános **képkódolást** is, amelynek célja a tömörítésen kívül pl. **hibavédelem**, **adattitkosítás** is lehet.

A **hagyományos képtárolás** adathordozói, pl. az ezüstvegyületek alkalmazásán alapuló fotóanyagok (**film**), illetve - különösen mozgókép rögzítésére - a képmagnetofon szalagja. A számítógépes képfeldolgozás és ezen adattárolási formák találkozására számos példa létezik (pl.: mikrofilm-alapú nagykapacitású képvisszakereső rendszer, számítógépes vezérlésű videokép archiváló rendszer). Sajnos a tapasztalat azt mutatta, hogy ezek a rendszerek sem **megbízhatóságban** (időbeli stabilitásban), sem **adathűségben**, sem pedig **adatkezelési sebességben** nem váltották be a hozzájuk fűzött reményeket, s általában **áruk** is igen magas. A korszerű digitális adathordozók (nagy tárolókapacitású mágneslemez, optikai lemez, CD-ROM) megjelenése ezeket a korábbi kísérleti megoldásokat háttérbe szorították.

Hasonló a helyzet a **képek átvitele** (hírközlő csatornán való átküldése) terén is. Bebizonyosodott, hogy az adatok digitális kezelése a tárolási adatsűrűségen kívül gyakorlatilag minden paraméter (adathűség, átviteli sebesség, tömörítési lehetőség) szempontjából előnyösebb az analóg megoldásoknál.

Mivel azonban mind az adattárolók, mind pedig a hírközlési csatornák használata költséges, természetes, hogy a digitális képtárolás és képátvitel megjelenésével egyidejűleg jelentkezett az az igény is, hogy a képi adatokat **minél hűebben, de ugyanakkor minél kisebb adatmennyiséggel lehessen ábrázolni**.

Ez **képtömörítéssel** érhető el.

A **képtömörítés** (*image compression*) iránti érdeklődés több mint 25 éves múltra tekinthet vissza. A kezdeti időszak erőfeszítései elsősorban az analóg módszerekre korlátozódtak, melynek tipikus példája a képi jelekre vonatkozó **sáv szélesség-csökkentés**. (Ez megfordítható jelszint- és frekvenciatranszformációkat jelentett, amelyek azonban mindig információvesztéssel is jártak, azaz a tömörítés végül is a jelhűség és a zajosság rovására történt. Napjaink valamennyi hírközlési rendszere tartalmaz ilyen jellegű jeltranszformációkat.)

A digitális képfeldolgozás terjedésével az érdeklődés eltolódott a **digitális képtömörítés** felé. Fontosságára nézve talán elegendő a **telefaxra**, vagy **HDTV** (*high defi-*

tion television) **szabványra** gondolnunk; egyikük sem lenne elképzelhető digitális képtömörítés nélkül.

A **telefax** a kétszintes képek igen hatékony **CCITT Gx (x=3 vagy 4)** kódolásán alapul; így vált lehetővé a szokásos telefonhálózaton keresztül a képátvitel. (Az eljárás a fekete-fehér **futamhosszak**at kódolja.)

A születőben levő **HDTV szabvány** a műsorszórásra jelenleg használatos TV közvetítési csatornakiosztás mellett teszi lehetővé a kb. hússzor akkora adatmennyiség átvitelét. (Az egyes képek kódolása veszteséges transzformáción alapuló kódolással történik, a sávszélességet az egyes **képfázisok** közti összefüggések kihasználásával csökkentik.)

Mint már említettük, a korszerű képbeviteli eszközök igen nagy adatmennyiség szolgáltatására alkalmasak. Pl.: a szokásos - 300 dpi - felbontású színes optikai lapolvasók egy **A4** méretű oldalról 25 Mbyte adatot szolgáltatnak. (Megemlítjük, hogy a hagyományos 1200 baud átviteli sebességű modem vonalakon ekkora képek átviteléhez kb. 50 órára lenne szükség. A szükséges átviteli idő a korszerű digitális telefonhálózat - ISDN B (64 kbit/s) esetén is kb. 1 óra lenne.) A tömörítés tehát alapvető igényként jelentkezik.

Napjainkban a **multimédia** eszközök használata szintén elképzelhetetlen lenne a kép és a hang rögzítése terén egyaránt alkalmazott adattömörítés nélkül.

A mindennapi élettől távolabbi területeken is igen nagy a képtömörítés jelentősége. Ilyen területek a katonai hírközlés, az űrtávközlés, vagy a videokonferencia.

A **tömörítési arány** kifejezés az eredeti tömörítetlen és a tömörített képek fájlméreteinek arányát adja. Amennyiben az eredeti tömörítetlen képem mérete 480 kByte, a tömörített képfájl mérete 10 kByte, akkor a:

$$\text{Tömörítési arány} = \frac{\text{Tömörítetlen fájl mérete}}{\text{Tömörített fájl mérete}} = \frac{480}{10} = \frac{48}{1} = 48:1$$

## 8.2. Redundanciák

Az **adattömörítés** (*data compression*) kifejezés azt a folyamatot jelöli, mely bizonyos információt reprezentáló **adatok mennyiségét** csökkenti.

Tisztán meg kell különböztetni egymástól az **adat** és az **információ** szavak jelentését; ezek nem egymás szinonimái. Egyszerűen megfogalmazva: adat az, ami az információt **hordozza, ábrázolja**. Ugyanazt az információt más és más adatokkal is meg lehet adni, s az adatok mennyisége is különböző lehet. Arra az adathalmazra,

mely nem a legkisebb mennyiségű adattal jellemzi ugyanazt az információt, azt mondjuk, hogy **redundáns**.

A redundanciának fontos jelentése van az adattömörítés szempontjából. A számítógépes képfeldolgozás területén három típusát különböztetjük meg:

### 1) Kódolási redundancia

**Lássunk egy egyszerű példát:** van egy olyan képünk, melyen a képpontértékek ábrázolása 1 byte-on történik. Ugyanakkor azonban, az ábrázolt képen összesen kétféle képpontérték található: a 0 jelöli a kép fekete, a 255 pedig a fehér részeit.

Világos, hogy ha a feketét továbbra is 0-val jelöljük, de a fehérhez ezentúl 1-et rendelünk, attól a kép még nem változik meg. Ugyanakkor így egy-egy képpont tárolásához az 1 byte helyett 1 bit is elegendő (ezáltal 8:1 tömörítést érünk el).

Az előbbi példa azt mutatta meg, hogy ugyanannak az információnak az eltérő kódolása csökkentheti a **kódolás redundanciáját**.

### 2) Képi redundancia

A kép többféle olyan belső összefüggéssel rendelkezhet, melynek kihasználása esetén az ábrázolásához kevesebb adat is elegendő. Néhány példa:

- **a képen csak azonos színű objektumok vannak homogén háttér előtt.** Hatékonyan kódolható pl. az objektumok határolóvonalainak megadásával. Az ún. **differenciális iránykód** alkalmazásával (azaz a határolóvonal végigjárásával úgy, hogy mindig az előző lépés irányától való **eltérést** kódoljuk) átlagosan 1.5 bit/határolópont érhető el.
- **egymás után következő képek csak kis mértékben különböznek.** Hatékonyan kódolható a képek közti **változások** megadásával.
- **a képen szabályos alakzatok vannak.** Hatékonyan kódolható az alakzatok generálását leíró függvénykapcsolatok megadásával.

### 3) Pszichovizuális redundancia

Bizonyos esetekben az az információ, melyet a kép tartalmaz ugyan, de az emberi látás számára nem hordoz információt, felesleges. (Például: ha csak a kép megjelenítése a feladat, de egzakt feldolgozására nincs szükség, a felesleges részletek megszüntetése még előnyös is, hiszen kevesebb adat tárolására, átvitelére van szükség.) A felesleges információ megszüntetésével adattömörítés érhető el.

Fontos megjegyeznünk, hogy míg az előző két csoport módszerei **nem eredményeztek** információvesztést, addig a pszichovizuális redundancia kihasználása **nem megfordítható**.

Néhány példa pszichovizuális redundanciára:

- **a képpontok ábrázolása 24 biten történik.** Ez kb. 16 millió szín megjelenítését teszi lehetővé, amire egyrészt a jelenlegi képmegjelenítő eszközök sem képesek, és ami az emberi látás felbontóképességét is jóval meghaladja. Nyil-

vánvaló, hogy elég csak a legjellemzőbb és egymástól érzékelhetően különböző színeket megtartani.

- **a képen látható objektumok** (pl. a képvétel gyenge fényviszonyai miatt) **erősen zajosak**. A képre alkalmazott zajszűrés a látás szempontjából felesleges információ mennyiségét csökkenti, s így a kép egyéb módszerek alkalmazásával hatékonyabban tömöríthető.
- **a kép „kevés látnivalót” tartalmaz, túlzottan nagy „méretben”**. Ekkor a kép mérete arányosan csökkenthető, amíg a csökkentés nem megy a látvány rovására. A képek tömörítése az előbbieken felsorolt redundanciák kihasználásán alapul.

### 8.3. Elvárások

**Általában** egy adattömörítési eljárástól a következőket várjuk el:

- legyen nagy hatékonyságú a **tömörítési arányt** illetően. (Ez más megfogalmazásban azt jelenti, hogy a tömörítés utáni adatmennyiség és a tömörítést megelőző adatmennyiség **aránya** legyen minél kisebb. Ehhez pl. az is szükséges, hogy a tömörítés algoritmus használja ki a tömörítendő adathalmaz speciális szerkezetét, illeszkedjen hozzá.)
- az eljárás **algoritmus** legyen hatékonyan implementálható. (Ez az elvárás a szélesebb körű felhasználásra számot tartó eljárások esetében azt jelenti, hogy mind szoftver, mind pedig hardver úton hatékonyan - gyorsan, kis ráfordítással - lehessen a tömörítést és a helyreállítást a gyakorlatban alkalmazni.)
- illeszkedjen meglévő rendszerek kötött lehetőségeihez (mind szoftver, mind pedig hardver tekintetében).

A **képek tömörítése** esetében a felsorolt általános elvárások kissé módosulhatnak. Például:

- az algoritmusok **használják ki a képek speciális adatszerkezetét**. Ez jelentős eltéréseket eredményezhet az általános adatstruktúrákon alkalmazható kódolási eljárásokhoz képest. Pl.: kihasználható, hogy a kép olyan kétdimenziós adathalmaz, melynek szomszédos adatai közt általában jelentős a korreláció. Továbbá: többsávos képek esetében az egyes sávok sem függetlenek egymástól; a mozgó képek esetében pedig az egymás után következő képek közti összefüggések is felhasználhatók.
- az esetek jelentős részében - bizonyos feltételek teljesülése esetén - **megengedhető az információvesztés**. (Ezek a feltételek pl.: az információvesztés ne befolyásolja a látványt; statikus képek átvitele esetén az átviteli eljárás tartalmazzon az átviteli hibákat idővel korrigáló mechanizmust.)

## 8.4. Veszteségmentes tömörítés

A **veszteségmentes képtömörítés** lehetővé teszi, hogy a helyreállítás során az eredeti képpel tökéletesen azonos képet kapjunk vissza. (Ez azt jelenti, hogy a bemeneti és kimeneti kép összes képpontja azonos.)

Ez az eljárás akkor indokolt, amikor a későbbi felhasználás nem teszi lehetővé azt, hogy a kép akár a legkisebb mértékben is módosuljon. Számos olyan alkalmazási terület van, amelynél a tömörítés során veszteség nem engedhető meg. Például az úrfelvételek, orvosi diagnózishoz szükséges röntgenfelvételek tömörítése csak a **kódolási-**, illetve a **képi- redundancia** kihasználásán alapulhat.

Sajnos, a szokásos összetételű tónusos (pl. a felvételi eszközök zaját is tartalmazó) képek esetében csak nagyon kis tömörítési arány érhető el, a 3:1 arány már jónak mondható. A zajszegény, vagy pl. számítógéppel létrehozott képek nagyobb arányban tömöríthetők.

### Egyszerű példa a veszteségmentes tömörítésre:

Ugyanannak a képnek a leírására a képpontértékek helyett azok pontról-pontra való változása is használható. Azonban, amíg a képpontértékek széles tartományban vehetnek fel értékeket, addig a változásértékek normális esetben (zajszegény képen) 0 átlagértékű, nagyon keskeny értékészletű függvényvel jellemezhetők. Ez abból következik, hogy a szomszédos képpontok nagy valószínűséggel közel azonos fényességek, illetve színűek. A leggyakoribb értékek tehát a 0, +1, -1, +2, -2,...; minél nagyobb a különbség abszolút értéke, annál kisebb valószínűséggel fordul elő. Ha tehát a felsorolt - gyakoribb - értékekhez rövid kódot, míg a ritkábban előforduló értékekhez hosszabbat rendelünk, s e változó hosszúságú kódokkal adjuk meg a változásokat, a kódolt eredmény tárigénye kisebb lesz, mint az eredeti képé. Ugyanakkor, az eredeti kép maradéktalanul helyreállítható. Az eljárás hatékonysága tovább növelhető, ha először meghatározzuk a változásértékek statisztikus eloszlását, s a változó hosszúságú kódokat az eloszlás ismeretében rendeljük a változásokhoz. (Ekkor természetesen a helyreállításhoz ennek a megfeleltetésnek az ismeretére is szükség van.)

A teljességre való törekvés nélkül lássunk néhány példát a veszteségmentes képtömörítésre.

### Változó hosszúságú kódolás VLC (*variable length coding*)

Ebben az esetben a tömörítés kizárólag a kódolási redundancia csökkentésén alapul. A szokásos esetekben elérhető a 2:1 tömörítési arány; nagyon jó esetben - pl. szkennelvel beolvasott kétszintes szöveges oldalak esetén - a 30:1 arány is elérhető.

Többféle változó adathosszúságú (VLC) kódolás létezik. A leggyakrabban használt



- az ún. **Huffman kódolás** - D. A. Huffman -tól származik, s az eljárás lényegét 1952-ben publikálta. Egy másik típus a az ún. **aritmetikai kódolás**.

### *1. Huffman kódolás*

Legyen  $n$ -féle bemenő adatunk, melyek előfordulási gyakorisága ismert. A Huffman kódolás minden egyes adathoz olyan kódot rendel, melynek hossza (bitszáma) fordítottan arányos az adathalmazban való előfordulásainak számával. A kódok megállapítására az alábbi egyszerű módszer is használható:

- Írjuk fel a bemenő adatokat egymás alá az előfordulási gyakoriságuk szerint rendezett sorrendben.
- Induljunk el a tábla két legkisebb gyakoriságú eleménél, és vonjuk össze őket úgy, hogy egyben adjuk össze a hozzájuk tartozó gyakoriságértékeket.
- A páronként történő összevonást mindaddig folytassuk, amíg az összeggel el nem érkezőnk a teljes képpontszámig.
- Ezután minden egyes adathoz olyan kódot rendelünk, melynek bitszáma megegyezik az adatot tartalmazó összevonások számával. A kód azon bitjei legyenek „1” értékűek, amelyekhez tartozó összevonásokban a rendezett sorban a párjánál feljebb volt.

A **statikus** Huffman kódtábla az adott kép statisztikai jellemzőitől független, előre elkészített táblázat, amely pl. nagyszámú hasonló kép elemzése útján alakult ki. Erre példa az egydimenziós faxkódolás.

A **globálisan adaptív** Huffman kódtábla az adott (teljes) kép **hisztogramjának** elkészítése után, a fentebb leírt generálási algoritmussal készül. Hátrányos, hogy ehhez a kép kétszeri feldolgozására van szükség (először csak a hisztogramszámítás érdekében, másodszer pedig a tényleges kódoláshoz). Az eljárás előnye, hogy a statikusnál nagyobb tömörítési arányt eredményez.

A **lokálisan adaptív** Huffman kódtábla a kódolás folyamán nem állandó; az adatokhoz tartozó gyakoriságértékek kísérik a kódolást, s a kódok jelentése is ennek megfelelően alakul. A legmagasabb tömörítési arányt ez adja, azonban hátrányos, hogy a folyamatos táblafelújítás meglehetősen nagy számításigényű. Alkalmazása a drága adatátviteli csatorna jobb kihasználása érdekében indokolt.

### *2. Aritmetikai kódolás*

Bár az aritmetikai kódolást a változó hosszúságú kóderek közé soroljuk, alkalmazásával végül is egy nagy **bitfolyam** áll elő, melyben a bemenő adatok és a kimenő adatok részletei közt nincs egyértelmű megfeleltetés. Az eljárás a bemenő adatok sorozatának egy részletéhez rendel kódot.

Minden egyes kimenő kód 0 és 1 közé eső valós szám, melynek a bitfolyam a számábrázolási korlátok következtében csak egy **szeletét** ábrázolhatja.

Anélkül, hogy a működést részletesebben ismertetnénk, megemlíjtük az eljárás két fő előnyét:

- A **lokálisan adaptív** változata is könnyen programozható, s így igen nagy tömörítési arány érhető el (ilyen például az **IBM** által kidolgozott ún. **Q-kóder**).
- Ellentétben a Huffman kódolással, az adatokhoz nemcsak egész számmal megadható hosszúságú kód tartozhat, így a kódolás annál nagyobb hatékonyságú lehet.

## Bitsík kódolás

Szintén a kódolási redundanciára épül, de már a képpontértékek közti korrelációt is kihasználja a **bitsík kódolás**.

Lényege: a bemenő képet kódolás előtt bontsuk annyi független képre, ahány bitek a képpontértékei: minden kép egy-egy adott bitsík adatait tartalmazza. Az így kapott egybites képeket tömörítsük hatékony (azaz nagy tömörítési arányt eredményező) veszteségmentes eljárással. A szokásos képek esetében a fenti felbontás különböző összetettségű képeket szolgáltat; a magasabb helyiértékű bitekhez tartozó képek lassan változó, azaz könnyen kódolható foltokat tartalmaznak. De a gyorsan változó alacsonyabb helyiértékű bitek kódolása is előnyösebbé válik, mert a szétválasztás általában hatékonyabb kódtábla használatát teszi lehetővé.

Megjegyezzük, hogy a helyreállításnál a kép **progresszíven** alakul ki, azaz először a durva részletek tűnnek elő, s a kép egyre finomabb felbontású lesz. Ez a tulajdonság számos alkalmazásnál előnyös.

## Homogén foltok kódolása

Az alábbi eljárás a túlnyomórészt nagy homogén foltokat tartalmazó kétszintes képek - vagy bitsíkok - hatékony tömörítésére használható. (A eljárás eredménye **quad-tree** néven is ismeretes, mert lényegében a képpontadathalmaz speciális fára való felűzését írja elő.)

Legyen a bemenőkép mérete:  $[m \times n]$ . Az algoritmus röviden a következő:

- Vizsgáljuk meg az összes képpontértéket a képen. Ha

- minden képpont azonos (azaz vagy „0”, vagy „1”), akkor a képhez rendeljük a „**homogén0**” vagy „**homogén1**”-kódot, és ezzel a kódolás véget is ér. Ha

- a képpontok nem azonosak, bontsuk a képet négy, azonos méretű részre (**subimage**) az eredeti oldalak megfelelő **homogén** kódolása érdekében. (Ekkor a részek mérete  $m/2$ , illetve  $n/2$  képpont lesz.
  - valamennyi részképre egymás után vizsgáljuk meg a homogenitást; amelyik homogén, ahhoz rendeljük a megfelelő **homogén** kódot, amelyik nem homogén, azt bontsuk tovább oldalfelezéssel, és vizsgáljuk tovább részenként.
  - az eljárás akkor ér véget, ha nem marad több részre bontható inhomogén folt.

## Határolóvonal kódolás

A homogén foltok hatékony kódolásának másik megközelítése: kódoljuk a foltokat határoló vonalakat, és jelöljük ki azokat a zárt tartományokat, amelyek nem a háttérhez, hanem az objektumokhoz tartoznak.

Többféle változat ismeretes, például:

**1D és 2D futamhossz kódolás** (*run-length encoding*). Sorfolytonosan dolgozza fel a képet, s az egyszínű képsorszakaszok **hosszát** kódolja hatékonyan. Ez lényegében a foltok határolóvonalainak pontonkénti kódolását jelenti, úgy, hogy e pontok közti szomszédsági összefüggést nem használjuk ki. Míg az 1D (egydimenziós) változat minden szakasznak önálló kódot feleltet meg, addig a 2D (kétdimenziós) változat kihasználja, hogy a futamhosszak a szomszédos sorok közt jelentős korrelációban vannak, s a kismértékű változásaikra külön kódokat használ.

(Ilyen kódok pl.: a szakasz kezdetének ill. végének helyzete nem változott; balra illetve jobbra tolódott 1,2, ill. 3 képponttal; a szakasz helyzete jelentősen változott.)

Tipikus példa e módszerekre: **CCITT G<sub>x</sub>** ( $x=3$  vagy  $x=4$ ) faxkódolás.

## Kontúrkövetés (*contour tracing*)

Hardver megvalósítás szempontjából a futamhossz kódolásnál előnytelenebb, viszont bizonyos típusú képekre hatékonyabban alkalmazható tömörítési eljárás. Lényege: használjuk ki az objektumok határolóvonalainak pontjai közti szomszédsági összefüggést, azaz menjünk végig a határolóvonalon pontról-pontra. Egy adott határpontból a szomszédosba nyolcféleképpen juthatunk el, ami (a vonalelágazásoktól, vonalvégződésektől, stb. eltekintve, azaz csak egyszeres zárt görbék feltételezve) három biten kódolható: ezek az ún. **iránykódok**. Ezt kihasználva csak a kontúr egyik pontjának ( $x, y$ ) koordinátáit kell megadni 2-szer 8 biten, ezt követően minden további pont megadásához már csak az iránykód, azaz 3 bit szükséges.

További kódolási redundancia az, ha nem használjuk ki, hogy a **szomszédos iránykódok közt is van korreláció**. A határolóvonal ugyanis általában bizonyos „tehetet-

lenséggel” változtat irányt, azaz a szomszédos iránykódok általában közel azonos irányba mutatnak. Célszerű ezért az előző iránytól való eltéréseket kódolni; s minél kisebb az irányváltoztatás, annál rövidebb kódot alkalmazni. Az így kapott kódokat nevezik **differenciális iránykódoknak**.

### Előrebecsléses (predictive) kódolás

A képi redundancia kihasználásának egyik legegyszerűbb módszere az **előrebecslés**. Ez azon alapszik, hogy a kép közeli részletei között nagymértékű a korreláció, azaz a kép adott részlete alapján annak közeli környezete többé-kevésbé megbecsülhető. Legegyszerűbb példa:

- ha sorfolytonosan haladunk balról jobbra, és az aktuális és azt megelőző képpontértékeket már ismerjük, akkor a sorban következő képpontról feltételezhető, hogy az ezeknek megfelelő tendencia arra is érvényes. Ha pl. a megelőző képpontértékek azonosak voltak, akkor a következő képpontérték is nagy valószínűséggel ugyanolyan lesz.
- ezért mondjuk azt, hogy a kép helyreállítása alapvetően előrebecsléssel történik, s a kódolás során a kép helyett csak azokat a korrekciós adatokat tároljuk el, melyek az előrebecslés tévedéseit kijavítják. Nagy valószínűséggel az előrebecsléseket csak kismértékben kell korrigálni, így a gyakoribb korrekciós adatokat rövidebb kóddal lehet ábrázolni. A módszer különösen nagy homogen foltok, lassú átmenetek (videók) esetén működik jól.

## 8.5. Veszteséges tömörítés

A **veszteséges képtömörítés** értelemszerűen nem teszi lehetővé az eredeti képtartalom maradéktalan helyreállítását.

A leggyakoribb elvárás a **látvány változatlansága**, ami (mint később erre részletebben is kitérünk) nem teszi szükségessé a képpontértékek pontról-pontra való azonoságát.

A veszteséges eljárásokkal viszonylag nagy tömörítési arány érhető el, a megengedett **minőségromlás** mértékétől függően. Tipikus a 20:1 arány, de ha csak a kép **felismerhetősége** a cél, a 100:1 arány is elérhető. Különböző összetett eljárások esetében (pl.: az alakzatok kontúrjának elkülönített kódolásával, valamint a foltok **textúrájának** az emberi látás sajátos törvényszerűségeit is figyelembe vevő kódolásával) a 70:1 arány még jó minőségű helyreállítást biztosít.

Számos veszteséges tömörítési eljárás ismeretes; ezek két fő csoportba sorolhatók:

- **előrebecslésen** (*prediction*) alapuló módszerek. Ezek általában kevés információ - pl. egy kis képrészlet - ismeretében nagyszámú képpontértéket becsülnek meg előre. A becsléssel előállított és tényleges értékek közti eltéréseket a tömörítés pontossági igényétől függő mértékben korrigálni kell; az eljárás lényegében a kezdőfeltételek rögzítése után ezeknek a korrekciós adatoknak az előállításán alapul. Jó esetben a korrekció mértéke is viszonylag kicsi. Ekkor a korrekciós paraméterek hatékonyan adhatók meg változó hosszúságú kódokkal.
- **matematikai transzformáción** alapuló módszerek. A kép leírása a képpontértékek helyett bázisfüggvények együtthatóinak megadásával történik. A kép helyreállítása értelemszerűen ugyanezen bázisfüggvények adott pontbeli súlyozott összegzésével történik. A tömörítésre az adja meg a lehetőséget, hogy ugyanannak a **látványnak** az előállításához nem feltétlenül szükséges, hogy az együtthatók változatlanok legyenek. Az emberi látás sajátosságai nagymértékű pontatlanságot, elhanyagolásokat engednek meg az együtthatók kódolása során.

## 1. Veszteséges előrebecsléses (predictive) kódolás

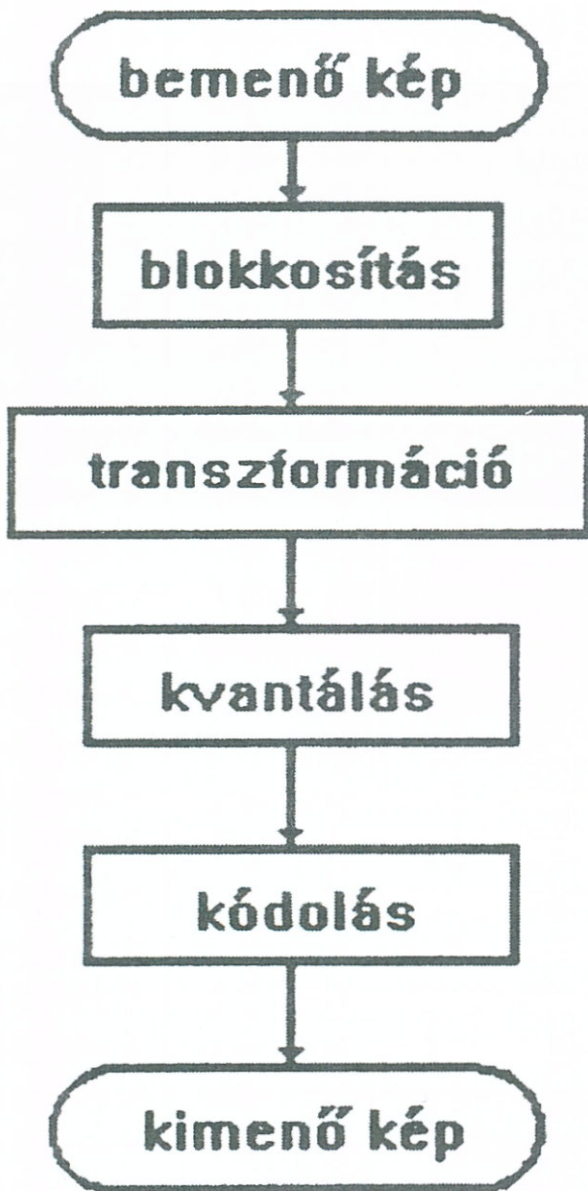
Az előző pontban ismertetett **előrebecsléses módszer** jó példát szolgáltat arra, hogyan lehet megengedni a veszteséget a kódolás során. A lényeg rendkívül tömören: csak akkor kerüljön az előrebecslés hibáját korrigáló adat a kódok közé, ha az előrebecsléssel elkövetett hiba **egy bizonyos küszöböt meghalad**. E küszöb értékének változtatása befolyásolja a helyreállítás hűségét, s egyben a tömörítési arányt.

## 2. Transzformációs kódolás

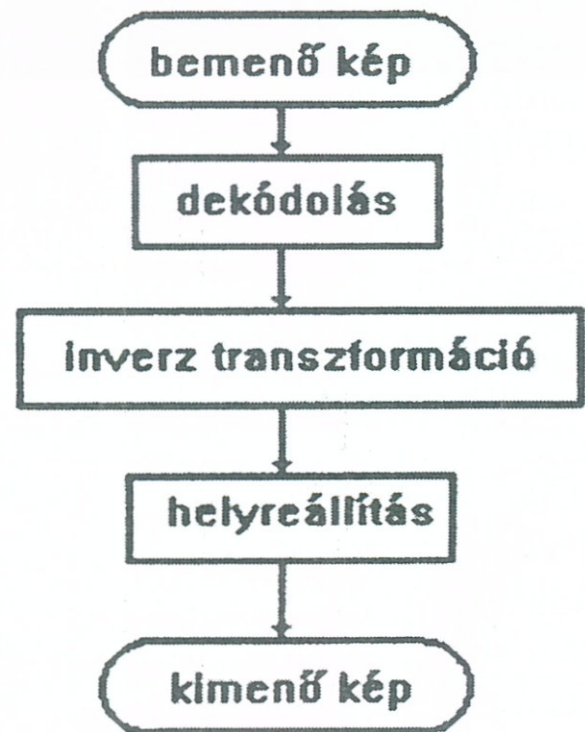
Az előző pontokban ismertetett módszerek közvetlenül a képpontértékek felhasználásán alapultak (közös angolszász megnevezésük is erre utal: *spatial domain methods*). A **transzformációs kódolás** olyan megfordítható lineáris transzformáció, mely a képet (illetve általában annak kisebb részleteit, az ún. blokkokat) transzformációs együtthatókra képezi le, melyeket kvantál és kódol. A legtöbb **természetes képre** a transzformációs együtthatók között sok a kis abszolút értékű; ezek torzítás megengedése esetén **elhanyagolhatóak**.

A sokféle transzformációs eljárás közül (például: **Karhunen-Loeve (KLT)** transzformáció, diszkrét **Fourier (DFT)** transzformáció, **Walsh-Hadamard (WHT)** transzformáció, **Fractal** transzformáció, **Wavelet** transzformáció) a hatékonyság, használhatóság alapján lehet választani.

Megállapítható, hogy az előállított együtthetők **függetlensége** szempontjából a Karhunen-Loeve transzformáció optimális, de sajnos a gyakorlati felhasználáshoz megfelelő hatékonyságú implementáció nem ismeretes.



18. ábra



19. ábra

Közel ugyanilyen jó tulajdonságokkal rendelkezik a **diszkrét koszinusz (DCT)** transzformáció, melynek használatához ráadásul igen hatékony számítási eljárások ismeretesek. Érthető ezért, hogy ennek a transzformációnak a szerepe az elmúlt időszakban megnőtt.

A 18. folyamatábrán a **transzformációs képkódolás** modellje látható. A 19. ábra a transzformációs kódolással nyert adatokból a **kép helyreállításának** modelljét szemlélteti.

## 8.6. JPEG szabvány

A **JPEG** (*Joint Photographic Experts Group*) ajánlások, melyeket az **ISO** (*International Standards Organization*) és a **CCITT** (*Consultation Committee on International Telephon and Telegraph*) szervezetek által 1986-ban felállított munkacsoportban születtek, napjainkra az álló képek veszteséges tömörítésére szabványossá váltak. (Így például: a közismert **TIFF** képfájl-formátumban külön tag van fenntartva a **JPEG** eljárással tömörített képformátum jelzésére.)

A diszkrét koszinusz transzformáción alapuló eljárással a látvány kismértékű romlása árán kb. 30:1 tömörítési arány is elérhető. A tömörítési arány illetve az esetleges minőségromlás mértéke felhasználói paraméterrel előírható.

Az eljárás eleve úgy lett megválasztva, hogy hardver és szoftver úton egyaránt hatékonyan lehessen megvalósítani.

Az eljárás a **szín összetevőket** egymástól függetlenül kezeli, azaz színes kép esetében a három összetevő mindegyikére végre kell hajtani. Ha a tömörítés veszteségmentes lenne, (eltekintve az esetleges kódolási redundanciától) nem lenne túl nagy jelentősége annak, hogy a képpontértékek milyen színrendszerben vannak ábrázolva. Mivel azonban a JPEG szabvány megengedi a veszteséget, értelemszerűen célszerű azt a képpontérték ábrázolást választani, amely az elkövetett hibával szemben a legkevésbé érzékeny, illetve amelyben az adatok belső összefüggései a legnagyobb tömörítési arány elérését teszik lehetővé.

**Pszichovizuális kísérletek** bebizonyították, hogy többféle előnnyel jár, ha a szokásos **RGB** színrendszer helyett az **YUV** színrendszert használjuk. A leglényegesebb: ezáltal a színösszetevők adatai a látás szempontjából fontosabb és kevésbé fontos adatokra válnak szét. Az emberi látás ugyanis az U- és V- (ún. **krominancia**-) összetevőkre sokkal kevésbé érzékeny, mint az Y- (ún. **luminancia**-) összetevőre, azaz sokkal pontatlanabb helyreállításuk is elfogadható a látvány szempontjából. Ebből következően, célszerű a színes képet a tömörítés előtt ebbe a rendszerbe transzformálni.

Az eljárás első lépése: az adott színösszetevőt ábrázoló képsávot (a továbbiakban egyszerűen csak: képet) egymástól független,  $8 \times 8$  képpontból álló blokkokra bontjuk. A továbbiakban ezeket a blokkokat (csaknem teljesen) függetlenül kezeljük. (Megjegyezzük, hogy a blokkok közti egyedüli összefüggést azok képpontérték-átlagainak különbségi kódolása jelenti.)

A **JPEG szabvány (adaptív) diszkrét koszinusz transzformáción (DCT)** alapuló képtömörítési eljárás.

A transzformáció végrehajtásával blokkonként  $8 \times 8 = 64$  db, **2D** (kétdimenziós) diszkrét bázisfüggvényhez tartozó együtthatót kapunk.

Az eljárás képlettel a következőképpen fejezhető ki:

$$DCT(k_1, k_2) = \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} x(n_1, n_2) * \cos(2\pi(2n_1+1)k_1*(4N)^{-1}) * \cos(2\pi(2n_2+1)k_2*(4N)^{-1})$$

ahol:

$N$	a blokk mérete (a JPEG szabványban: 8)
$x(n_1, n_2)$	a képpontérték
$n_1, n_2, k_1, k_2$	blokkrelatív pozíciók

Az előbbi képlet közvetlen számításokra alkalmatlan, mivel valós adatokon nagyszámú művelet elvégzését igényli. A transzformáció igen hatékonyan számítható közelítése pl. az **RVFFT** (*Real Valued Fast Fourier Transform*) eljárásen alapul.

A **DCT** transzformáció inkább több adatot eredményez, mint kevesebbet, hiszen a kapott együtthatókat nagyobb pontossággal kell ábrázolni, mint az eredeti képpontértékeket. Akkor tehát mi értelme van az egésznek?

Míg a blokk 64 képpontja egymással többnyire erősen összefügg (*korrelált*), a transzformáció után kapott 64 együttható gyakorlatilag nem mutat belső összefüggést. (A diszkrét koszinusz transzformáció a **Karhunen-Loeve transzformációt** megközelítő mértékben tudja az adatokat független paraméterekké alakítani.) A **pszichovizuális kísérletek** bebizonyították, hogy e 64 együttható nem azonos mértékben fontos az eredetit közelítő látvány létrehozásához. Kiderült, hogy a magasabb frekvenciájú képtartalom változást reprezentáló bázisfüggvények még akkor sem játszanak a látvány létrehozásában túl nagy szerepet, ha a hozzájuk tartozó együttható értéke nagy. (Megjegyzendő, hogy a kísérleteket megadott képernyőméret-látótávolság összefüggések mellett végezték, s a **DCT** tömörítési eljárás alkalmazása esetén is feltételezik a képhelyreállítás bizonyos szinten **standard** körülményeit.)

Az előbbiekből tehát az következik, hogy a magasabb frekvenciájú összetevőkhöz tartozó együtthatók sokkal durvábban (pontatlanabban) kódolhatók, mint az alacsony frekvenciához tartozóak. A kódolás pontossága úgy is megközelíthető, hogy megszabjuk, melyik együttható **hányféle** értéket vehet fel, azaz milyen pontosságú a **kvantálása**. (A kvantálás pl. az együttható adott számmal való **osztásával** valósítható meg, az egészre kerekített hányados lesz az eredmény. Minél nagyobb az osztó, annál kevesebbféle eredmény jöhet ki, azaz - az osztóval való szorzással - annál durvábban állítható helyre az eredeti együttható.)



Pl.: ha a DCT utáni együttható értéke 11, a megfelelő kvantálási együttható viszont 30, akkor az eredmény 0 lesz. Ugyancsak nulla lesz az eredmény, ha az együttható értéke a  $\{-15, \dots, 15\}$  tartományban bármilyen más értéket vesz fel.

Tipikus kvantáló - (azaz 64 db, együtthatóról-együtthatóra változó osztót tartalmazó) - „mátrix” luminanciához:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Tipikus kvantáló - (azaz 64 db, együtthatóról-együtthatóra változó osztót tartalmazó) - „mátrix” krominanciához:

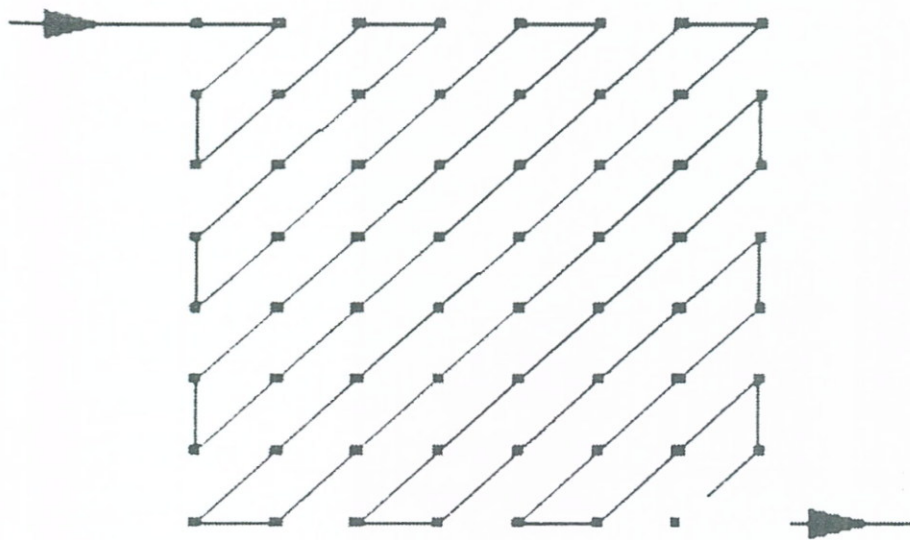
17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

A szokásos JPEG -alapú rendszerekben a képminőséget (s ezzel a tömörítési arányt is) pl. a kvantálási szintek változtatásával lehet befolyásolni. Minél „feljebb” toljuk a kvantálási szinteket, annál durvább lesz a kép, viszont annál nagyobb lesz a tömörítés aránya.

- A kvantálás után kapott 64 db. együttható túlnyomó része 0, vagy igen kis abszolút értékű szám lesz. (Már itt megjegyezzük, hogy a blokk eredeti képtartalma - illetve annak közelítése - úgy állítható elő, ha az előbbieken kapott számokat megszorozzuk a megfelelő kvantálási együtthatókkal, s az így kapott - közelítő - együtthatókkal a **DCT** inverzét hajtjuk végre.)
- A kvantálási adatok megválasztásából következik, hogy a kapott számok közül a legtöbb nem zérus értékű a blokk balfelső sarkának közelében található,

s attól távolodva, egyre több lesz a zérus együtttható. Ez a tény igen jól kihasználható az együttthatók kódolásához.

- Rendezzük át az együttthatókat a blokk balfelső sarkától a jobbsó sarok felé haladó cikk-cakk útvonalon való elhelyezkedés sorrendje szerint. A 20. ábra táblázatában minden együtttható helyén az a szám szerepel, amely megmondja, hogy az átrendezés után hányadik helyre kerül (felül az útvonal képe látható):



20. ábra

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64
36	37	49	50	58	59	63	64

Az első együtttható a konstans-függvényt reprezentálja, amely a blokkon belüli képpontértékek átlagértéke. Ezt az együttthatót (az elektronika „egyenáram” - DC elnevezésének analógiájára) **DC együttthatónak** nevezik. A többi együtttható a nemkonstans bázisfüggvényekhez tartozik; ezeket AC együttthatóknak nevezik.

A JPEG szabvány szerint a blokkok **DC együttthatói** az **AC együttthatóktól** függetlenül vannak kódolva: ehelyett a blokkról-blokkra való változásukat tárolják. Ezután kódolják változó hosszúságú kódokkal (adaptív VLC [variable length coding] eljárással) az **AC együttthatókat**.

## Sorfolytonos, illetve progresszív képfelépítés

A tömörítés eredménye az előbbieken leírtak szerint olyan **bitfolyam** mely közvetlenül felhasználható az eredeti képtartalom rekonstrukciójára. Ekkor **sorfolytonos képfelépítésről** beszélünk. Ez a JPEG eljárás leggyakoribb felhasználási módja; pl. képi másolat készítéséhez erre van szükség.

Egyes felhasználási területeken kiemelt fontosságú, hogy a kép a helyreállítás során minél hamarabb kerüljön felismerhető állapotba, még ha ennek az is az ára, hogy a teljes rekonstrukció az előbbinél kissé tovább tart. A helyreállításnak ezt a változatát nevezik **progresszív képfelépítésnek**. Előbb a **DC együtthatók** alapján épül fel egy igen durva kép, melynél minden  $8 \times 8$  pontos blokkot egy-egy átlagérték jellemez. A következő ciklusban az első **AC együtthatók** szerinti módosítás történik meg minden blokkban; ez  $4 \times 4$  pontos mintákat eredményez. A módszer így halad az egyre finomabb felbontás felé. Tipikus alkalmazási területe: képi adatbázisok kezelése, amelynek során a helyreállítás durva képből egyre finomabb képet hoz létre, s kívánságra - ha kiderül, hogy másik képre van szükség - a folyamat bármikor megszakítható.

### JPEG2000

Jelenleg fejlesztés alatt áll a JPEG2000 elnevezésű tömörítő eljárás, amely az eddig használt hatékony (elsősorban matematikai transzformáció(k)ra épülő) eljárások közül több módszert (algoritmust) is használ. Illeszkedik majd, a digitális kamerák és videók által készített állóképek több különböző felbontást is igénylő elvárásaihoz, ugyanakkor veszteségmentes vagy közel veszteségmentes tömörítésre is lehetőséget ad. Bevezetését 2002 közepére tervezik.

## 8.7. Fraktál tömörítés

A DCT transzformáción alapuló eljárásoktól gyökeresen eltérő, rendkívül hatékony képtömörítési módszer, az ún. fraktál (önhasonló, matematikai alakzatok) transzformáción alapuló eljárás. Az eljárás veszteséges képtömörítést takar. A kifejlesztője Michael Barnsley által létrehozott Iterated Inc. cég.

A tónusos vagy színes képek fraktálokra alapuló tömörítésének lényege, hogy természetben előforduló képek zöme éles kontúrokat, folytonos átmeneteket és általában ismétlődő (önhasonló) mintákat tartalmaz.

A tömörítő program először felosztja a teljes képet apró részecskékre, majd minden egyes részhez (domain) megkeresi a hozzá leginkább hasonlító, de eltérő méretű és helyzetű másik részletet (range). Tároláskor a domainek színét és a leképezés adatait szükséges csupán menteni. A leképezéshez szükséges matematikai transzformációs el-

járások jól ismertek. Létezik a transzformációnak hatékony számítógépes implementációja. Ugyan jelentős a tömörítés irányú számítási igénye az eljárásnak (800x600x24 pixelxpixelxbit-es kép esetén, Pentium 100 processzorral kb. 1-2 perc, Pentium MMX 200 esetén ez kb. 30 másodperc, PII 300 esetén kb. 15 másodperc, ekkor a jó minőségű képek mérete kb. 25-30 Kbyte), azonban a jelenlegi „populáris” processzorok számítási teljesítménye már alkalmas a feladat elvégzésére.

A kitömörítés során a domaineik rekurzív módon épülnek fel. Az egyes domaineik tartalmát az előző domain tartalma és a tárolt leképezés határozza meg. Amint a két egymás utáni kicsomagolási folyamat között minimálissá válik a különbség, a kitömörítés befejeződik. Mivel a kitömörítés (helyreállítás) sokkal gyorsabban elvégezhető (800x600x24 pixelxpixelxbit-es kép esetén, Pentium 100 processzorral kb. kb. 1-2 másodperc) mint a betömörítés, a multimédiás és prezentációs alkalmazásoknál gyakran használt állóképek esetén a felhasználók szinte nem is veszik észre. A helyreállítás után a kép nagyítható, s ekkor látszatra új részletek jönnek elő a globális mintázatok lokális ismétlésével. Az eljárással rendkívül nagy tömörítési arány érhető el (akár 100:1 is), kis minőségromlással.

Számos előnye van a széles körben elterjedt JPEG szabvánnyal szemben:

- sokkal nagyobb a tömörítés mértéke közel azonos minőség esetén,
- az emberi agy számára idegen „digitális” hatás nagyításkor kevésbé érezhető,
- jobb a kontúrok visszaadása,
- a színárnyalatok visszaadása feltűnően jobb FIF formátum esetén.

A legtöbb program ugyan még nem tartalmazza a FIF formátumú képek olvasását jelenleg, azonban az állóképek ki- és betömörítésére szolgáló ún. TWAIN modul vagy önálló program a cég Internetes szerveréről szabadon letölthető és kereskedelmi forgalmat leszámítva szabadon használható is. Az Internetes és multimédiás alkalmazások mellett rendkívül hasznos lehet prezentációk, képi adatbázisok használatakor.

A fraktál és JPEG tömörített képek pszichovizuális összehasonlítását végezheti el az olvasó az alábbi címen: <http://www.georgikon.hu/visual.htm>.

## 8.8. Mozgó képek tömörítése

A mozgó képek számítógépbe vitele olyan eszközöket kíván, melyek képesek a szabványos videojel valósidejű digitalizálására és eltárolására. Ez a feladat mind sebességi (adatbeviteli), mind pedig tárolóhely problémát jelent. (Amennyiben pl. 768\*576 képpontos, 24 bites színes képeket szeretnénk valós (25 kép/s) időben eltárolni, az másodpercenként kb. 32 Mbyte adatot jelentene.) Ugyanígy - a felvett és eltárolt képsorozat lejátszásához is szükség van az adatok valósidejű átalakítására, s az eredeti mozgó képek - legalább közelítő - helyreállítására.

A probléma megoldását a mozgóképek valós idejű tömörítése jelenti.

Számos olyan program és algoritmus ismeretes, mely e feladatot kisebb „ablakban”, kizárólag szoftver segítségével oldja meg. A teljes képernyőt betöltő mozgó képek tömörítéséhez illetve helyreállításához azonban a jelenlegi (PC alapú) számítástechnikai eszközök mellett speciális hardver alkalmazására is szükség van. Természetesen csak olyan eljárás esélyes a szabványossá válásra, mely szoftverben és hardverben egyaránt hatékonyan és olcsón megvalósítható.

A tömörítési feladatot legegyszerűbben az egyes képek egymástól független kezelésével lehet megoldani. Ekkor elvileg bármely eddig ismertett tömörítési eljárás alkalmazható, de - mivel látásunk ebben az esetben még kevésbé érzékeny a helyreállítás során elkövetett hibákra - a nagyobb hatékonyság érdekében célszerű veszteséges eljárást választani. Így használható pl. a 8.6. pontban ismertett JPEG eljárás. (Ezt **M-JPEG** eljárásnak nevezik, ahol „M” - motion - a mozgást jelzi, bár a JPEG eljárástól lényegileg nem különbözik.) Az átlagos tömörítési arány 24:1, ami önmagában kevés: a kb. 1.3 Mbyte/s adatfolyamot a szokásos merevlemezek nem képesek valós időben fogadni. Ezért az M-JPEG eljárást a képfelbontás csökkentésével együttesen használják. Az elfogadott 352\*288 (PAL szabványhoz illeszkedő) képméret kb. 0.32 Mbyte/s adatfolyamot eredményez, ami folyamatosan eltárolható - viszont a kép minősége éppen csak megközelíti a VHS minőséget.

A képek független tömörítése nyilvánvaló luxus, hiszen így olyan redundancia marad kihasználatlanul - az egymás utáni képek nagyfokú azonossága - mely az adatmennyiség további csökkentését teszi lehetővé.

Az 1988 végén alakult **MPEG (Motion Pictures Experts Group)** bizottság azt a célt tűzte maga elé, hogy mozgóképek és hang tömörítéséhez szoftverben és hardverben egyaránt hatékony eljárást dolgozzon ki. Az első változat (**MPEG-1**) 1990 végére született meg: ez kb. 0.18 MByte/s adatfolyammal a VHS minőséget biztosította. Elsősorban multimédia CD készítésére volt alkalmas. Azonnal felmerült a magasabb minőségi követelményeket is kielégítő szabvány iránti igény. Az **MPEG-2**-nek nevezett szabvány kb. 0.5-1.0 MByte/s adatfolyammal az interlaced, műsorszórásra alkalmas minőségű átvitelt biztosítja, támogatja a 16:9 képarányt és a többcsatornás videojelátvitelt is. 1991-től a harmadik fázisként (**MPEG-3**) a HDTV (High Definition Television) volt a bizottság fő célkitűzése. Az 1992-től használható **MPEG-4**, tovább fokozott minőségi/hatékonysági igényeket elégít ki. Céljául az Internet-en történő átvitelt tűzte. Napjainkban az MPEG a mozgókép rögzítésének valamennyi területén alapvető szerepet játszik. Az olvasó különböző formátumú videókat tekinthet meg a CD - VIDEO könyvtárban.

## Mi az MPEG eljárás lényege?

Az MPEG tömörítés bonyolult, soklépéses eljárás. Egyes részmegoldásai más eljárásokon alapulnak - így pl. a JPEG -en, mely DCT és Huffman kódolást alkalmaz -, míg az előzőekben tárgyalt módszerekhez képest alapvetően új részmegoldásokat is tartalmaz. A leglényegesebb új vonása az, hogy az időben szomszédos képek közti redundanciát csökkenti, azaz amikor lehetséges, csak a képek közti változást kódolja. Biztonsági okokból azonban időről-időre „újrakezdi” az egymásra épülő kódolást, s ezzel meggátolja az esetleges adatátviteli hibák „továbbterjedését”.

Leegyszerűsítve, az adatfolyam három, különböző szerepet betöltő képtípus egymásutánjából áll:

- **Különálló kép.** Kb. negyed másodpercenként egy teljes kép szerepel az adatfolyamban, JPEG kódolt formában. Miután ez a kép a következő adatok számára mintegy referenciaként szolgál, viszonylag alacsony tömörítési arányú (12:1), viszont jó minőségű.
- **Előrebecsült kép.** Az előző képhez viszonyított változást kódolja, s ezzel magas tömörítési arányt biztosít. Minél több előrebecsült kép következik egymás után, annál inkább eltérhet a helyreállított képsorozat az eredetitől.
- **„Kétirányú” kép.** A megelőző és következő kép átlagát használják fel, ezáltal nagyon magas tömörítési arányt érnek el.

Annak meghatározása, hogy mikor melyik képtípust kell kódolni, a képsorozat tartalma alapján az MPEG kóder feladata - ez alapvetően befolyásolja a felhasználó által beállított minőségi faktor.

## Gyakorlatok a képkódoláshoz és tömörítéshez

### 25. Gyakorlat:

Színrendszerek\_Indít\_Szín\_Színes\_Képforrás\_Tesztkép\_Színskála

*Az egérrel kattintsunk a színskála egyes mezőire! Figyeljük meg az adott ponthoz tartozó összetevők értékeit a különböző színrendszerekben!*

*Ismételjük meg a gyakorlatot a Képl-gyel!*

### 26. Gyakorlat:

*A gyakorlat elindítása előtt állítsuk át a VGA felbontást 16 színűre!*

## Dithering\_Indít\_Szín\_Színes\_Képforrás\_File\_Kép1

*A korábban már többször alkalmazott kép a VGA lerontott felbontásának eredményeként a megszokottnál sokkal rosszabb minőségű. Célszerű a gyakorlat folytatása előtt elolvasni a Help-ben lévő elméleti ismertetőt. Ezután folytassuk a gyakorlatot és próbáljuk végig a különböző módszereket, figyeljük meg az eredményt! A gyakorlat azt illusztrálja, hogy a kép gyenge minőségű megjelenítését javítani lehet különböző speciális kódolási eljárásokkal.*

*A gyakorlat után ne felejtjük visszaállítani a VGA felbontását!*

### **27. Gyakorlat:**

Zajszűrés\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép2\_Szűrőméret\_1  
\_Végrehajt\_Képforrás\_Utolsó eredménykép\_

*Jelöljük ki teljes szélességében a kép felső harmadát az egérrel, majd folytassuk:*

Szűrőméret\_3\_Végrehajt\_Képforrás\_Utolsó eredménykép\_Kilép\_Átszínezés  
\_Indít\_Átszínezőtábla\_Identikus\_Végrehajt\_Képforrás\_Utolsó eredménykép\_

*Jelöljük ki a kép alsó harmadát teljes szélességben, majd:*

Átszínezőtábla\_Szintrevágás\_BW\_Végrehajt\_Képforrás\_Utolsó eredménykép\_Profil\_Indít\_

Ha jól állítottuk be a Szintrevágást (80 körül), akkor az input képmező alsó harmada egy jó minőségű kétszintre vágott kép, a középső harmada az eredeti, a felső harmada a 3\*3-as átlagoló szűrővel szűrt kép. Jelöljük ki a képen teljesen végigfutó vízszintes szakaszokat az alsó, középső és felső harmadban! Milyen hosszú vízszintes szakaszok találhatóak a profilban a kijelölt szakaszok mentén? Milyen kódolást lenne érdemes alkalmazni ennél a képnél, ha tömöríteni szeretnénk a képet?

### **28. Gyakorlat:**

Hisztogram\_Indít\_Szín\_fekete/fehér\_Képforrás\_File\_Kép7\_Számol  
Konvolúció\_Szűrő mérete\_3\*3\_Szűrési mód\_Def.szűrővel\_Szűrődef.

*Adjuk meg a 13. Gyakorlatban definiált vízszintes irányú differencia szűrőt!*

Végrehajt\_\_Képforrás\_Utolsó eredménykép\_Kilép\_Hisztogram\_Indít\_  
Számol

*A hisztogram alapján eredményes lenne-e a különbség képen alapuló, változó hosszúságú kódok alkalmazása tömörítésre?*

## Ellenőrző kérdések

1. Mi a különbség a kép tömörítése és a kép kicsinyítése között?
2. Milyen változást eredményez a kép veszteséges tömörítése?
3. Mely eszközökben használják a CCITT G3 tömörítési eljárást és mi az eljárás lényege?
4. Hasonlítsa össze képtömörítési szempontból a redundancia főbb típusait!
5. Mit jelent a tömörítési arány?
6. Az általános adattömörítéshez képest milyen speciális jellemzői vannak a képtömörítésnek?
7. Ismertesse a Huffman kódolás lényegét!
8. Ismertesse az aritmetikai kódolás lényegét!
9. Mit jelent a bitsík kódolás?
10. Mi a quad-tree kódolás lényege és milyen képekre alkalmazható hatékonyan?
11. Mi a különbség az 1D és a 2D futamhossz-kódolás között?
12. Mi a különbség az iránykódok és a differenciális iránykódok között?
13. Mi az előrebecsléses kódolás lényege és milyen képekre alkalmazható hatékonyan?
14. Mi a veszteséges tömörítési eljárások két fő típusa?
15. Melyik transzformációs kódolási eljárás optimális az együtthatók függetlensége szempontjából?
16. Mi a diszkrét koszinusz transzformáció előnye?
17. Ismertesse röviden a JPEG képtömörítés főbb lépéseit!
18. Hogyan alkalmazható a JPEG eljárás színes képekre?
19. Mit jelent a progresszív képfelépítés?
20. Mi a fraktál tömörítés lényege?
21. Mit tud elmondani a fraktál és JPEG tömörített képeket összehasonlítva?
22. Hogy épül fel az MPEG tömörítő eljárás?
23. Mi a különálló kép?
24. Mekkora lenne a  $800 \times 600 \times 24$  pixel<sup>2</sup>xbit felbontású videó, digitalizált változatának veszteségmentes adatfolyama / Mbyte/s-ban adjuk meg /?



## 9. TÁVÉRZÉKELÉS

### 9.1. Bevezetés

A könyv következő fejezeteiben a digitális képfeldolgozás alkalmazásaival foglalkozunk. Reménytelen vállalkozás lenne teljes körű áttekintésre törekedni, ezért csak a legnagyobb alkalmazási területekre térünk ki. Az egyes alkalmazási területeknél az adott területre jellemző feldolgozási módszereket, eszközöket és eljárásokat próbáljuk meg összefoglalni, és csak röviden utalunk konkrét alkalmazásokra. A képekkel, esetenként video részlettel illusztrált alkalmazási példákat a CD lemezen helyeztük el.

Az alkalmazási területek kiválasztása több szempontból sem volt egyszerű. Egyrészt mindenki nyilván a hozzá közelebb álló területet tartja fontosnak, és biztos sok olvasó csalódott lesz, hogy egyes alkalmazások kimaradtak. Másrészt, ha az alkalmazási területeket a mindennapi felhasználási területek oldaláról közelítjük azt tapasztaljuk, hogy különböző területeken lévő feladatokat képfeldolgozás oldalról teljesen azonos megközelítéssel lehet megoldani. (Például egy növénybetegség meghatározása mikroszkópi képfeldolgozó rendszerrel a mezőgazdasági alkalmazásokhoz sorolandó, de lényegében azonos rendszert igényel, mint egy cytologiai vizsgálat, ami viszont orvosi alkalmazási terület.)

A fenti okok miatt az alkalmazások felosztása biztos vitatható, de abban valószínűleg egyetértés van, hogy a távérzékeléssel kapcsolatos alkalmazások elválaszthatatlanok a digitális képfeldolgozástól, és a távérzékelés az elmúlt évtizedekben a digitális képfeldolgozás fejlődésének egyik alapvető ösztönzője, hajtó motorja volt, ezért indokolt, hogy az alkalmazások sorát ezzel a területtel kezdjük, és kicsit bővebben foglalkozunk a témakörrel.

**Távérzékelés** alkalmazásával két- vagy háromdimenziós objektumok vizsgálhatók úgy, hogy az érzékelő eszközök nincsenek közvetlen kapcsolatban a vizsgálat tárgyával. Bár ennek az általánosan elfogadott definíciónak a mikroszkópi képfeldolgozás is eleget tesz, általában ha távérzékelésről beszélünk, akkor az űr- és légifelvételekre gondolunk. A műholdakon elhelyezett szenzorok a földfelszín különböző tárgyairól visszavert, illetve kisugározott elektromágneses energiát érzékelik, így ezek az adatok bizonyos feldolgozások után információforrásként szolgálnak a vizsgált objektumról.

A távérzékelési módszerek szélesebb körű gyakorlati alkalmazásának kezdete a II. világháború idejére tehető. Ez akkor a légifelvételek vizuális és optikai kiértékelését jelentette. A hatvanas években a jelentős technológiai fejlődés lehetővé tette a mikrohullámú és termális infra eszközök nem csupán katonai alkalmazását is. Erre az időre tehető a számítógépek alkalmazásának kezdete a képkiértékelésben. Egy évtized eltel-

tével megjelentek a műholdakon elhelyezett érzékelők, és ezeket már széles körben alkalmazták polgári célokra is. Természetesen ezeket az eszközöket az űrkutatásban is használják

## 9.2. Távérzékeléssel kapcsolatos alapfogalmak

A távérzékelésben az **elektromágneses hullámokat leggyakrabban a hullámhosszukkal és az elektromágneses spektrumon belül elfoglalt helyükkel jellemezzük**. A hullámhossz jellemzésére a mikrométert ( $1 \mu\text{m} = 1 \times 10^{-6} \text{ m}$ ) használjuk.

Az elektromágneses (EM) spektrum egyes tartományai ( kozmikus, röntgen, ultraibolya, látható, infra, mikrohullám és rádió) nem választhatók szét élesen. A spektrumon belüli sugárzástípusok elkülönítése a szenzorok érzékenységének növekedésével egyre finomabbá vált.

A EM sugárzás 0,03 nm hullámhosszig terjedő tartománya a kozmikus és gamma tartomány. Ez fontos szerepet tölt be az Univerzumról alkotott elképzeléseink tisztázására. Gyakorlatilag a fizikusok és csillagászok „használják” csupán.

A röntgen sugárzást (0,03 nm-3,0 nm), amely a földi életre általában káros, a légkör szinte teljesen elnyeli. Az ultraibolya sugárzás a teljes spektrum látható részének kék-ibolya színű végétől a kisebb hullámhosszú tartomány felé található (0,03-0,4  $\mu\text{m}$ ). **Az emberi szem érzékenysége a 0,4 és 0,7  $\mu\text{m}$  közötti sávra terjed ki, ez a látható tartomány**. Ezen belül a kék 0,4-0,5  $\mu\text{m}$ , a zöld 0,5-0,6  $\mu\text{m}$ , a vörös 0,6-0,7  $\mu\text{m}$  hullámhosszú. Szemünk és a fotótechnikában használt negatív- és diafilm érzékenysége az ilyen erősségű és hullámhosszú energiára, vagyis a látható tartományra van hangolva. A látható fénynél hosszabb hullámú kibocsátott energia csak speciális érzékelők segítségével vizsgálható. A látható fény vörös végétől az infravörös tartomány három különböző tartománya található: közeli-infravörös (near-infrared-IR) 0,7  $\mu\text{m}$ -tól 1,3  $\mu\text{m}$ -ig, a középső infravörös (mid-infrared) 1,3-3  $\mu\text{m}$  és a hőtartomány (thermal-infrared) 3  $\mu\text{m}$ -100  $\mu\text{m}$ . A 3  $\mu\text{m}$ -es hullámhossz határvonal a visszavert és a kibocsátott infravörös sugarak között. Ez alatti hullámhosszakra a visszavert energia, míg föllette a kibocsátott energia detektálása a jellemző. A spektrum következő tartománya (0,1 mm - 30 cm) képezi mikrohullámú sugárzást. Azon túl található a rádió tartomány (30 cm-től nagyobb hullámhosszú sugárzás).

A legáltalánosabb érzékelő rendszerek a látható fény egy vagy több sávjával, infravörös sávokkal vagy mikrohullámokkal dolgoznak. A távérzékelés szempontjából nagyon fontos, hogy az infravörös tartományon belül csak a hőtartományú infravörös sáv alkalmas a hőmennyiség direkt érzékelésére a másik két sáv, a közeli- és a közép-infravörös sáv nem.

A földfelszín átlagos hőmérséklete kb. 300 °K (27°C). A Wien-féle törvény szerint

a maximális energiát a földfelszín a 9,7  $\mu\text{m}$ -es hullámhosszon bocsátja ki, ez a termális infravörös csúcs. A Nap energiacsúcsa 0,5  $\mu\text{m}$  hullámhossz közelében van.

A Wien-féle eltolódási törvény alapján, fekete test sugárzási görbéje az alábbi összefüggés szerint függ a fekete test hőmérsékletétől:

$$\lambda_{\max} T = A$$

ahol

$\lambda_{\max}$  = a maximális spektrális sugárzás hullámhossza  $\mu\text{m}$ -ben,

$T$  = abszolút hőmérséklet,  $^{\circ}\text{K}$ -ben

$A$  = 2898  $\mu\text{m}^{\circ}\text{K}$ .

Bizonyos szenzorok, pl. a radarok, saját energiaforrásukkal pásztázzák a vizsgált felszínt. Az ilyen típusú rendszereket „aktív”, a természetes sugárzást érzékelő rendszereket pedig „passzív” rendszereknek nevezzük. A villanófényt használó kamera aktív rendszernek, napfényt alkalmazó fényképezőgép passzív rendszernek minősül.

Az atmoszféra távérzékelésre gyakorolt hatása függ a sugár által megtett út hosszától, a sugárzás energiájának nagyságától, az atmoszféra összetételétől, a részecskék nagyságától és a hullámhossztól. **A távérzékelés szempontjából a szóródás és az elnyelés a két legfontosabb atmoszférikus hatás.**

Az atmoszférikus szóródás az atmoszférában megtalálható részecskék okozta hatás. Több típusa létezik, aszerint, hogy a részecske átmérője hogyan viszonyul a vele kölcsönhatásba lépő sugárzás hullámhosszához. Ha a részecskék (pl. nagyobb molekulák) átmérője sokkal kisebb, mint a sugárzás hullámhossza, a szóródást **Rayleigh-szóródás**nak nevezzük. Mivel a Rayleigh-szóródás hatása a hullámhossz negyedik hatványával fordítottan arányos, ezért sokkal intenzívebb a rövidebb hullámhosszak szóródása, mint a hosszabb hullámhosszaké.

A kék ég a Rayleigh-szóródás következménye. A napsugárzás kölcsönhatásba lép az atmoszférával, a rövidebb (kék) hullámhosszak erőteljesebben szóródnak, mint más látható hullámhosszak. Napkeltekor és napnyugtakor, a napsugarak sokkal hosszabb utat tesznek meg, mint napközben. A hosszabb út miatt, a rövidebb hullámhosszak szóródása és elnyelése szinte teljes, ezért csak a kevésbé szóródó, hosszabb hullámokat (narancs és vörös) látjuk.

A Rayleigh-szóródás következtében a műholdas fényképek homályosak. Ez abban nyilvánul meg, hogy csökken a kép élessége és romlik a kontrasztja. Színes fényképezésnél, légi- és űrfelvételeknél, különösen nagy magasságokban, kékes-szürke árnyalata lesz a képnek. Ezek kiszűrése megfelelő fényszűrő alkalmazásával lehetséges.

A másik típusú szóródás, az ún. **Mie-szóródás**, mely akkor lép fel, ha a kölcsönhatásba lépő részecske átmérője közel egyenlő a sugárzás hullámhosszával. A vízgőz és a por okoz Mie-szóródást. Hatása inkább a nagyobb hullámhosszak esetén érezhető. A Mie-szóródás már csekély felhőzet esetén is jelentős lehet.

Nem elhanyagolható szóródástípus a **nem szelektív szóródás**, mely akkor jön létre, ha a részecske átmérője sokkal nagyobb, mint a vele kölcsönhatásba lépő sugárzás hullámhossza. A vízcseppek átmérője általában 5 és 100  $\mu\text{m}$  között változik, így az átmérő nagyobb, mint a látható- és infravörös sugárzás hullámhossza. Ezért a szóródás nem szelektív a hullámhosszal kapcsolatban. A kék, a zöld és a vörös fényt teljesen egyenlő mértékben szórják a vízcseppek, ezért fehér színű a felhő és a köd.

A szóródással ellentétben, az atmoszférikus abszorpció valódi energia veszteséget jelent. A vízgőznek, a széndioxidnak és az ózonnak van a legjelentősebb abszorpciós hatása. Ezek a gázok az elektromágneses energiát egyes hullámhosszokban intenzíven vagy teljes egészében elnyelik. Így döntően befolyásolják azt, hogy mely spektrális sávokat alkalmazhatjuk az adott távérzékelő rendszerekben. Azokat a tartományokat, melyekben az atmoszféra részlegesen vagy teljes egészében átengedi az elektromágneses energiát „**atmoszférikus ablakoknak**” nevezzük.

atmoszférikus ablak	hullámhossz tartomány ( $\mu\text{m}$ )
1	0,3 – 1,3
2	1,5 – 1,8
3	2,0 – 2,6
4	3,0 – 3,6
5	4,2 – 5,0
6	7,0 – 15,0

Az emberi szem spektrális érzékenységi tartománya kapcsolatban van egy atmoszférikus ablakkal és a napsugárzás kibocsátási csúcsával. A földfelszín által kibocsátott hőenergia a 3 és 5  $\mu\text{m}$ , valamint a 8 és 14  $\mu\text{m}$  közötti atmoszférikus ablakokon figyelhető meg. Ezekben a tartományokban működő érzékelőket termális szkennereknek (thermal scanners) nevezzük. A többsávú érzékelők vagy multispektrális szkennerek (MSS) a látható fény és az infravörös tartomány több, vékony sávját egyidejűleg képesek vizsgálni. A radar és a passzív mikrohullámú rendszerek az 1 mm és az 30 cm közötti atmoszférikus ablak adta lehetőségeket használják ki.

**Az elektromágneses energia és a földfelszín találkozásakor, három alapvető energia-kölcsönhatást különböztetünk meg;** a beérkező energia egy része visszaverődhet, elnyelődhet vagy/és elvezetődik. Az energiamegmaradás elve alapján igaz, hogy:

$$E_b(\lambda) = E_r(\lambda) + E_a(\lambda) + E_t(\lambda)$$

ahol

- $E_b$  a beérkező energiamennyiség,
- $E_r$  a reflektált (visszavert) energia,
- $E_a$  az abszorbeált (elnyelt) energia,
- $E_t$  a továbbított energia.

Mindegyik komponens a  $\lambda$  hullámhossz függvénye.

A visszavert, az elnyelt és a továbbított energia aránya változik a földfelszín különböző tárgyain, függ az anyag típusától, állapotától. Lehetséges, hogy egy adott földfelszíni forma (pl. növényzet) másképpen reflektálja az energiát a különböző tartományokban. Ezért lehetséges az, hogy bizonyos növények teljesen hasonlóak egy spektrális tartományban, míg különbözőek egy másik sávban. A látható fény tartományában, ezeknek a spektrális variációknak az eredményei a színek. Akkor mondjuk, hogy egy tárgy kék, ha döntően a spektrum kék sávját veri vissza, zöld színű, ha a spektrum zöld sávját veri vissza és így tovább.

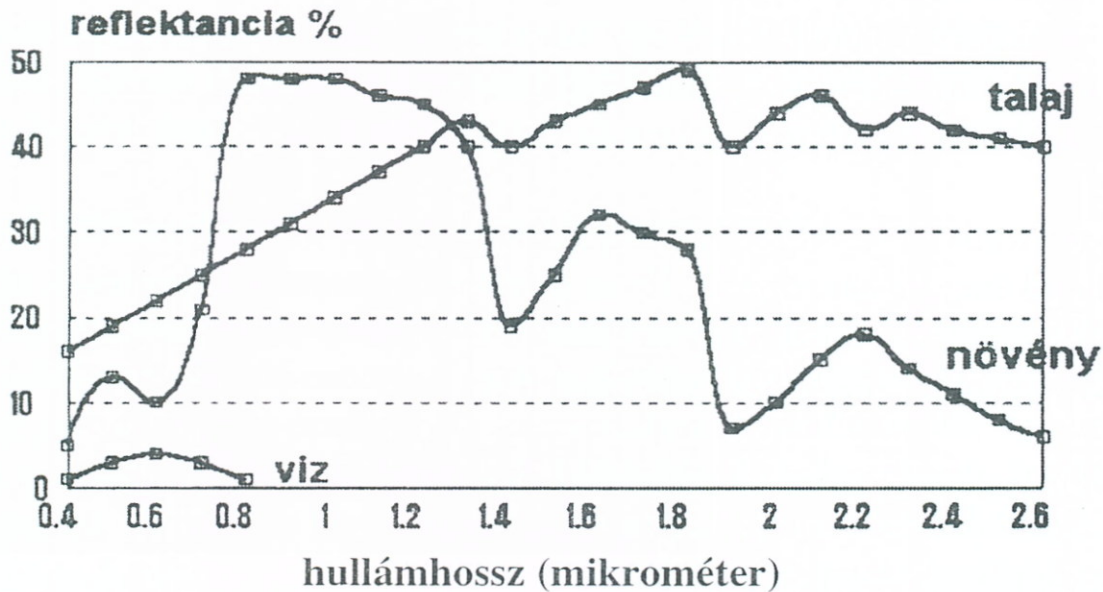
**Spektrális visszaverődési görbének** nevezzük egy objektumnak a hullámhossz függvényében kifejezett visszaverődési értékeit. Az egészséges zöld vegetáció, a talaj és a víz visszaverődési görbéje látható a 21. ábrán. A görbék ezeknek a felszíni formáknak a fő jellegzetességeit mutatják, de ezeken belül jelentős eltérések vannak.

A spektrum látható tartományában jelentkező minimum a növény leveleiben található színtestekhez (pigmentekhez) kötődik. A klorofill erősen elnyeli a 0,45 és 0,67  $\mu\text{m}$  közötti hullámhossz tartomány energiáját. A növény levelei erősen abszorbeálják a kék és a vörös fényt, míg a zöld tartományban sugárzott energiát nagyon erősen visszaverik, ezért a szemünk az egészséges vegetációt zöld színben látja.

Ha egy növény valamilyen betegségben, vagy káros hatástól szenved, akkor lelassul a normális növekedése, ami a klorofill tartalom csökkenésében is megnyilvánul. A kisebb klorofill mennyiség kisebb kék és vörös abszorpcióhoz vezet. Gyakran a vörös visszaverődés növekedése úgy jelentkezik, hogy a növény a zöld és a vörös színek kombinációjaként sárga színű lesz.

Az infravörös tartomány felé haladva, a 0,7  $\mu\text{m}$ -es hullámhossznál, az egészséges növényzet visszaverődése ugrásszerűen megnő. A 0,7 és 1,3  $\mu\text{m}$  közötti sávban, az egészséges növényzet a beérkezett energia 40-50%-át visszaveri. A maradék energia legnagyobb része elvezetődik, hiszen az abszorpció ebben a sávban minimális (kevesebb, mint 5 %). A növényzet visszaverő képessége a 0,7 és 1,3  $\mu\text{m}$  közötti sávban főleg levélzetének belső szerkezeti sajátosságaiból következik. Eltérő fajoknál a belső szerkezet nagyon különböző is lehet. Így a visszaverődés mérése lehetőséget ad a fajták elkülönítésére még akkor is, ha a látható fényben ezek a fajták nagyon hasonlítanak egymásra.

# Reflektancia görbék



21. ábra

Hasonlóképpen, ezen sávban működő szenzorokat használhatunk az egyes fajták betegségeinek kimutatására is, hiszen a betegség is megváltoztatja a visszaverődési tulajdonságokat.

A méréskor zavaró lehet a növény levélzetének rétegzettsége, amely lehetővé teszi a többszörös visszaverést és elvezetést. Ezért az infravörös visszaverődés növekszik a korona levélrétegeinek a számával.

1,3 μm-nél nagyobb hullámhosszúságú energia legnagyobb része elnyelődik, vagy visszaverődik a növényzeten, kevés az energiaelvezetés. Az 1,4 és a 2,7 μm-nél előforduló minimumok azért alakulnak ki, mert a levelek magas víztartalma elnyeli ezeket a hullámhosszakokat. Ezeket a sávokat „vízelnyelési sávoknak” nevezzük. Visszaverődési csúcsokat találunk az 1,6 μm és a 2,2 μm-es sávban, az elnyelési sávok között. Az 1,3 μm-nél nagyobb hullámhosszakra igaz, hogy a levél visszaverődése hozzávetőlegesen fordítottan arányos a levél teljes víztartalmával. A teljes víztartalom pedig a növény nedvességtartalmától és a levél vastagságától függ.

A talaj reflexiós képességét a nedvességtartalma, a szerkezete (fizikai összetétele, a homok, iszap és az agyag aránya, stb.), a felszín egyenetlensége, vasoxid jelenléte és a szerves anyag tartalma határozza meg. Ennek megfelelően a különböző talajfajták reflektancia görbéje is eltérő.

A talaj nedvességtartalma szoros kapcsolatban van a talaj fizikai összetételével. A kavicsos, homokos talajok rendszerint jól vezetik a vizet. Ennek következménye a homoktalajok alacsony víztartalma és relatíve magas visszaverése, ugyanakkor a rossz

vízvezető talajok (agyagos, iszapos talajok) általában alacsony visszaverőképességűek. A talajnedvesség hiányában a talaj szerkezete döntő a visszaverődés szempontjából és a kavicsos, durvább szemcseösszetételű talajok sötétebbnek látszanak, mint a finomabb szerkezetű talajok. Ezért egy talajtípus visszaverő képessége csak bizonyos feltételek mellett marad változatlan. A visszaverődés mértékét a talajfelszín egyenetlensége, a talaj szervesanyag tartalma és vasoxid tartalma szintén jelentősen befolyásolja.

A víz spektrális visszaverődési tulajdonságai közül a legjellemzőbb az infravörös sávban fellépő energia abszorpció. Ez az infravörös abszorpció nemcsak nyílt vízfelületek (óceán, tengerek, tavak, folyók) esetén figyelhető meg, hanem a növényzetben és a talajban jelenlévő víz is hasonló elnyelést okoz.

### 9.3. Az EM sugárzást detektáló berendezések

A képfeldolgozás egyik legnagyobb alkalmazási területe az űrkutatás és a távérzékeléssel nyert adatok feldolgozása. Ezeken a területeken gyakran nem a látható fényt érzékelő eszközöket, hanem egyéb speciális rendszereket alkalmaznak. Ezek az alkalmazások a hagyományos képfeldolgozástól többnyire csak az érzékelés fizikai elvében térnek el, bár vannak olyan területek is, melyek teljesen új problémákat vetnek fel. A képérzékelés, képalkotás többnyire az adott kölcsönhatás erősségének változásain alapszik (pl.: röntgen-, illetve elektronmikroszkóp-felvételek készítése esetén), de vannak olyan területek is, ahol a kölcsönhatás időbeli lefolyása szolgáltat adatokat (pl.: ultrahang képek készítése, radar, sonar), illetve ahol a fázisviszonyoknak van nagy szerepük (pl. hologramok készítése).

A következőkben néhány példán keresztül mutatjuk be, hogy a távérzékelés esetén (de természetesen más képfeldolgozási feladatnál) milyen jelforrások fordulhatnak elő, és azok milyen eltérő feladatok megoldására is alkalmazhatók.

#### Hőkamerás képérzékelő rendszerek

A képérzékelő eszköz az elektromágneses hullámokat az infravörös tartományban érzékeli, ami lényegében a hőszugárzást jelenti. A szokásos működési tartományok: 2-5  $\mu\text{m}$ , illetve: 10-12  $\mu\text{m}$ . A hőfelvételt az érzékelő felületére speciális optikával képezik le, amely e fenti hullámhossz tartományban is megfelelően működik. (A hagyományos objektívek ugyanis ebben a tartományban nem használhatók.) Az érzékelőt a kelő jel/zaj viszony elérése érdekében hűteni kell. Erre a célra pl. folyékony nitrogént használnak. A hőkamerák kimenőjele - az eltérő képalkotási frekvencia következtében

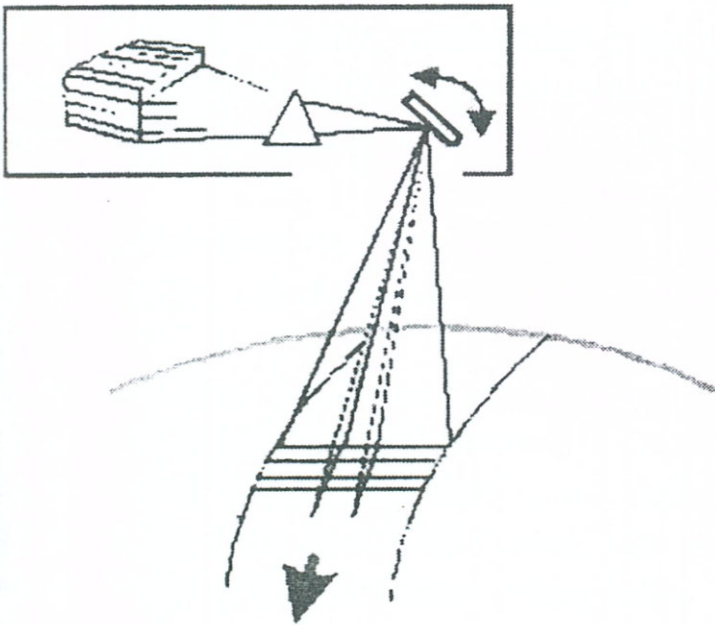
-nem a szokásos videojel szabványok valamelyikével történik, így számítógépbe vitelükhöz speciális illesztőkártyára van szükség.

A hőkamerás képérzékelő drága eszköz, és többek között a hűtési igény miatt felhasználási területe is viszonylag korlátos. A távérzékelésen túlmenően a leggyakoribb felhasználási területek: hőszigetelés, felmelegedés képi ellenőrzése; orvosi diagnózis elősegítése; égési folyamatok analízise.

### Multispektrális képérzékelő rendszerek (szkennerek)

A műholdakon elhelyezett képérzékelő eszközök egyik igen fontos típusa a **multi-spektrális képérzékelő** (*szkenner*). Ennek legfontosabb jellemzője az, hogy nem egy hullámhossz tartományban érzékeli az elektromágneses hullámokat, hanem több tartományban is rendelkezik érzékelőkkel. E tartományok nem csupán a látható tartományt, hanem pl. a közeli és távoli infravörös tartományt is tartalmazhatják. Az egyik korai műhold, a **LANDSAT-D** ún. **MSS** (multispectral scanner) érzékelő rendszere a következő sávokban működött:

1. sáv: 0,5 – 0,6  $\mu\text{m}$
2. sáv: 0,6 – 0,7  $\mu\text{m}$
3. sáv: 0,7 – 0,8  $\mu\text{m}$
4. sáv: 0,8 – 1,1  $\mu\text{m}$



22. ábra

A 22. ábra az **MSS**-képképzés lényegét mutatja: a műhold kering a Föld körül, a haladás irányára merőleges irányban pedig egy tükör rezgő mozgást végez a műhold fedélzetén. E két mozgás eredményezi a felszín sorsor-sorra való leképezését: a rendszer folyamatosan tapogatja a Föld felszínéről visszavert elektromágneses hullámokat, s a kiértékelő központba sugározza a nyert adatokat. A **LANDSAT-D** esetében a látómező a Föld felszínén kb. 185 km széles; a rendszer elektronikai megfontolásokból egyidejűleg, sávonként 6 képsort képez le, azaz 24 érzékelővel dolgozik.



Az alábbiakban röviden kitérünk napjaink korszerű és széles körben alkalmazott űrérzékelőinek alapvető jellemzőire. Az érzékelők által készített felvételeket találunk a CD IMAGES/MUHOLD könyvtárában.

A **Landsat 5 TM** (Thematic Mapper) elnevezésű érzékelő 7 sávban készít felvételeket, amelyből 6 sáv 30 m-es, míg a termális-infra tartományban készített felvétel 120 m-es terepi felbontású (CD – IMAGES\MUHOLD\LANDSAT).

Kék	0.45 - 0.52 $\mu\text{m}$
Zöld	0.52 - 0.60 $\mu\text{m}$
Piros	0.63 - 0.69 $\mu\text{m}$
Közeli infra	0.76 - 0.90 $\mu\text{m}$
Középső infra	1.55 - 1.75 $\mu\text{m}$
Középső infra	2.08 - 2.35 $\mu\text{m}$
Termális infra	10.4 - 12.4 $\mu\text{m}$

A **napszinkron** (a pálya változásai kiegyenlítik a Földnek a Nap körüli keringésből adódó változásait, azaz a műhold egy adott hely fölött mindig azonos helyi időben halad el), 705 km magasan keringő műhold által készített felvétel 185x170 km-es területről ad képi adatokat, melyeket 16 napos gyakorisággal készíthet ugyanarról a területről. Adatai a legszélesebb alkalmazásokhoz nyújtanak információt kezdve a vegetáció térképezéstől a környezetvédelmi alkalmazásokon keresztül egészen a sarki megfigyelésekig.

Az **IRS** (Indian Remote Sensing Satellite) család három műholdból (IRS 1A, 1B és 1C) áll (CD – IMAGES\MUHOLD\IRS). Az egyes eszközöket többféle érzékelővel is felszerelték (LISS-I, LISS-IIA, LISS-IIB, LISS-III, Pancromatic, WiFS), melyek adatait az alábbi táblázatban foglaltuk össze:

Jellemző\érzékelő	<b>IRS-1A</b> LISS-I és LISS-II	<b>IRS-1B</b> LISS-I és LISS-II	<b>IRS-1C</b> LISS-III, Pancromatic, WiFS
Érzékelési tartomány	<b>Kék</b> 0.45 - 0.52 $\mu\text{m}$ <b>Zöld</b> 0.52 - 0.59 $\mu\text{m}$ <b>Piros</b> 0.62 - 0.68 $\mu\text{m}$ <b>Közeli infra</b> 0.77 - 0.86 $\mu\text{m}$	<b>Kék</b> 0.45 - 0.52 $\mu\text{m}$ <b>Zöld</b> 0.52 - 0.59 $\mu\text{m}$ <b>Piros</b> 0.62 - 0.68 $\mu\text{m}$ <b>Közeli infra</b> 0.77 - 0.86 $\mu\text{m}$	<b>Pankromatikus</b> 0.5 - 0.75 $\mu\text{m}$ <b>LISS-III</b> <b>Zöld</b> 0.52 - 0.59 $\mu\text{m}$ <b>Piros</b> 0.62 - 0.68 $\mu\text{m}$ <b>Közeli infra</b> 0.77 - 0.86 $\mu\text{m}$ <b>Középső infra</b> 1.55-1.70 $\mu\text{m}$ <b>WiFS</b> <b>Vörös</b> 0.62-0.68 $\mu\text{m}$ <b>Közeli infra</b> 0.77-0.86 $\mu\text{m}$
Felbontás	LISS-I 72,5m LISS-II 36,25 m	LISS-I 72,5m LISS-II 36,25 m	Pan 5,8 m LISS-III 23 m Középső infra 70 m WiFS 188 m
Felvételkészítés gyakorisága (ugyanarról a területről ugyanazon látószögből készül felvétel)	22 naponként	22 naponként	Pan és LISS-III 24 naponként  WiFS 5 naponként
Képméret	<b>LISS-I</b> 148 km széles sávban <b>LISS-II</b> 146 km széles sávban	<b>LISS-I</b> 148 km sávban <b>LISS-II</b> 146 km sávban	<b>Pan</b> 70 km sávban <b>LISS-III</b> 142 km sávban <b>LISS-III</b> középső infra 148 km sávban <b>WiFS</b> 774 km sávban

Adatait elsősorban vegetáció térképezésre és földhasználat felmérésére használják. Mind globális, mind lokális vizsgálatokhoz kiváló.

A 798 km magasan napszinkron pályán keringő **RADARSAT** (CD – IMAGES\MUHOLD\SAR) műhold érzékelője az 5,6 cm-es hullámhosszban érzékel. Az elkészített felvétel felbontása az érzékelőfej beállításától függ: 8 m, 25 m, 30 m, 35 m, 50 m, 100 m. Az elkészített kép mérete is ennek megfelelően 50x50 km-től 500x500 km-ig terjed. Adatai térképészeti, urbanisztikai, mezőgazdasági termésbecsési, geológiai valamint sztereófelvételek készítésére is használhatók.

Az **IKONOS** műhold pankromatikus ( $0.45 - 0.90 \mu\text{m}$ ) felvételei 1-1,5 m-es (pályán történő pozíciótól függően), multispektrális (érzékelési tartományuk azonos a LANDSAT TM első 4 sávjával) képei 4 m-es terepi felbontásúak. A rendkívül nagy terepi felbontást tovább javítja az is, hogy a multispektrális adatokkal „kiegészíthetők” a pankromatikus képi adatok, ezáltal nagyfelbontású „színezett” képet kapunk. Ilyen felvétel a: IMAGES\MUHOLD\IKONOS\IKONOS\_pan\_szinezett\_becs\_1m.tif. Az érzékelő 681 km magas, napszinkron pályán  $13 \times 13$  km-es állóképeket vagy  $11 \times 100$  km-es sávól egészen  $11 \times 1000$  km-es sávig készíthet vizuális adatokat. Alkalmazási területei közül kiemelt szerepet kapott a vegetáció pontos feltérképezése, a felszíni mesterséges objektumok és mozgások (épületek és egyéb mesterséges objektumok - áruszállítás, tömegközlekedés) valamint a turizmus környezetre gyakorolt hatásának nyomon követése.

A **SPOT** (CD – IMAGES\MUHOLD\SPOT) 832 km magas, napszinkron pályán 26 naponkénti gyakorisággal készíthet felvételeket. Jelenleg a műhold a SPOT-4 elnevezésű érzékelővel dolgozik, amely pankromatikus ( $0.51 - 0.73 \mu\text{m}$ ) 10 m terepi felbontású és multispektrális ( $0.50 - 0.59 \mu\text{m}$ ;  $0.61 - 0.68 \mu\text{m}$ ;  $0.79 - 0.89 \mu\text{m}$ ) 20 m terepi felbontású adatokat szolgáltat. Rendkívül megbízhatóan dolgozik a rendszer, amelyet az is bizonyít, hogy a SPOT-1 érzékelő a tervezett élettartama (2,5 év) helyett 4 évig dolgozott, közben közel 3 millió felvételt készített. A jelenleg működő rendszer naponta több száz felvételt készít! Adatait egy - Interneten is elérhető - adatbázisban tárolják, amelyből a felhasználó bármikor kikeresheti a számára megfelelő felvételeket. Alkalmazási területei talán az egyik legszélesebb palettát tükrözik, így hazánkban is kedvelt. A térképészet, a geológia, a mezőgazdaság, az erdészet, a környezetvédelem, stb. csak néhány kiragadott példa. A felhasználó a CD lemezen talál egy a Kis-Balaton térségét szemléltető eredeti 1994-es SPOT felvétel ( $3000 \times 3200$  pixel<sup>2</sup>) alapján készített képrészletet ( $550 \times 550$  pixel<sup>2</sup> – sp1994ms.tif file), amely egy erdőkárosodás felmérését elvégző nemzetközi munka alapját képezte.

## Ultrahang kamerák

A képképzésre nagy frekvenciájú mechanikai rezgések terjedését, visszaverődését felhasználó eszközöket **ultrahang kamerának** nevezzük. A működés alapelve a felhasználási területtől függetlenül ugyanaz: a kamera meghatározott időpillanatban meghatározott erősségű, hullámhosszúságú, irányú és időtartamú ultrahanghullámot bocsát ki, majd a kibocsátás megszüntetése után vevőjével érzékeli a visszavert hullámok erősségét. Arra, hogy a visszaverődések milyen távolságban levő térfogatelemekből történtek, az eltelt időből, arra pedig, hogy a visszaverő határréteg milyen anyag-

szerkezet változást takar, a jel erősségéből következtet a készülék. Mivel a nagyfrekvenciájú hanghullámok kellően jól fókuszálhatók, a vizsgált térrész pontról-pontra letapogatható. Az eredő kép elsősorban a vizsgált térrészben észlelt sűrűségváltozások helyének lokalizálására alkalmas, de kiegészítő műveletekkel sűrűségeloszlási kép is készíthető. A működés fontos feltétele a rezgéseket jól vezető, a hullámokat kevésbé szétszóró közeg; ezért a módszer jól használható víz alatt, illetve szilárd testekben, de gyenge eredményekre vezet levegőben.

## Radár

A képképzés a kibocsátott mikrohullámú (**azaz néhány centiméter hullámhosszúságú**) sugarak visszaverődésének elemzésén alapul. A radar impulzusüzemű eszköz, s a következő három lépést ismétli gyors egymásutánban:

1. Nagy energiájú mikrohullámú sugárnyalábot bocsát ki a letapogatni kívánt irányba. A sugár fókuszálását paraboloid tükör biztosíthatja.
2. A visszavert (igen gyenge) jeleket felfogja és elemzi. (A folyamat időbeli lefolyása a visszaverő felületek távolságáról, a jelek erőssége pedig az objektumok nagyságáról, fémességéről hordoz információt.)
3. A sugárkibocsátó illetve jelérzékelő antennát pozicionálja (lépteti) a következő irányba.

Megjegyzendő, hogy a korábbi mechanikus kivitelű antennákat - főleg a hadiiparban - felváltják a sok kis antennából felépített ún. fázistolásos antennák.

## 9.4. Távérzékeléssel kapcsolatos alkalmazások

A távérzékeléssel kapcsolatos alkalmazásoknak négy nagyon fontos jellemzője van, nevezetesen:

- multispektrális képek alkalmazása,
- multitemporális feldolgozások,
- hatalmas adat mennyiség,
- egyéb információs rendszerekhez való kapcsolódás.

A **multispektrális képek** alkalmazása csak a feldolgozás nagyon tudatos előkészítésével lehet eredményes: A felhasználónak alaposan ismernie kell a vizsgálni kívánt jelenséget, ugyanis a rendelkezésre álló képsávok közül a megfelelők kiválasztása meghatározza a feldolgozás eredményességét. Más hullámhossz tartományban készített felvételek alkalmasak a vegetáció vizsgálatára, mint a geológiai vagy vízügyi vizs-

gálatokra, de az évszak is befolyásolja milyen felvételeket érdemes alkalmazni. Nagyon sok spektrális sáv egyidejű alkalmazása megsokszorozza a feldolgozandó adatmennyiséget és nem ad többletinformációt. Nem azért készülnek több hullámhossz tartományban a felvételek, hogy mindig mindegyiket egyidejűleg felhasználjuk, hanem azért, hogy minden feladathoz a legmegfelelőbbet tudjuk alkalmazni. Általában két-három, maximum öt-hat sáv egyidejű alkalmazása elegendő egy adott feladatnál.

A vizsgálatok célja többnyire a változások nyomonkövetése, ezért a felvételek értékelésekor azonos területről különböző időpontban készített felvételeket használnak, ezt nevezik **multitemporális** feldolgozásnak.

Ha az előbbieken túlmenően figyelembe vesszük a mai szenzorok felbontását (a Föld felszínéről ma már 1 méter alatti felbontással is készülnek felvételek, és egy-egy felvétel 10000 km<sup>2</sup>-nél is nagyobb területet fedhet le, akkor nem szorul különösebb magyarázatra, hogy a távérzékelési alkalmazások egyik jellemzője a hatalmas adatmennyiség. Egy-egy távérzékeléssel kapcsolatos feladat megoldása a vizsgált terület nagyságától, a sávok számától és a felbontástól függően néhányszor tíz Megabájtól több száz vagy ezer Megabájtig terjedő képanyag feldolgozását jelenti.

A más információs rendszerekhez való kapcsolódás alapvetően két dolgot takar. Az egyik a távérzékeléssel kapott **képek transzformálása** valamilyen elfogadott térképi ábrázolásnak megfelelően, hiszen a kapott információkat, eredményeket csak így tudjuk felhasználni. A másik, hogy a távérzékeléssel kapott adatok kiértékeléséhez általában szükség van olyan **referenciaadatokra**, amelyek alapján az egyes spektrális reflektancia értékekhez hozzárendelhetők az adott felvételi körülmények között a tényleges felszíni jelenségek. Ezek az adatok a mintaterületeken végzett felszíni mérések eredményei.

A fentiek alapján elmondható, hogy a távérzékeléssel kapcsolatos feladatokban a digitális képfeldolgozás szinte teljes fegyvertárát felhasználják, de a különböző képjavítási, multispektális osztályozási és geometriai transzformációs eljárások kiemelt szerepet kapnak ezekben az alkalmazásokban. A feladatok sokrétűségének jellemzésére az alábbiakban csak vázaltszerűen felsoroljuk az egyik legismertebb, távérzékelési feladatokra kifejlesztett képfeldolgozó rendszer az ERDAS IMAGINE funkciókészletét.

### Alap modulkészlet

**Lekérdező:** megjeleníteni, lekérdezni, kiemelni és feliratozni képes egy vagy több réteget a lekérdező ablakban.

**Térképszerkesztő:** térképek, szemléltető grafikák előállítására egy vagy több kép felhasználásával, feliratok készítése, térképkeret, lépték, jelmagyarázat készítése.

**Állománykezelő:** direkt hozzáférés sokféle egymástól különböző raszteres adattípushoz.

**Segédprogramok:** a képekhez kapcsolódó statisztikák megjelenítése, vetületi rendszerekről információ és a térképpel kapcsolatos egyéb adatok nyilvántartása, statisztikák készítése és frissítése.

**Képküigazító:** a képek transzformációja földrajzi koordinátákkal rendelkező térképpé, vagy egy kép egy másik képhez való hozzárendelése.

**Képelemző:** sokfajta, képfeldolgozási és térinformatikai funkcióhoz ad gyors hozzáférést, mint: spektrális és terepi képminőség fokozó eljárások, osztályozás, Fourier analízis, topográfiai analízis, térinformatika elemzés és számos segédprogram.

### **Kiegészítő modulok**

**Képkatalógus:** a képkönyvtárat és az információ kezelő rendszert segíti.

**Vektor:** raszteres és vektoros feldolgozásokhoz, ARC/INFO adatmodell.

**Radar:** radarkép feldolgozó, számos radar formátumot támogat.

**Perspektivikus kép készítő:** a domborzat perspektivikus képének előállítása, multispektrális, vagy pankromatikus kép alapján.

**IMAGINE OrthoMAX:** „polcra levehető” hardveren futtatható, precíz ortofotó térképeket, űr- vagy légifelvételekből előállító rendszer.

**Fejlesztői eszköztár:** C programozói eszköztár.

**IMAGINE helyreállítás:** minták összehasonlításán alapuló, (resampling) modul, ami éles, kristálytiszta képet eredményez.

**VirtualGIS:** A 3-D-s megjelenítő. Egyszerre több réteg, vektor, annotáció, tematikus vagy nem tematikus adatok teríthetők rá szőnyegszerűen a 3D-s terepre.

A továbbiakban felsoroljuk a távérzékeléshez kapcsolódó legfontosabb alkalmazási területeket (a CD-n széleskörű képanyag szemlélteti az alábbiakat):

**Mezőgazdaság és erdőszet:** A vegetáció, illetve művelt területek és erdőfajták egyes típusainak megkülönböztetése. Termésbecslés, fakitermelés nyomon követése, biomassa kalkuláció, a növényzet életképessége, illetve betegsége, talajok állapota, talajtársulások, erdő ill. legelőtüzek károsodásának felmérése, védett állatok élőhelyeinek térképezése.

**Területfelhasználás:** területfelhasználás alapú osztályozás, térképezés és térképfrissítés, területi potenciál meghatározása, települések növekedésének nyomon követése, regionális tervezés, közlekedési hálózat térképezése, vízfelületek határainak térképezése, árterek gazdálkodási kérdései, új közlekedési útvonalak és infrastruktúrális hálózati elemek optimális helyének megállapítása.

**Geológia:** nagyobb geológiai egységek térképezése, geológiai térképek felülvizsgálata, bizonyos kőzettípusok felismerése, laza kőzetek és talajok azonosítása, vulkanikus kőzetek térképezése, tereptárgyak térképezése, bányászati lehetőségek kutatása

a felszíni jegyek alapján, regionális struktúrák meghatározása, vonalas jellegű formák térképezése.

**Vízgazdálkodás:** vízfelületek határainak megállapítása, felszíni vizek térképezése, árvizek és árterületek térképezése, hó és jég fedettségi térkép készítése, gleccserek jellemzőinek mérése, üledék lerakódás és szennyeződés terjedési minták ábrázolása, öntözött területek lehatárolása, tavak és vizenyős területek számbavétele, a hó olvadásból származó belvívveszély becslése.

**Partmenti területek védelme:** a szennyeződés terjedési mintái és a vízmozgás meghatározása, partvonalak mozgásának térképezése, zátonyok és sekély partmenti vizek felmérése, jégborítottság térképezése a hajózás számára, vízpartok eróziójának térképezése, környezet szennyezések, olajkiömlések nyomkövetése, vízmélység mérése.

**Környezetvédelem:** a felszíni bányászat, illetve a tájképi helyreállítás monitoring jellegű figyelése, vízszennyezések térképezése ill. monitoringja, természeti katasztrófák hatásainak megállapítása, az emberi tevékenység környezeti hatásainak monitoringja, aszálykárok felmérése, szilárd hulladék lerakóhelyek kijelölése, erőművek ill. más gyártelepek helyének kiválasztása.

**Meteorológia:** rövid és középtávú előrejelzések, frontok, felhőzetek mozgásának megfigyelése.

## Ellenőrző kérdések

1. Mit nevezünk távérzékelésnek?
2. Milyen tartományokra oszlik az EM spektrum?
3. Mire jellemző a Wien-féle eltolódási törvény?
4. Ismertesse a legfontosabb atmoszférikus hatásokat!
5. Milyen hatása van a műholdképekre a Rayleigh-szórásnak?
6. Milyen hatása van a műholdképekre a Mie-szórásnak?
7. Milyen összefüggés írja le az EM energia és földfelszín kölcsönhatását?
8. Mit nevezünk reflektancia görbének?
9. Miért különíthetők el az egészséges ill. beteg növénykultúrák?
10. Hogyan befolyásolja a reflektanciát a növények víztartalma?
11. Mi befolyásolja a talaj visszaverő képességét?
12. Ismertesse a leggyakoribb sugárzás detektáló rendszerek jellemzőit!
13. Láthatók-e ultrahangos képalkotással a színek?
14. Mit jelent a multitemporális feldolgozás?
15. Milyen képfeldolgozási és térinformatikai elemeket tartalmazhat egy képfeldolgozó szoftver rendszer (pl. ERDAS IMAGINE)?
16. Milyen globális feladatok megoldására használható a távérzékelés?

## 10. ORVOSI ALKALMAZÁSOK

### 10.1. Bevezetés

A képfeldolgozás kialakulása és fejlődése szorosan összefügg az orvosi alkalmazásokkal. Már az 1970-es években megjelentek az első rendszerek különböző mikroszkópi képkieértékelési feladatokra. Az orvosi alkalmazások célja részben különböző orvosi képkalkotó rendszerek képeinek javítása, részben a nagytömegű laboratóriumi (mikroszkópi) vizsgálatok elvégzésének automatizálása. Természetesen mindkét vonatkozásban a kutatás is jelentős szerepet kap. Az elmúlt időszakban egyre inkább előtérbe kerül a távdiagnózis, illetve a tele-konzílium igénye, amelynek során képek átvitelére is sor kerülhet, és az adott szakterület specialistájával való konzultáció lehetővé válik időrabló utazás nélkül.

Képfeldolgozás szempontjából tehát ugyanazok a feladatok jelentkeznek, mint más alkalmazásoknál, a kép minőségének javítása, archiválás és tömörítés, szegmentálás és alakfelismerés, osztályozás, de néhány speciális igény és adottság mindenképpen említésre érdemes.

### 10.2. Képkalkotó berendezések képeinek javítása

Az évtizedek óta alkalmazott röntgenfelvételek mellett ma már számos olyan berendezés segíti a diagnózis felállítását és a gyógyítást, amelyek valamilyen elektromágneses vagy igen gyenge radioaktív sugárzást bocsátanak ki, és az ember belső szerveiről visszaverődő jelekből képet alkotnak. Ilyen például a Computer Tomográf, a különböző ultrahangos berendezések, a szcintillációs kamera, az NMR berendezés. Ezeknél a berendezéseknél az első feladat a képrekonstrukció, ami szorosan összefügg magának a berendezésnek a működési elvével, így ezt a feladatot maga a műszergyártó oldja meg, és általában a megoldást nem publikálja, más képfeldolgozó berendezés hozzáillesztését nem támogatja. Megemlítjük, hogy a röntgen berendezéseknél képjavításra gyakran használt eljárás a képek időben történő átlagolása, mivel a képforrás jele gyenge a jelentkező zajhoz viszonyítva, viszont a zaj időben közel 0 átlagú.

A különböző képkalkotó rendszereknél - a képrekonstrukciótól eltekintve - alapvető feladat a kép minőségének javítása és a képek tárolása. A kép az esetek többségében monokróm, színek legfeljebb álszínes megjelenítésként, egyes jelenségek kiemelésére szolgálnak. Az értékelés a szürkeségi fokozatokban való eltéréseken alapul, ezért a különböző kontrasztfokozó eljárásoknak jelentős szerepe van. Tekintettel arra, hogy a kieértékelés szempontjából érdekes részek denzitás értékei a környezetüktől ál-



talában igen kis mértékben különböznek, és az orvosi szempontból érdekes eltérések igen nehezen algoritmizálhatók, ezért automatikus értékelő rendszerek ezen a területen még nem léteznek, a képfeldolgozás feladata az értékelő számára a kiértékelés megkönnyítése. Jelentős szerepe van még a méréseknek, ugyanis objektív adatok kaphatók az ember belső szerveiről (pl. ultrahangos vizsgálatok). Ezek a mérések sem automatikusak azonban, hanem a vizsgálatot végző orvos a kurzorral jelöli ki a mérendő távolságot vagy átmérőt.

A képtárolásnak az a jelentősége, hogy a korábbi leletek (képek) összehasonlításával a páciens egészségi állapota nyomon követhető. Természetesen itt nagy szerep jut a képtömörítési eljárásoknak, ugyanis egy-egy kép tárolása önmagában is jelentős területet igényel. Egy mellkasi röntgenfelvétel például tömörítés nélkül 3-4 Mbyte, vagy több helyet is elfoglalhat a digitalizálás módjától függően. Itt említjük meg, hogy a röntgenfelvételeket előszeretettel képpontonkénti 12 bitre digitalizálják, éppen a finomabb gradációs felbontás miatt. A CD - IMAGES/FORMATUM könyvtárban több olyan kép is található, amely sávonként 16 bites. Mindig nehéz döntés, hogy minőségromlást megengedjünk-e vagy sem a tömörítés során. Nyilván a későbbi felhasználás módjától függ, melyik utat válasszuk. Ma sok esetben az archiválás - hagyományos módon - filmen, vagy ultrahangos berendezéseknél videó kazettán történik. Ezeknek az anyagoknak az utólagos digitalizálása horribilis költségeket jelenthet.

Talán itt érdemes megemlíteni, hogy a különböző orvosi képalkotó berendezések képeinek egységes kezelése komoly kihívást jelent. Gondoljunk arra, hogy ma már egyre inkább általánossá válik egy-egy kórházon belül integrált szoftverrendszerek alkalmazása számítógépes hálózaton. Ezekben a rendszerekben a különböző helyeken és különböző műszereken keletkező képeket is egységesen kell kezelni. Megoldást az elmúlt években egyre jobban elterjedt és elfogadottá vált - ma már szabványnak tekinthető - DICOM (Digital Imaging and Communications in Medicine) előírások és ajánlások alkalmazása jelentett. A DICOM olyan felületeket és protokollokat ír elő, amelyek betartásával a legkülönbözőbb berendezések együtt tudnak működni, beleértve a képek átadását és fogadását is.

### 10.3. Kéпкиértékelési feladatok

A mikroszkóp megjelenésétől fogva az orvosok vizsgálati eszköze. Bátran elmondhatjuk, hogy a mikroszkóp kifejlesztésében jelentős szerepe volt az orvosi felhasználás igényeinek. A mikroszkópi vizsgálatok lehetőségeit megsokszorozhatjuk a számítógépes képfeldolgozás segítségével, de nem mikroszkópi képek értékelése is hasonló képfeldolgozási feladatok megoldását igényli.

Itt érdemes felhívni a figyelmet a mikroszkóp és a videokamera vagy digitális fény-

képezőgép korrekt illesztésének fontosságára. Ahhoz, hogy a mikroszkóp képét számítógéppel feldolgozzuk, a mikroszkópnak kamera csatlakozási lehetőséggel kell rendelkeznie. Ugyanazt az optikai megoldást kell használni mint a mikroszkópi kép fényképezésénél, de tisztában kell lenni azzal, hogy a hagyományos fényképezőgéphez, a különböző videokamerákhoz és digitális fényképezőgépekhez különböző közdarabok szükségesek. Ez abból adódik, hogy a fényképezés esetén jóval nagyobb felületre kell leképezni a mikroszkóp alatt lévő tárgyat mint egy CCD kamera esetén. A szokásos filmméret 36x24 mm, míg például az 1/2" méretű CCD 6,4x4,8 mm. Videokamerák közül minimum 1/2"-es CCD-vel rendelkező kamerát érdemes választani, de igényesebb megoldásoknál 2/3" vagy akár 1" méretű CCD-vel rendelkező kamerákat is használnak.

A mikroszkópgyártók opcióként biztosítják a legkülönfélébb csatlakozási lehetőségeket, de a felhasználók sokszor anyagi megfontolásból egyedi megoldásokat keresnek, ami nem egyszer csalódást okoz, ha nem megfelelő szakértelemmel tervezik meg a közdarabot.

Gondos mérlegelést igényel a videokamera és a digitális fényképezőgép közötti választás is. Általánosságban elmondható, hogy a digitális kamerák felbontása ma már meghaladja a videokamera és digitalizáló kártya kombinációval elérhető felbontást, viszont a képbeviteli és a feldolgozási idő a videokamera esetén rövidebb.

### A képfeldolgozás különböző szintjei

Az orvosi alkalmazások során a következő lehetőségeket nyújtja a képfeldolgozás:

- Képszerkesztés,
- Lényeges részek kiemelése,
- Jellemzők objektív, nagy tömegű mérése,
- Statisztikai kiértékelés,
- Automatikus elemzés,
- Képadatbázis.

Lássunk egy-egy példát a felsorolt lehetőségekre.

Genetikai vizsgálatok **egyszerű képszerkesztési funkciókkal** is nagymértékben meggyorsíthatók és a számítógépbe bevitt szakértői ismeretek segítségével a döntés folyamata elősegíthető. A szülők kromoszómáinak párosításával, esetleges eltérések felderítésével következtetni lehet arra, hogy az utód milyen örökletes tulajdonságokkal fog rendelkezni. A vizsgálat korábban (sok helyen most is) úgy történik, hogy a mikroszkópi képet, amelyen a kromoszómák találhatóak lefényképezik, majd erős nagyítást készítenek a fényképről. A kromoszómákat jellemzőik alapján párosítani kell,

amit megnehezít, hogy a kromoszómák egymáshoz képest elfordulva, a képen szét-szórva helyezkednek el. Az asszisztens ezért ollóval kivágja a képről a kromoszómákat, majd összepárosítás után az összetartozókat egymás mellé ragasztva állítja elő a további értékelés számára alkalmas tablót. Ez a körülményes és lassú eljárás a számítógépes feldolgozással hatékonyabbá tehető, sőt a feladat részben automatizálható.

Mikroszkópi képek kiértékelésénél az esetek jelentős részében a képen látható objektumoknak csak egy része (pl. a sejtmag) érdekes az adott vizsgálat szempontjából. Ilyenkor egyszerűbb vagy bonyolultabb **szegmentálási eljárásokkal** a számunkra érdekes objektumok elválaszthatók a háttértől.

Az egyes objektumok jellemzőinek mérése korábban nem, vagy csak nagyon körülményes módon volt lehetséges. A terület mérése például úgy történt, hogy lefényképezték a mikroszkópban látható képet, majd a képre helyezett milliméterpapíron számolták le a fedett terület nagyságát, vagy a mérendő területeket kivágták, és mérlegen mérték meg a súlyát, majd ebből következtettek a területre. Képfeldolgozó rendszerrel a **területen kívül a vizsgált objektum számos jellemzője (kerület, minimális átmérő, maximális átmérő, szimmetriatényező, stb.) meghatározható**. Az esetek többségében a vizsgálathoz nem egyetlen objektum, hanem statisztikailag értékelhető számú objektum valamely jellemzőjének mérésére van szükség. Ez gyakorlatilag csak számítógéppel valósítható meg. Egy képen az egyes objektumok jellemzői sorban lemérhetők és az egymás után következő képeken található objektumokra vonatkozó eredmények halmozottan tárolhatók. **Az objektumok jellemzőik szerint osztályozhatók**.

Mikroszkópi vizsgálatoknál általában több száz objektumra vonatkozó mérési eredmény statisztikája a döntés alapja. Nem az a lényeges például, hogy egy vagy két sejt szimmetria tényezője eltér a normálistól, hanem az, hogy a normális és eltérő sejtek eloszlása milyen képet mutat. Ezért felbecsülhetetlen a **nagy tömegű mérés és a statisztikai kiértékelés** lehetősége, hiszen olyan vizsgálati módszereket tesz lehetővé, amelyek egyébként nem lennének megoldhatók.

Egyes esetekben (pl. bizonyos vizeletvizsgálati módszerek) a **képfeldolgozáson alapuló teljes laboratóriumi automatikus kiértékelés** megvalósítható. Ilyenkor a minta behelyezését, mikroszkóp tárgyasztalának a mozgatását, a fókuszálást is a számítógép végzi a képfeldolgozási és az eredmény jegyzőkönyvezési feladatokkal együtt. Ezek a műszerek a nagy tömegű szűrővizsgálatok szempontjából fontosak.

A patológiai gyakorlatban nagy szerepe van a korábbi szövettani vizsgálatok eredményeivel való összehasonlításnak. A gyakran előforduló esetekben a patológus az agyában lévő „képadatbázisban” keresi a mintákat, amelyekkel a vizsgált képet összehasonlítja. Ritkán előforduló kóros eseteknél, illetve kevésbé gyakorlott patológus számára felbecsülhetetlen egy olyan képadatbázis, amelyben megfelelő mintákat talál-

hat a vizsgálatot végző orvos. Korábban színes képeket tartalmazó albumokat használtak erre a célra, napjainkban ezt a szerepet egyre inkább a **számítógépes képadatbázisok** veszik át.

A továbbiakban néhány, orvosi célra fejlesztett képfeldolgozó rendszert mutatunk be.

### Cytologiai kéпкиértékelés

A cytológiai vizsgálatoknál, a határesetek eldöntéséhez és a helyes diagnózis felállításához nagy mennyiségű sejt (több száz) morfológiai jellemzőinek statisztikai értékelését kell elvégezni. A mikroszkóp látómezőjében azonban általában csak 10-20 értékelhető sejt található, ha a feldolgozáshoz megfelelő nagyítást alkalmazunk. A mikroszkóp látómezőjében lévő kép a mikroszkópra helyezett videó kamerán és digitalizáló kártyán keresztül jut a számítógépbe, majd megfelelő programmal kiértékelik. A feladat a képen belüli sejtek szegmentálása, a morfológiai jellemzők mérése, majd azonos mintából vett több képen a mérés megismétlése, a mérési eredmények összeállítása és statisztikai kiértékelése.

PICTRON Kft által fejlesztett CYTOSOFT program a digitalizált képet először intenzitásérték szerint szegmentálja, azaz az egyes foltokat interaktívan egymástól illetve a háttértől elválasztja, a keletkezett foltokat azonos színnel kiszínezi.

Ezt követi a tényleges mérés. A mérés nélkülözhetetlen része a kalibrálás. Mivel egy digitális képen a terület vagy hosszúság mérték nélküli szám, azaz az illető alakzathoz tartozó pontok száma, megfeleltetést kell létesíteni az alakzat tényleges mérete és a hozzá tartozó pontszám között. Ez a megfeleltetés a kalibráció. A program leszámolja az illető alakzathoz tartozó pontokat, majd a kalibrációs tényező segítségével természetes mértékegységre alakítja. Kalibrálás után számos paramétert határozhatunk meg. Így kiszámíthatjuk egy alakzat kerületét, területét, konvexitását, lapultságát valamint zárványok és lyukak terület arányát. Fontos adat lehet centrumok meghatározása, illetve a centrumok közötti távolságok megmérése.

E közvetlen mérési eredmények felhasználásával méretosztályok határozhatók meg. A program a képen található objektumokat a megfelelő méret osztályokba sorolja. A mérés végeredménye egy olyan táblázat, amely az egyes méretosztályokba eső foltok számát tartalmazza. A program több képen végzett mérési eredményeket összegzi, és ez alapján statisztikát készít. A mérések segítségével megállapítható a kóros és egészséges sejtek aránya egy mintában, illetve a betegség előrehaladottságának mértéke.

## Agyinfarktus következményeinek vizsgálata

Kísérleti állaton mesterségesen előidézett agyinfarktus következményeinek vizsgálata során a szegmentálás szín szerinti osztályozással történik. Az elhalt állat agyvelőjét vékony szeletekre vágják, majd egy kiválasztott szeletet felülről megvilágítva mikroszkópra szerelt kamerával leképezik és digitalizálják.

A feladat az infarktus következtében megsérült és az épen maradt agyvelőterületek megmérése. Ennek érdekében a kalibrálás után a teljes kép pontjait három osztályba kell sorolni. A besorolás előtt az úgynevezett tanítási eljárással az ép, és az elhalt részből valamint a háttérből kiválasztunk néhány,  $16 \times 16$  képpontból álló mintaterületet. A mintaterületek kijelölését egér segítségével egy változtatható méretű négyzet (szonda) megfelelő helyre mozgatásával végezzük el. A kijelölés után a program kiszámolja az RGB térben a minta színátlagát és szórását. A keletkezett mintákat a program archíválja, hogy hasonló színösszetételű képek későbbi méréséhez ne legyen szükség újabb betanításra.

A megadott minták alapján a COLIM program az egyes képpontokat a három osztály valamelyikébe besorolja. A döntést úgynevezett euklideszi távolságmérés alapján végzi a program, azaz a program megnézi, hogy az adott képpont euklideszi távolsága melyik mintához képest a legkisebb, és annak az osztálynak a színére színezi, amelyikhez a minta tartozik, így egy három szintre vágott kép jön létre. A szakértő felülbírálhatja a program értékelését, egy változtatható méretű ceruza és egy radír segítségével egy pontthalmaz manuálisan átsorolható egy másik osztályba. Manuálisan a két félteke elválasztható, és az értékelés ennek figyelembevételével történik. Ezt követően a program még egyszer végighalad a képen, megkeresi az összefüggő értékes foltokat, és leméri azok területét. A rendszer egy adott küszöbnél kisebb területű foltokat automatikusan letörli, minden megmért objektum után kiírja annak területét, és felkínál egy listát, melyből az objektumhoz címkét lehet hozzárendelni. A címkéket a program „Szótár” menüjébe lehet felvenni és módosítani. Ez a címke gyűjtemény globális adatbázist alkot. Egy mérési ciklus a táblázat egy sorát alkotja. E sor első eleme a forrás kép neve, ezt követi egy magyarázó szöveg, amelynek tartalmát a felhasználó a mérési ciklus végén, szabadon megadhatja. A sort a mért paraméterek folytatják. Ezek lehetnek távolságra, két távolság arányára, illetve objektum részek területére utaló adatok. E számszerű eredmények további értékelését más statisztikai programokkal végezhetjük el.

E mérési eljárás segítségével következtetések tehetők az agyvelő elhalásának térbeli mértékére is, ha a vizsgálatot helyes sorrendben az összes agyvelőszeletre elvégezzük és rekonstruáljuk az elhalást háromdimenzióra. A kísérletet gyógyszeres kezelés hatékonyságának mérésére is fel lehet használni oly módon, hogy a mesterséges in-

farktust követően a kísérleti állatot gyógyszeres kezelésnek vetik alá. A kérdés az, milyen összefüggés van az adagolt gyógyszer mennyisége és a károsodott agy területe között. Erre ad választ egy ilyen mérési sorozat.

### **Mikrobiológiai-válasz kiértékelés**

Gyógyszerek hatóanyagainak értékelésénél általában egyenletesen terített táptalajon baktérium tenyészetet hoznak létre egy üvegtálcán. A baktérium tenyészetet „beoltják” a vizsgálni kívánt hatóanyag vagy hatóanyagok (különböző antibiotikumok) különböző hígítású változatával, 8\*8-as latin négyzet elrendezésben. Minden mérés során az azonos koncentrátumú hatóanyaggal 8 helyen végzik el az oltást, a latin négyzetes elrendezésnek az a célja, hogy a tenyészetben jelentkező egyenetlenségből adódó mérési eltérések hatását minimalizálja. Az antibiotikum, a beoltás centrumából kiindulva többé-kevésbe körszimmetrikus alakban, elpusztítja a baktériumokat, ennek hatására a táptalaj áttetszővé válik, az alkalmazott hatóanyagtól és a hígítástól függően kisebb vagy nagyobb „lyukak” keletkeznek a baktérium tenyészetben. A feladat a lyukak átmérőjének meghatározása, és ez alapján a hatóanyag gyógyszer vizsgálati szabvány eljárásoknak megfelelő értékelése. A feladat automatizált megoldásához a képvételt végző kamera x-y irányú léptetését és a minták stabil megvilágítását kellett megoldani.

A MIKROBI rendszerben (fejlesztő: PICTRON Kft) a mozgatást a két léptetőmotor végzi, az egyik a kamerát, a másik a kamerát tartó szánt mozgatja. A megvilágítás a mintákat tartalmazó tálca alatt, a kamerát tartó szánnal együtt mozog. A stabil megvilágítás érdekében, az egész mozgató/megvilágító rendszer burkolattal van beborítva.

Képfeldolgozás szempontjából a feladat kritikus része a határátmenetek megtalálása, és megfelelő kör illesztő algoritmus kidolgozása.

A rendszer működése közben az aktuálisan mért terület megjelenik, és ha az antibiotikumok a körtől nagyon eltérő alakzatban pusztították el a baktériumokat, akkor manuálisan korrigálható a mérendő objektum. A program folyamatosan jelzi, melyik objektum mérésénél tart a feldolgozás.

A mérés megkezdése előtt szükség van a különböző paraméterek beállítására a kalibrálás elvégzésére, a mérési elrendezés megadására. Ehhez különböző segédfilek állnak rendelkezésre (pl. antibiotikum adatbázis).

A feldolgozás eredményét képernyőn, vagy mérési jegyzőkönyv formájában, kinyomtatva kaphatjuk meg.

## Gyorsan változó folyamatok mérése

Az előzőekben ismertetett problémák és rendszerek azt a képzetet kelthetik az olvasóban, hogy az orvosi alkalmazások esetén kizárólag álló képek, nem túl nagy feldolgozási sebességet igénylő kiértékelésre van szükség. Ennek cáfolataként röviden két érdekes mérési feladatot említünk meg.

- **In-vivo kísérletek értékelése.** Számos kutatási modell vizsgálja a véráramlási viszonyokat, elsősorban a sebességet élő állatokban. Ezeknél a kísérleteknél a kísérleti állat erekkel átszőtt szervét mikroszkóp alá helyezik. Megfelelő megvilágítás és nagyítás esetén láthatóvá válik az erekben áramló vér. A vizsgálatok célja a véráramlás körülményeinek vizsgálata bizonyos jelenségek, anyagok hatására. A gondot az okozza, hogy az alkalmazott nagyítás miatt a képernyőn észlelt mozgás igen gyors, a képminőség meglehetősen gyenge és a jelenség nehezen vagy egyáltalán nem megismételhető.
- **Spermium motilitás vizsgálata.** A megtermékenyítés esélye nem csak spermiumok számától, hanem nagymértékben azok mozgékonyaságától (motilitásától) is függ. Ez utóbbi jellemző meghatározásához részben a sebesség nagyságát, részben a mozgás irányát kell vizsgálni, ami sok objektum nagyon gyors mozgásának követését igényli.

Az említett két problémával részletesebben foglalkozunk a CD- lemezen.

## Ellenőrző kérdések

1. Soroljon fel néhány orvosi műszert, amely képet állít elő!
2. Milyen jellegzetes feladatok merülnek fel az orvosi képfeldolgozásban?
3. Hol használják az időbeni képátlagolást a kép minőségének javítására?
4. Milyen előnyei vannak a számítógépes képparchiválásnak az orvosi gyakorlatban?
5. Mit takar a DICOM rövidítés?
6. Milyen feladatokat lehet számítógéppel megoldani a mikroszkópi vizsgálatoknál?
7. Milyen előnyei vannak a számítógépes képfeldolgozásnak a mikroszkópi vizsgálatok során?
8. Léteznek-e olyan vizsgálatok a mikroszkópia területén, amelyek gyakorlatilag csak számítógépes képfeldolgozással oldhatók meg?

# 11. IPARI KÉPFELDOLGOZÓ RENDSZEREK

## 11.1. Bevezetés

Széles körben alkalmazzák a képfeldolgozás módszereit az ipar területén is. Ipari alkalmazásnak tekinthetnénk az anyagvizsgálati laboratóriumokban alkalmazott mikroszkópi képelemző rendszerek használatát, vagy a távérzékelési módszerekkel végzett geológiai törésvonal vizsgálatokat is, amelyek az ásványi kincs lelőhelyek felkutatásához nyújtanak értékes segítséget, de általában ipari képfeldolgozó rendszerek alatt a gyártás során végzett minőségellenőrzést, folyamatvezérlést végző rendszereket értünk. A gyártás közbeni minőségellenőrzés előnye a gyorsaság, a termelés dokumentálása, megbízhatóság, objektivitás. Alkalmazásuk a gyártás költségeit is csökkentheti, hiszen a hibás darabok kiválasztása menet közben megtörténik, így a technológiai folyamat későbbi fázisaiban történő ráfordítások már nem jelentkeznek. Ezen feladatok speciális követelményeket támasztanak a rendszerrel szemben, és ebből adódóan speciális megoldásokra van szükség.

Általában alapvető elvárás az ipari képfeldolgozó rendszereknél:

- a real-time feldolgozás,
- az adott ipari alkalmazáshoz illeszkedő pontosság,
- a mérések reprodukálhatósága, dokumentálása,
- az ipari környezetben való alkalmazhatóság,
- a vezérlési, visszacsatolási funkciók biztosítása,
- a nagy megbízhatóság.

Az alábbiakban ezeket az elvárásokat és ezekkel kapcsolatos megoldásokat részletezzük, majd néhány példát mutatunk be erről a területről. A fenti elvárások a különböző alkalmazásoknál különböző súllyal jelentkeznek, ezért többnyire az ipari képfeldolgozó rendszerek egyedi megoldásokat igényelnek. Ennek ellenére néhány közös jellemző megfogalmazható, ezeket kívánjuk ismertetni.

## 11.2. Ipari képfeldolgozó rendszerek jellemzői

Az ipari alkalmazásoknál alapvető követelmény, hogy a gyártás folyamatát ne lassítsa le a beiktatott képfeldolgozó rendszer. A **real-time feldolgozás** ebben az esetben azt jelenti, hogy a mérési, képelemzési és ezekhez kapcsolódó vezérlési funkciók végrehajtási idejének (beleértve esetleg mechanikai, pneumatikus mozgások elvégzését) a gyártósor sebességével összhangba kell lennie. Ha például a feladat nyomtatott áramkörön lévő furatok meglétének és méretének ellenőrzése, akkor az egész képe-



lemzési, mérési vezérlési feladatnak rövidebb idő alatt kell lezajlania, mint amennyi idő alatt a munkadarab a mérési helytől a beavatkozás helyéig (pl. a hibás munkadarabot leterelő automatáig) eljut. Ez az idő alkalmazásonként változik, de az időkorlát a rendszer fejlesztése szempontjából többnyire kritikus tényező. Ebből a követelményből adódik, hogy általában ad hoc megoldásokat kell keresni, és maximálisan ki kell használni a konkrét feladatból adódó könnyítési, egyszerűsítési lehetőségeket. Többnyire a teljes kép feldolgozására nincs idő, ezért csak az előzetesen kijelölt vizsgálati helyeken (angol terminológiából átvett kifejezéssel - Region Of Interest = ROI) végezzük el a vizsgálatot. Szerencsére a gyártósoroknál általában a munkadarabok mozgása bizonyos tűréshatáron belül meghatározott, és a technológiai folyamatok jelentős részében a munkadarab rövid időre történő megállítása sem okoz gondot, mivel egyes munkafázisok elvégzéséhez azt egyébként is meg kell állítani.

Léteznek olyan feladatok is, ahol nem, vagy csak bizonyos korlátok között valósítható meg a vizsgált objektumok jól pozicionált mozgása. Erre jó példa a gyümölcs vagy egyéb mezőgazdasági termény minőségi osztályozása méret, szín, alak vagy egyéb jellemzők szerint. Ezekben az esetekben a folyamatosan mozgó darabok azonosítását is meg kell oldani az egyéb paraméterek mérése mellett.

További csoportot alkotnak a folyamatos vizsgálatot igénylő alkalmazások, mint például a szövetgyártás közben történő minőségellenőrzés. Ekkor nem jól definiált egyedi darabokat kell vizsgálni, hanem a folyamatosan mozgó, adott szélességű anyagban a szövési hibákat kell megtalálni. Ilyen feladatoknál általában vonal kamerákat alkalmaznak, amelyek egyetlen, esetenként több ezer elemből álló CCD érzékelők. Az érzékelő sor előtt egyenletes sebességgel elhaladó minta soronkénti értékelését kell megoldani, a mozgás irányában végtelen képen.

A fenti kiragadott néhány példa is egyértelműen bizonyítja, hogy általános célú képfeldolgozó rendszerek ezekben az esetekben legfeljebb a kezdeti tapasztalatgyűjtés során alkalmazhatóak. Minden egyes alkalmazásnál néhány, a feladatra optimalizált funkcióra van szükség, tehát egy általános célú rendszer funkciókészlete felesleges. Amire meg szükség lenne az nem elég gyors, éppen az általános, többcélú felhasználás biztosítása miatt. Ebből adódik, hogy sok esetben az ipari alkalmazásoknál nem általános számítógépet, hanem a specifikus feladatra „felprogramozott” speciális processzort tartalmazó célszámítógépet használnak.

Az ipari rendszerekben meghatározó fontosságú kérdés a **mérési pontosság**. Az elvárt pontosság adódik a technológiai előírásokból, a másik oldalon - a sebességi elvárások miatt - többnyire adott a TV kamera+digitalizáló kártya kombináció alkalmazása. Ezek geometriai felbontása a TV rendszer sajátosságából fakadóan korlátozott, így a tényleges fizikai méret és a mért objektum pixelekben kifejezhető méretei közötti viszonyt az alkalmazott optikai rendszer (a kamerára szerelt objektív) határozza meg.

Ebből következik, hogy ha abszolút méreteit tekintve kis tárgyakat mérünk, akkor erős nagyítást biztosító optikát kell alkalmazni. Problémát az okoz, ha az elvárt pontosság és a munkadarab méretei között több nagyságrend eltérés van, mivel ebben az esetben a tárgy nem fér bele a látómezőbe, ha a tűréshatárnak megfelelő nagyítást alkalmazunk. Például egy 50 cm-es munkadarabon bizonyos részek pozícióját kell ellenőrizni tized milliméteres pontossággal. A probléma megoldását jelentheti, ha például a gyártási folyamat során meg lehet állítani a mérendő munkadarabot, és a vizsgálandó pozíciók környezetében külön kamerákkal, erős nagyítás mellett végezzük el a mérést.

A mérések pontosságához, de egyben a **mérések reprodukálhatóságához** kapcsolódik az ipari rendszerek **kalibrációja**. Ebben a fázisban a rendszer hitelesítése történik etalon tárgyak elhelyezésével, és annak megállapításával, hogy egy-egy pixel az adott rendszerben milyen fizikai méreteket reprezentál. Ez az adott rendszer összes komponensének paramétereitől és beállításától függ (kamera, digitalizáló kártya, objektív, megvilágítás, szoftver, algoritmusok, stb.), ezért a rendszer bármelyik elemének megváltoztatásakor a kalibrálást újra el kell végezni, és az adott gyártási-mérési folyamat során az állandóságukat biztosítani kell. Példaként megemlíjtjük, hogy a megvilágítás változása is befolyásolja a mérés eredményét. Gondoljunk arra, hogy feldolgozás során többnyire a kamera képet valamilyen eljárással két szintre vágjuk, és a bináris képen történik a mérés. Ha beállítunk egy vágási küszöbszámot a két szintre vágásnál, és időközben a megvilágítás megváltozik, teljesen más mérési eredményeket kapunk.

A reprodukálhatóság érdekében a kalibrációs adatokat **dokumentálni** kell. Ezen túlmenően általános elvárás az ipari rendszereknél, hogy a mérési folyamat eredményeit naplózza a rendszer, a gyártási technológia szempontjából fontos adatokat tárolja, szükség szerint kinyomtassa. Ez a típusú **dokumentálhatóság** nem képfeldolgozási, hanem hagyományos adatfeldolgozási feladatokat jelent.

Az **ipari környezetben való alkalmazhatóság** nem képfeldolgozási probléma, de a tervezéskor nem hagyható figyelmen kívül, hogy esetleg a műszerekre káros, vagy a mérést zavaró környezetben (por, rázás, pára stb.) kell telepíteni a rendszert.

A **vezérlési, visszacsatolási funkciók** megteremtése a szokásos vezérlő rendszereknél alkalmazott csatoló kártyák és szabványosított interfészek alkalmazását jelenti. Itt említjük meg, hogy léteznek olyan digitalizáló kártyák, amelyekre eleve ráépítenek vezérlő jelek fogadására és kiadására alkalmas csatlakozásokat. Más esetekben a képfeldolgozó és vezérlő rendszert külön célgépként építik meg (az ipari igényeknek megfelelő robusztus kivitelben), és a számítógépet ezek programozására, és/vagy egyetlen soros vagy párhuzamos interfészen megvalósított kapcsolat révén a rendszer magasabb hierarchia szintjén használják. Nagyobb rendszerek esetén több képfeldolgozó/vezérlő célgép munkáját vezérli egy számítógép.

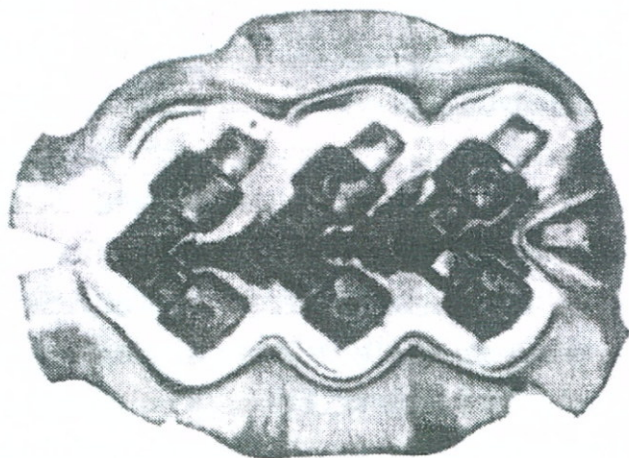
A **nagy megbízhatóság** két vonatkozásban is érthető elvárás az ipari képfeldolgozó rendszereknél. Egyrészt a hibás felismerés rossz döntést eredményez, aminek a hatására a hibás munkadarab a gyártási folyamatban marad, így további költséges munkafázisokat hajtanak végre rajta, és a végtermék selejt lesz. Más vonatkozásban elvárás a rendszerrel szemben, hogy a rendszer meghibásodása és az ebből fakadó leállás minimális legyen.

### 11.3. Alkalmazási példák

Az elmondottak szemléltetésére lássunk néhány konkrét feladatot. Az alábbi alkalmazások képekkel illusztrálva a CD lemezen is megtalálhatók.

#### Pozicionáló robot vezérlése

Az első egy **munkadarab helyes pozicionálását** oldja meg. A 23. ábrán látható csatlakozó idomokról a gyártási folyamat során le kell vágni a képen jól látható felesleges széleket.



23. ábra

A vágószerszámhoz azonban csak adott pozícióba érkezhetsen a munkadarab, különben a munkadarabot rossz helyen vágja el, és maga a vágószerszám is rövid időn belül tönkremegy. A képfeldolgozó rendszer feladata annak megállapítása, hogy a négy lehetséges helyzetből melyikben van a munkadarab, (a csődarabok „könyöke” előre vagy hátra áll, illetve a hosszabb csont lefe-

lé vagy felfelé helyezkedik el), és a megfelelő robotnak kiadni az utasítást, hogy a munkadarabot forgassa a helyes pozícióba, mielőtt az a vágószerszámhoz ér. Látható, hogy a három csődarab közötti árnyék zavarja az egyes darabok elkülönítését és megnehezíti a helyes szegmentálást. Az adott feladatban azonban nem érdemes és nem is szükséges a kép korrekt szegmentálása, hiszen négy lehetséges pozíció között kell választani, és a szállító rendszer lehetővé teszi rövid időre a munkadarab megállítást. Ezek tudatában néhány jól kiválasztott helyen végzett mintavétellel és a betanítás során kapott mintákkal való egyszerű mintaillesztéssel eldönthető, hogy milyen pozícióban áll a munkadarab. Természetesen más munkadaraboknál esetleg más helyeken kell

a mintákat kiválasztani, és a betanítást elvégezni. Ez a példa jól szemlélteti, hogy ad-hoc, egyszerű megoldások sokszor eredményesebbek, mint a probléma általános megközelítése.

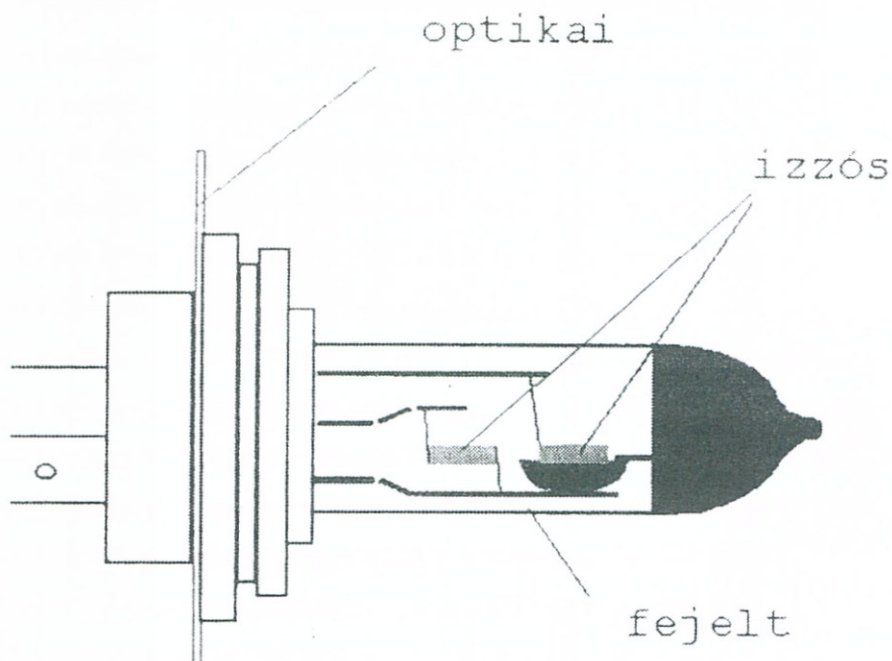
### Műszerfal ellenőrzés

Másik példaként a **személygépkocsik műszerfalának** ellenőrzését említjük. Lényegében arról van szó, hogy jól meghatározott bemenő jelek hatására a műszerfalon adott lámpáknak ki kell gyulladni, megadott frekvenciával villogni, illetve adott mérőeszközök hatására adott műszerek mutatójának megadott mértékben ki kell térniük (fordulatszám-láló, sebességmérő). A lámpák felvillanásának ellenőrzése, illetve a mutatóknak adott terhelés hatására adott pozícióra való beállításának ellenőrzése képfeldolgozási szempontból elvileg nem okoz gondot. Az első problémát az okozza, hogy ha egyetlen kamerával nézzük a műszerfalat, akkor a felbontás már nem lesz elég pontos a mutatók beállításának ellenőrzésére, tehát 2, esetleg 3 kamerát kell alkalmazni a műszerfal befogására. Külön feladat (nem képfeldolgozási!) a tesztelő padnál a szimulátor kialakítása - egyszerű és biztos csatlakozási felülettel - amire a vizsgálatához érkező műszerfal egyetlen mozdulattal rögzíthető. Feltétlenül gondoskodni kell az egyenletes megvilágításról, a rendszer árnyékolásáról, ugyanis a szerelő csarnokban az év- és napszaktól, sőt az időjárástól függően változik a megvilágítás. Láthatjuk tehát, hogy az ipari képfeldolgozó rendszereknél sokszor nem a képfeldolgozási feladat, hanem az egyéb körülmények biztosítása okozhat problémát.

### Autólámpák gyártásvezérlése

A harmadik példa a FALCON-VISION Kft által megvalósított rendszer, amely a H4 típusú **autólámpák képfeldolgozáson alapuló gyártás vezérlését és a gyártás utáni termékválogatást valósítja meg.**

A gyártás befejező fázisában hegesztik össze a spirált tartalmazó, ún. fejtelt lámpát - 24. ábra - az optikai tárcsával (ezt kell betenni a foglalatba az autóban). Nagy jelentősége van annak, hogy ez a két alkatrész megfelelően van-e összeillesztve. A megfelelő összeillesztés vonatkozik az alkatrészek által bezárt szögére, valamint az izzószálaknak az optikai tárcsához viszonyított magasságára. A feladat megoldására kifejlesztett képfeldolgozó és gyártásvezérlő rendszer tízmikronos (!) pontossággal méri az alkatrészek helyzetét, megállapítja, hogy azok beállíthatók-e a követelmények szerint, elvégzi a beállítást, majd összehegesztés után újabb mérést végez, melynek alapján minőségi osztályokba sorolja a kész lámpákat (természetesen a válogatást is vezérli). A gyártó gép sebessége 1800 db/óra.



24. ábra

A rendszerben három mérőhelyen összesen kilenc darab CCD kamera készít képet, nemzetközi szabvány által előírt mérőirányokból. A mérések bevetített fénypon-  
ton (mint abszolút referencia-  
ponton) alapulnak. Ebben a rendszerben sem a teljes képet dolgozzák föl, hanem csak a kezelőfelület segítségével definiált mérőablakokba eső képrészleteket. Ennek a feladatnak nem a képfeldolgozás a legnehezebb része

(bár a kívánt pontosság speciális szoftver- és hardvermegoldásokat igényel). Az igazi nehézséget az jelentette, hogy a lámpának kb. húszféle paramétere van, amelyek változásának bonyolult kölcsönhatása a beállítás optimalizálását, tehát megfelelő szoftver kidolgozását kívánta meg. (A programnyelv C, illetve assembly.)

A kezelőfelület (amelynek feladata a hosszútávú adatgyűjtés, a rendszerparaméterek beállítása és a dokumentálás) PC-n fut.

Hangsúlyosan fontos a rendszer kalibrálása, amelyhez külön erre a célra kifejlesztett eszközöket használnak.

A rendszer egységei között forgó jelátvitellel megvalósított, kétirányú, infravörös kommunikáció zajlik (9600 bit/sec sebességgel). A tápegységek, a rendszervezérlő és a léptetőmotorok között hálózatos, hibatűrő jelforgalmat alakítottak ki. Folyamatos önteszt gondoskodik a biztonságos üzemelésről.

A rendszer bonyolultságát érzékelteti néhány számadat:

a rendszerbe épített processzorok száma:	1 db	TMS320c30 (DSP)
	10 db	TMS34010
	2 db	M68000
	120 db	PIC
a léptetőmotorok száma:	96 db	
a beépített kártyák száma:	235 db	
programozható, processzorvezérelt tápegység:	9 db.	

## Szelepülék minőség ellenőrzése

Az ipari képfeldolgozó rendszerek egy része minőség ellenőrzési feladatokat old meg. A **gázkészülékekben lévő szelepek záródásának ellenőrzése** is ebbe a csoportba tartozik. A szelepek akkor zárodnak jól, ha az elzáró szelep egyenletes, hibátlan, sík felületre fekszik fel záródáskor, ezért az úgynevezett szelepüléket megfelelő számmal csiszolják egyenletesre. A vizsgálandó munkadarabon 5 darab átmenő furat van, ezek végére fekszik fel a zárószelep. Miután a furatok különböző mélységben és egymáshoz elég közel helyezkednek el, a megfelelő képminőség biztosítása érdekében speciális megvilágító rendszert és 5 CCD kamerát tartalmazó mérőpadot kellett kialakítani. A megvilágítás féligáteresztő tükrökön keresztül érkezik a furatokra, és ezek a tükrök vetítik a kamerák CCD érzékelőjére a képet.

A mérés akkor indul, amikor a munkadarab a felfogó sín hátfalán elhelyezett kapcsolót benyomja. A számítógép vezérli a megvilágítás és a képvétel időzítését. A program ellenőrzi a belső kör sugarát és a körgyűrű szélességét is. A minősítés a körgyűrűn található foltok méretétől és számától függ. A mérési eredményeket és a hibás munkadarabok képét a program naplófájlban tárolja.

A rendszer kalibrálása a mérési paraméterek beállítása a rendszer fontos része, ezt a feladatot csak az arra feljogosított személy, megfelelő jelszó ismeretében végezheti el. A hibák jellegének értékelését külön program funkció végzi, amely a gyűrű kina-gyítását és egyéb vizsgálatokat is lehetővé tesz.

## Faminták automatikus válogatása

A faiparban a képfeldolgozás kb. 10-15 éves múltra tekint vissza. Az első években csak szenzorok alkalmazásával próbálkoztak, de ma már teljes körű CCD érzékelős fa-vizsgáló és termékellenőrző rendszerek léteznek. Használják a faipari minőségellenőrzésben lézert is, ez alaktani eltérések, ill. felületi fizikai hibák felderítésére alkalmas.

Világviszonylatban jelenleg kb. 300 millió m<sup>2</sup> parkettát készítenek évente a legkülönbözőbb eljárásokkal. Ennek a mennyiségnek jelenleg 15-25 %-át vizsgálják a számítógépes minőségellenőrző rendszerek, de ez az érték folyamatosan nő. A jelenleg üzemelő rendszerek átlagsebessége 150 m/perc, ami a gép mellett 2-3 fős kiszolgáló személyzetet igényel. Egyes elemzők szerint a hardware eszközök már megfelelnek a teljesen automata rendszerek követelményeinek, csak a tökéletesen megfelelő software hiányzik, ill. az ipari képfeldolgozás e területét kell olyan szintre fejleszteni, hogy megoldást nyújtson e problémára.

Cél, hogy különböző minőségű famintákról a lehető legtöbb információt szerez-

zünk egy automatikus szétválogató rendszer számára előre rögzített elvárások szerint, majd elvégezzük a minták automatikus osztályozását.

A probléma az emberi látórendszer vizuális befolyásolhatósága alapján merült fel. A faanyagokat feldolgozó cégek zöménél a faanyagot teljesen szubjektív megítélés alapján osztályozzák. Természetesen a dolgozók által is ismert technológiai utasításokban szerepelnek különböző minősítési szempontok. Ezek részben általános faipari szempontok alapján készültek (fa színe, egészségi állapota, növekedési ill. a feldolgozás során keletkezett hibák, stb.), részben pedig a vevői/piaci elvárásokat tükrözik.

A fa mintázatát színezetét jelentősen befolyásolja a termőhely, az éves csapadék mennyiség, mikor vágták ki a fát, hol tárolták, hogyan szárították, hogyan gőzölték, stb. Mivel az osztályozás teljesen manuálisan történik, így mind látórendszeri, mind a tényleges munkavégzés közbeni lokális-intenzitásbeli problémák felmerülhetnek. Ezen utóbbira világít rá az alábbi gyakorlatban előforduló probléma: amennyiben egy világosabb rönkből származó lapkasorozatba bekerül egy sötétebb lapka (a lapka egy 65x450x5 mm-es méretre bevágott fa, ami a szalagparketta fedő, járórétegét képezi), akkor az egy egész más osztályba is bekerülhet a szubjektív megítélés alapján, mint ahova tőle kismértékben eltérő környezetű minták esetén került volna. Ezen probléma kiszűrésére merült fel a számítógép használata.

A jelenleg működő rendszerek alapja, hogy elsősorban a fa szerkezete, színe, hibái szolgálnak az elemzés alapjául. A rendszert minták alapján betanítják, majd a futószalagon érkező alapanyagokat a CCD kamera segítségével beviszik a számítógépbe. Egy ilyen felvétel felbontása nagyon magas is lehet akár 0,007 mm-es pontosság is elérhető. A nagy teljesítmény elérése érdekében egy berendezésen belül több processzor dolgozik párhuzamosan, és egy vezérlőegység osztja szét a rögzített képet az éppen szabad egységnek. Majd ezeket a képeket - ismert minőségi követelményeknek megfelelő - mintákról készült képekhez hasonlítják. A kezelhetetlen alapanyagot egy külön csoportba teszik, amelyet emberi segítséggel utólag felülvizsgálunk, majd osztályozunk.

## Ellenőrző kérdések

1. Milyen speciális elvárások merülnek fel ipari képfeldolgozó rendszereknél?
2. Mit jelent a real-time feldolgozás az ipari képfeldolgozásban?
3. Miért célszerű az ipari alkalmazásokhoz egyedi megoldásokat felhasználni?
4. Soroljon fel néhány megoldást a mérési pontosság fokozására!
5. Mi a célja a kalibrációnak?
6. Milyen körülményeket kell figyelembe venni a mérés reprodukálhatósága érdekében?

7. Ismertessen egy alkalmazási példát az ipari képfeldolgozás területéről!
8. Mi okozza a legnagyobb nehézséget faipari alkalmazások során?
9. Mi képezi a jelenleg működő faipari képfeldolgozó rendszerek alapjait?



## 12. MEZŐGAZDASÁGI ALKALMAZÁSOK

### 12.1. Bevezetés

A mezőgazdaságban folyó kutatások során alkalmazott mérési és kiértékelési módszerek gyakran pontatlanul, esetleg szubjektív ítéleteket is tartalmazó módon végezhetőek el. Nagyon sok feladat a természetes környezet gyakori megváltoztatását teszi szükségessé. A kutatók előtt örökös problémaként vetődik fel az a kérdés is, hogy a növényekről vizuálisan nyert információkat hogyan lehet számszerűsíteni. E célt szolgálják a különböző szubjektív megítélésen kialakított és alkalmazott skálák (pl. EWRC, borítási százalék, stb.). Az így nyert információk biometriai egzakt feldolgozása számos problémát vethet fel. Ugyanakkor számos kísérlet kiértékelése szubjektív ítéleteket is tartalmazhat, amely a kapott eredményeket bizonytalanná teheti, esetleg további kísérleteket igényelhet. Mindez az esetek többségében hosszú idő (hónapok, esetleg évek) kitartó munkáját és jelentős anyagi költségeket kíván. A kísérletek kiértékelésének matematikai alapjai régóta ismertek. A kivitelezést számos számítógéppel vezérelt eszköz támogatja. Ezek gyakran céleszközök, amelyek egy adott feladat megoldására alkalmasak. Általában a beépített mérési és/vagy kiértékelési funkciók rugalmas változtatására nem alkalmasak, vagy komoly szakirányú felkészültséget igényelnek. A legtöbb esetben azonban szükség lehet egyedi problémák megoldására is.

Ezen gondok megoldásában jelenthet segítséget a digitális képfeldolgozás alkalmazása egyes mezőgazdasági kísérletek mérésénél és a mérési eredmények feldolgozásánál.

Ha a képfeldolgozás mezőgazdasági alkalmazásait akarjuk felsorolni, esetleg csoportosítani, igen nehéz feladatot vállalunk magunkra, miután nagyon szerteágazó és nagyon eltérő jellegű alkalmazásokkal találkozunk. Csak példaként utalunk arra, hogy az állattenyésztés is a mezőgazdasághoz tartozik, és ezen belül az állatgyógyászat a humán gyógyászatnál alkalmazott teljes technikai arzenált, így a képfeldolgozó rendszereket tartalmazó műszereket is felhasználja munkája során. Az orvosi alkalmazások fejezetben említett spermium motilitás vizsgáló rendszert például szélesebb körben használják az állattenyésztésben, mint a humán területen, hiszen az állatok esetén a mesterséges megtermékenyítés gyakran alkalmazott módszer, és a vizsgálat elvégzése a sikeres megtermékenyítési százalékot jelentősen növeli. További érdekességként megemlítjük, hogy van olyan hely, ahol a nagy értékű CT (Computer Tomográf) berendezést nappal emberek, éjszakai műszakban állatok vizsgálatára használják. Képfeldolgozás szempontjából a két felhasználás között nincs különbség, más kérdés, hogy a műszakváltások között a teljes rendszert sterilizálni kell, sőt baleset esetén 15

perc alatt át kell váltani a humán vizsgálatra. A példa jól mutatja, hogy a mezőgazdasági alkalmazások között minden előfordulhat. Ez meg is könnyíti a dolgunkat, hiszen az előző fejezetekben ismertetett alkalmazási területek mindegyike megtalálható a mezőgazdasági alkalmazások között is. Gondoljunk bele, hogy a távérzékelési alkalmazások igen nagy hányadát teszik ki a mezőgazdasági felhasználások, ebben a vonatkozásban tehát elegendő utalni az ott leírtakra.

Ha tovább vizsgálódunk, akkor megállapíthatjuk, hogy az egyéb alkalmazások lényegében két nagy csoportra oszthatók, az egyik a mikroszkóp közbeiktatásával vagy közvetlen videokamerával felvett és digitalizált képek kutatási, vagy minőség vizsgálati célú értékelése, a másik az ipari jellegű alkalmazások köre. Mindkét témakörrel volt már szó ezért a továbbiakban megpróbáljuk megfogalmazni, és néhány példával megvilágítani azokat a jellegzetességeket, amelyek eltérést mutatnak a korábbiakhoz képest. A fejezet további részében az első csoportba tartozókat nevezzük képelemzési feladatoknak, a másik csoportba tartozókat pedig ipari jellegűeknek.

Összefoglalásként megállapíthatjuk, hogy a mezőgazdasági alkalmazások szétválasztása az alábbi szempontok alapján tehető meg:

1. A feladat eredete alapján:

- Növényi
- Állati
- Mindkét

2. A képfeldolgozási feladat alapján:

- Képelemzési feladat
- Osztályozási, alakfelismerési feladat
- Mesterséges látórendszeri feladat

3. A folyamat időbeli lefolyása alapján:

- Időben állandó
- Időben változó

4. A folyamatról készült vizuális információ alapján:

- Állóképek feldolgozása
- Mozgóképek feldolgozása
- Álló- és mozgóképek feldolgozása

5. A folyamat térbeli kiterjedése alapján:

- Mikro (pl. mikroszkópi feldolgozások)
- Makro (pl. digitális kamerával készült lokális terepi felvételek)
- Globális (légi- és űrfelvételek)

## 12.2. Képelemzési feladatok

A digitális képelemző rendszereket többnyire kutatási vagy minősítési feladatokra alkalmazzuk. Ha a különböző alkalmazásokat áttekintjük, megállapíthatjuk, hogy a mezőgazdaság területein, más alkalmazási területekhez képest gyakrabban van szükség:

- szín információ felhasználására,
- háromdimenziós objektumok vizsgálatára.

Elég arra gondolunk, hogy egy levélen a növénybetegség vizsgálatánál a szín lehet a legfontosabb információ, ugyanakkor a levél, a mag a gyökérzet nem tekinthető két-dimenziós alakzatnak. Ez sokszor megnehezíti a kiértékelést (zavaró árnyékhatások lépnek fel, a kétdimenziósra leképezett felület területe lényegesen eltérhet a tényleges területtől stb.) ezért indokolt, hogy itt hívjuk fel a figyelmet a mezőgazdasági képkiértékelési feladatoknál (de általában minden videokamerás bevitt alkalmazó rendszer-nél) a felvétel során felmerülő gyakorlati szempontokra.

Nézzünk egy példát. Egy előkészített (pl. mosott gyökér, kiválasztott felvételezési irány, mikroszkópi metszet, stb.) **objektum**ról kamera segítségével készítünk felvételeket. Ekkor a következőket célszerű figyelembe venni:

- A felvételezés során mindig készítsünk az objektumhoz és a mérési feladathoz illeszkedő **referencia felvételt**.
- A zavaró képrészeket próbáljuk meg már felvételezéskor kiküszöbölni vagy csökkenteni.
- A megvilágítás megfelelő erősségű és közel homogén legyen.
- A TV kamerával készített képek széleinél fellépő fényszegényedés okozta hibák (shading) kiküszöbölése érdekében igyekezzünk az értékes információkat a közel lineáris érzékelési tartományban, a kép középpontjában elhelyezni.
- Az objektum(ok) részei, részletei könnyen elkülöníthetők, egyértelműek, a feldolgozási célnak megfelelőek legyenek.
- A felvételezéskor megfelelő minőségű és fókusztávolságú lencsét használjunk.

Valódi színes felvételek készítése teljesen hasonló az előzőekben ismertetett felvételezési technikához, a következő eltéréssel: a CCD videojelet **24, 36 vagy 48 bites képpontokká** alakítjuk. Az átalakítás során három képsíkot veszünk igénybe, a három alapvető (Red, Green, Blue) érzékelési sávnak megfelelően. Gyakran célszerű az egyes sávoknak megfelelő digitális információt külön-külön is tárolni. Gondolni kell arra is, hogy az olcsóbb, multimédia célú digitalizáló kártyák általában YUV formában állítják elő a képet, és az RGB komponenseket szoftver úton, közelítő eljárásokkal szolgáltatják. Ezért ezek a kártyák csak korlátozottan használhatók fel képfeldolgozási célokra, azokon a területeken, ahol a jelentkező hiba nem befolyásolja a feldolgozás eredményét.

Valódi színes felvételek készítésekor ügyelni kell a megfelelő, színhelyes felvételezésre is. Korrekciókra általában kevert fényviszonyok (pl. napfény és fénycső együtt) mellett történő felvételezéskor van szükség. Ennek korrigálására automatikus korrekció hiányában **színhőmérséklet mérő készüléket** és **színhőmérséklet módosító szűrőket** alkalmazhatunk.

Mikroszkóppal készített felvételek során a kamera CCD érzékelője közvetlenül a mikroszkóp lencserendszere által leképezett képet érzékeli. Itt is problémát okozhat a fehér- és **színegyensúly** helyes beállítása. A felvételek készítésekor kamerához illesztett mikro-közdarabot és megfelelő szemlencsét használhatunk. A nagyítás mértékét, feladattól és az alkalmazott objektívek nagyságától függően választhatjuk meg.

A jó képfelvétel alapvető a feldolgozás eredményessége szempontjából. A további feldolgozás szempontjából elmondható, hogy bár a szokásos képelemzési eljárások többsége használható ezen a területen, de sokszor van szükség egyedi megoldásokra is. Az általános módszerek közül gyakran alkalmazzák az automatikus vagy interaktív úton vezérelt terület és kerület mérést, különböző morfológiai jellemzők szerint történő osztályozást, szín szerinti szegmentálást és osztályozást.

Továbbiakban néhány konkrét alkalmazást sorolunk fel röviden.

- Magok ill. vegetatív részek alakjának, méretének és felületének meghatározása.
- Vetőmagvizsgálat, fajtaazonosítás, növekedésdinamikai vizsgálat
- Növényi leveleken a sejtfal időbeli változásának analízise.
- Lucernaszár szöveteiben a lignin koncentráció meghatározása.
- Hagyma, paradicsom, búza gyökér növekedésének mérése.
- Fitofág rovarok által okozott levél területi veszteség meghatározása.
- Különböző gabonamagvak elkülönítése 3D képeken a magok profilja, valamint 2D képeken a magok kontúrja és néhány jellemző morfológiai paraméter alapján (maximális hossz, maximális szélesség, kontúrterület, a kontúrponatok átlagos távolsága egy önkényesen választott centrumtól, a profil első deriváltjának átlaga, stb.).
- Sérült kukorica magvak automatikus kiválasztása.
- Növényi szövetek degradációjának meghatározása.
- Színes (3\*8 bites) növény identifikáló rendszer.
- Kórokozó gombák számítógépes felismerése.

A CD „Mezőgazdasági alkalmazások” fejezetében további érdekességek és információk található az alábbi témaköröket illetően:

- Irodalmi áttekintés
- A szubjektív emberi ítéletek minimalizálása a méréseknél

- Képekkel illusztrált feldolgozási példák
- Alkalmazási lehetőségek

### 12.3. Ipari jellegű alkalmazások

A mezőgazdasághoz sorolható élelmiszeripar területén az alkalmazások jelentős része semmiben sem különbözik más ipari rendszerektől. Például a képfeldolgozó rendszer szempontjából teljesen közömbös, hogy valamilyen élelmiszer vagy gyógyszer csomagolásán elhelyezett címke vagy felirat meglétét és olvashatóságát ellenőrzi. Léteznek azonban jellegzetesen mezőgazdasággal kapcsolatos feladatok, amelyek a képfeldolgozási eljárások szempontjából eltérő megközelítést igényelnek.

A élelmiszeriparban jelentkező osztályozási, minősítési feladatoknál például a méret és alak meghatározás annyiban tér el a megszokottól, hogy a **háromdimenziós „munkadarabon”** (alma, uborka, burgonya stb.) nem jelölhető ki sík felületek a mérés elvégzéséhez, hanem az árnyékhatásokkal torzított felület kétdimenziós vetületi képen kell a mérést elvégezni. További problémát okoz, hogy nem biztosítható a mérendő objektumok megállítása, pozicionálása. Ezekben a rendszerekben többnyire a **szín információt** is fel kell használni az értékelés során (érettségi állapot, károsodások). A nehézségek ellenére léteznek képfeldolgozó rendszerek uborka, burgonya, gyümölcs és más mezőgazdasági termények osztályozására. A manuális munka helyettesítésére vagy gyorsítására jó példa, hogy burgonyagumók közül a sütésre alkalmas gumókat képfeldolgozó rendszer segítségével szeparálták. A rendszer **„neural network”** alapú eszközt tartalmazott, amelyre hatékony algoritmusok kidolgozását is elvégezték az angol kutatók. Szintén neurális hálózati feldolgozó eszközt alkalmaztak egyes sárgarépa típusok elkülönítésére. A rendszert 80 mintának 5 osztályba való sorolásával „tanították”.

Ipari jellegű alkalmazásnak tekinthető az élő állatok mérése is, ebben az esetben annyival bonyolultabb a helyzet, hogy a „mérendő objektum” saját lábán halad a mérési pozíció felé és az időzítés a pásztorok ostorának függvénye. Magyarországon is alkalmaztak képfeldolgozó rendszert szabadon tartott szarvasmarhák méreteinek meghatározására. Az állatokat korlátokkal kialakított keskeny folyosóba terelték, amit egy sorompóval zártak le mindkét végén, amikor „a munkadarab” a megfelelő pozícióba érkezett. A súlyt mérleggel, az állat egyéb méreteit a folyosó fölé, és oldalt elhelyezett kamerák digitalizált képen mérték le. Az azonosítás az állatok nyakörvére szerelt azonosító lappal hagyományos módon történt, de érdekességként megemlítjük, hogy a szarvasmarhákon lévő foltok mérete és elhelyezkedése alkalmas az egyedi azonosításra, éppen úgy mint az ember esetén az ujjlenyomat.

Sikerrel kecsegtető fejlesztések folynak Izraelben mezőgazdasági robotok valós

idejű látórendszerének fejlesztése és az érett dinnye detektálására történő alkalmazása érdekében. A rendszer egy pipeline alapú képfeldolgozó modult és valós idejű feldolgozó hardver eszközöket tartalmaz. A továbbfejlesztést a rendszerhez illesztett (közeli-) infrakamera alkalmazásával próbálják pontosítani, illetve bizonyos károsodások detektálására felhasználni.

A fenti példák remélhetően jól illusztrálták a képfeldolgozás alkalmazásának széleskörű lehetőségét, és a mezőgazdaság területén jelentkező speciális problémákat is.

## Ellenőrző kérdések

1. Miben jelent segítséget a digitális képfeldolgozás a mezőgazdasági kísérletek értékelésénél?
2. Milyen csoportosítási szempontokat ismer a mezőgazdasági felhasználások esetén?
3. Mi jellemző a mezőgazdaságban használt képfeldolgozási technikákra?
4. Milyen alkalmazási területeket ismer?
5. Mire érdemes figyelni digitális képek készítésénél?
6. Hol használnak valós idejű látórendszert?
7. Milyen képfeldolgozási módszerekkel szeparálható a beteg és az egészséges levél?
8. Analóg vagy digitális módszerrel mérhető könnyen és pontosan mezőgazdasági objektumok területe? Miért?

## 13. IRODAI ALKALMAZÁS

### 13.1. Bevezetés

Az irodai alkalmazások és képfeldolgozás az 1980-as évek közepéig két egymástól teljesen eltérő világot jelentettek. Ebben az időszakban jelentek meg azok az eszközök, amelyek lehetővé tették, hogy képfeldolgozási megoldásokat alkalmazzanak az irodaautomatizálás területén. Divatba jött a papírmentes iroda jelszó, amely azt tűzte ki célul, hogy a dokumentumokat elektronikus úton állítsák elő, tárolják és osszák szét a felhasználók között. Ez működött is abban az esetben, ha minden dokumentum számítógépes úton keletkezett, de problémát jelentett a korábbi, vagy más felhasználtól érkező, papíron lévő dokumentumok rendszerbe való bevitele és felhasználása. A szkennerek megjelenése a dokumentumok képszerű bevitelét megoldotta, és a karakterfelismerő programok a másik problémára, nevezetesen a képformátumú dokumentumok karakteres formába történő átalakítására is megoldást kínáltak. Képfeldolgozási szempontból az irodai alkalmazásnak két vonzata van, az egyik a dokumentumok képi formában történő archiválása, a másik a karakterek felismerése.

### 13.2. Dokumentumarchiválás

A dokumentum archiválás során nagymennyiségű dokumentumot (befizetési bizonylatot, számlát stb.) kell képként tárolni, illetve visszanyerni. E feladatok elvégzéséhez gyors szkennereket illetve nagy tárolókat használnak. Az alkalmazások egy részénél valójában nem kell, sőt nem is szabad ezeket a dokumentumokat megváltoztatni, hiszen az archiválás célja éppen a dokumentumok hosszú időre történő megőrzése és rendezett formában történő tárolása. Ezekben az esetekben az azonosító adatokat billentyűzetről viszik be a számítógépbe, és a visszakeresés ezek alapján történhet. Ekkor képfeldolgozás szempontjából egyetlen feladat jó: hatékony képtömörítési eljárás kiválasztása és alkalmazása (részletesen ezzel e témakörrel a 16. fejezetben foglalkozunk). Miután itt csak a dokumentum képszerű megjelenítése a cél, ezért az esetek többségében veszteséges képtömörítési eljárások alkalmazása megengedhető. Miután az összes többi probléma nem képfeldolgozással kapcsolatos, ezért ezzel a témával a továbbiakban nem foglalkozunk.

### 13.3. Karakterfelismerés

Abban az esetben, ha a digitalizált dokumentum további feldolgozására is szükség van, akkor meg kell oldani a karakterek felismerését, és a dokumentum karakteres formába transzformálását. Ezt követően lehetőség van könyvek, gépelt szövegek, faxok újraserkesztésére, táblázatok készítésére, kiértékelésére.

A karakterfelismerés egyszerűbb változata az adatlapok feldolgozása. Statisztikai felmérések, kérdőívek kiértékelésekor előre kinyomtatott űrlapokon, meghatározott helyre nyomtatott négyzeteket kell megjelölni. Egy alkalmas program az illető négyzet helyének ismeretében felismeri, hogy van-e rajta jelzés vagy nincs. Ilyen elven működnek a gépi lottószelvény értékelések, népszámláláskor, választáskor, közvélemény-kutatáshoz készített kérdőív kiértékelések. A karakterfelismerés legnehezebb - a mai napig általánosan meg nem oldott - esete a kézzel írt szövegek felismerése, bár speciális esetekre már léteznek megbízható rendszerek (pl. a postai levélirányítás). Megemlítjük hogy, különböző felhasználási területekre egymástól eltérő karakterfelismerési megoldásokat alkalmaznak (például a rendszámfelismerésről a következő fejezetben lesz szó). Ebben a fejezetben, a továbbiakban kifejezetten az irodai alkalmazásokban elterjedt karakterfelismerő programokról lesz szó.

A kinyomtatott vagy papírra írott szövegek digitális úton történő felismeréséhez (optikai karakterfelismerés) az alábbi eszközökre van szükség:

- számítógép (személyi számítógép),
- lapolvasó szkener (általában síkágyas) és
- OCR (Optical Character Recognition) program.

A 25. ábrán végigkísérhetjük az optikai karakterfelismerés egy lehetséges folyamatát (folyamatos határvonalú keret jelzi az egyes folyamatokat, míg a szaggatott határvonalú keretek a folyamatok utáni eredményt szemléltetik). A felismerés tárgyát képező nyomtatott vagy írott szöveget a legjobb minőség elérése érdekében lapolvasó szkener segítségével digitalizáljuk. Eredményül egy bittérképes képet kapunk. A következő lépésben elkülönítjük az egyes leendő karaktereknek megfelelő képrészeket. Eredményül egy bittérképes karakter mátrixot kapunk. Az előfeldolgozás során, eltávolítjuk vagy megpróbáljuk redukálni a számunkra zavaró komponenseket (zaj, üres részek, nem felismerendő részek, stb.). Az egyes mátrixokat normalizáljuk, azaz megfelelő méretre, pozícióra, elhelyezkedésre, stb. transzformáljuk.

A kapott normalizált képen meghatározzuk a felismeréshez szükséges jellemzőket (pl. kétdimenziós morfológiai paraméterek, textúrák, logikai kapcsolatok, stb.). Eredményül olyan halmazt kapunk, amely kapcsolatokat, paramétereket, függvényeket tar-



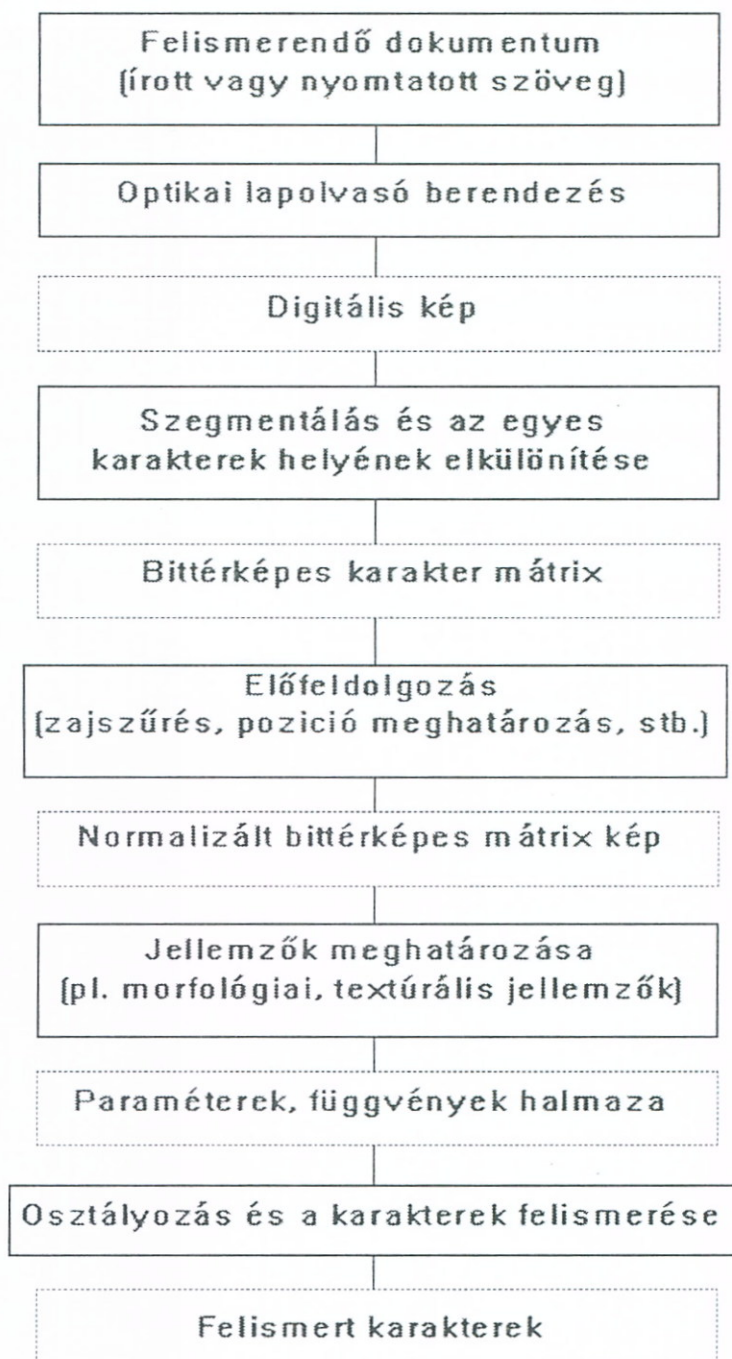
talmaz. Az utolsó lépésben egy adatbázissal történő összehasonlítás alapján megpróbáljuk a paraméterek alapján felismerni (osztályozással) a karaktereket.

A folyamat végrehajtását számos probléma nehezíti, korlátozza:

- Eredeti nyomtatás minősége:
  1. lézernyomtatóval készített szöveg minimális hiba (1-2%) mellett felismerhető,
  2. karbon szalagos írógéppel készített szöveg felismerésekor jelentkező hiba a gyakorlatban még elfogadható,
  3. hagyományos írógéppel készített szöveg felismerése gyakran nagy hibaszázalék mellett valósítható meg.
- Szkenner és vezérlőprogramjának paraméterei - optikai felbontás, beállítási paraméterek, stb.
- OCR program beállításai - nyelv, karakterkészlet, kép orientációja, stb.

A feladat láthatóan rendkívül összetett, ennek ellenére örömmel nyugtázzhatjuk, hogy az OCR programok területén, világviszonylatban az egyik legnagyobb példányszámban eladott és használt programnak (Recognita) magyar cég a kifejlesztője.

A program többféle operációs rendszer alatt is működik, használata egyszerű és magas találati aránnyal dolgozik. Több mint 100 különböző nyelv karakterkészletét tudja kezelni.



25. ábra

Néhány napjainkban forgalmazott karakterfelismerő program:

- Recognita Plus
- OmniPage Pro
- TextBridge Pro
- EasyReader Elite
- Presto!OCR Pro

Az OCR alapú rendszerek működéséhez tesztképeket és részletesebb leírást a CD lemezen találunk (IMAGES/OCR alkönyvtár).

## Ellenőrző kérdések

1. Sorolja fel az automatikus irodai alkalmazásokat!
2. Mik az archiváló rendszerek előnyei a hagyományos dokumentumkezelési eljárásokkal szemben?
3. Mi az űrlap-kiértékelés lényege?
4. Mit nevezünk optikai karakterfelismerésnek ?
5. Ismertesse az optikai karakterfelismerés folyamatát !
6. Milyen problémák jelentkeznek a felismerés során ?
7. Nevezzen meg néhány napjainkban használt OCR programot !
8. Lézernyomtatóval, karbon szalagos írógéppel vagy hagyományos írógéppel írott szöveg ismerhető fel a legjobb eredménnyel ? Miért ?
9. Ön szerint befolyásolja-e a felismerés pontosságát amennyiben a szöveg vízszintes sorai a képen elforgatásra kerülnek? Miért?

## 14. BŰNÜGY- ÉS BIZTONSÁGTECHNIKA

### 14.1. Bevezetés

A bűnügyi és a biztonságtechnikai alkalmazások céljukat tekintve eltérőek, hiszen az előbbi elsősorban a bűntények felderítésére, az utóbbi inkább a megelőzésére szolgál, eszköztárukban azonban sok hasonlót találunk, ez indokolja egy fejezetben történő ismertetésüket. Mindkét területen a képfeldolgozó rendszerek alkalmazása egyre nagyobb szerepet kap a számos egyéb eszköz és rendszer között.

### 14.2. Bűnügyi alkalmazások

A képfeldolgozó rendszereket a rendőrség és a nemzetbiztonsági szervek is széles körben alkalmazzák, részben a bűnügyi szakértői munka segítésére, részben operatív feladatok megoldására. A bűnügyi szakértői munka támogatására kifejlesztett rendszerek - képfeldolgozás szempontjából - leginkább az orvosi, laboratóriumi képek kiértékelésére alkalmazott rendszerekkel mutatnak rokonságot, azzal a különbséggel, hogy ebben az esetben a hiteles mérési, összehasonlítási feladatoknak van nagyobb szerepe, és a képen lévő objektumok statisztikai jellegű kiértékelése és feldolgozása háttérbe szorul. Az operatív felhasználások egy része nem igényel gyors (néhány másodperc alatti) feldolgozást, de vannak olyan feladatok, ahol az ipari rendszereknél megfogalmazott elvárásoknak kell eleget tenni. A bűnügyi alkalmazások közös jellemzője, hogy a képfeldolgozási feladat általában képadatbázisban való kereséssel is párosul. A továbbiakban a teljesség igénye nélkül néhány tipikus alkalmazást mutatunk be röviden.

#### **Kép összehasonlítás és felismerés**

A szakértői munka során visszatérő feladat, hogy a helyszínen talált nyomokat (karcolás, lőfegyver töltényhüvely, cipőtalplenyomat, bélyegző lenyomat, ujj- és tenyér lenyomat stb.) össze kell hasonlítani a feltételezett elkövetési eszközzel készített próbanyomokkal, illetve a feltételezett elkövető személy ujjlenyomataival. Ebben az esetben egy képfeldolgozó rendszer, viszonylag egyszerű képedítelési funkciókkal (nagyítás, kicsinyítés, geometriai transzformációk, képek egymásra úsztatása) és mérési lehetőségekkel már jelentős segítséget nyújt az azonosság vagy különbözőség eldöntéséhez. A PICTRON Kft által fejlesztett SoftScope rendszer ezeket a feladatokat látja el, kiegészítve egy adatbázis kezelővel, amely a kiértékelt képek dokumentálását, archiválását és visszakeresését teszi lehetővé. A rendszer szabvány video jelet tud fo-

gadni, így közvetlen CCD kameráról, videoszalagról, mikroszkópról, vagy speciális, okmány ellenőrző rendszerről is bevihetők a képek. A képek jellegzetes kulcsszavakkal való archiválása akkor jelent nagy segítséget, ha még nincs gyanúsított, és a helyszínen talált nyomok értékelése alapján az adatbázisban keresve ellenőrizhetjük, hogy az adott eszközt használták-e már korábbi bűnügyekben.

Különösen alkalmasak a helyszínen talált ujj- és tenyérlenymatok az elkövető személyének azonosítására, hiszen mint köztudott ezek minden személynél eltérőek. Több (köztük magyar) ujjlenymat felismerő rendszer létezik. Korábban az ujjlenymat globális jellegzetességeit (fodorvonal irányultság, elágazások, hurkok, örvények, boltozat) egyedi vizsgálattal, szakértők kódolták be, és a nyilvántartott személyek adataival együtt tárolták. A kézi feldolgozás rengeteg időt vett igénybe, és soha nem tudott naprakész lenni. Az automatikus, képfeldolgozáson alapuló rendszerek inkább lokális mintákat (villák, triádusok) és fodorszál-számlálást használnak, mivel ezek a jellemzők kevésbé hibaérzékenyek. A kép kiértékelés során jelentős szerepet kapnak a különböző élkeresési és irány meghatározási módszerek, és a jellegzetességek leírására bonyolult kódolási eljárásokat alkalmaznak. Az azonosítás többlépcsős döntési modell alapján, automatikusan történik, így a nyom rögzítése után az adatbázisban történő visszakeresés azonnal megkezdődhet. A megoldást nehezíti, hogy a helyszínen többnyire csak töredék és eltorzított nyomok maradnak, ezért olyan jellegzetességeket kell kiválasztani a felismeréshez, amelyek invariánsak a geometriai torzulásokra, és töredék nyomon is elegendő számban fordulnak elő.

### Okmányhamisítás felderítése

A pénzhamisítás és az okmányhamisítás felderítésére a legegyszerűbb infralámpától a bonyolult számítógépes kép összehasonlító rendszerig számos eszközt fejlesztettek ki. A védekezés az okmány előállításával kezdődik, speciális papírokat, tintákat és jelzéseket alkalmaznak a fontosabb dokumentumok hamisításának megelőzésére, azonban a bűnözők is lépést tartanak a technikai fejlődéssel, és az eredeti dokumentumokhoz nagymértékben hasonlító, szabad szemmel alig, műszerekkel is csak alapos vizsgálatok után felismerhető hamisítványok előállítására képesek. Az ellenük folytatott harcban jelentős szerepet játszanak a digitális képfeldolgozás alapján működő rendszerek.

A legfejlettebb rendszerek **funkcionálisan két részből** állnak.

Az **egyik a hagyományos optikai rész**, amelynek feladata, hogy a megvilágítási viszonyok illetve az érzékelt jelek, leginkább az okmányon, bankjegyen lévő ábrák, biztonsági jelek anyagi jellemzői alapján kihangsúlyozottan felismerhetőek legyenek. Egy ilyen rendszerben a vizsgálat azon alapszik, hogy a megvilágított objektumokról

visszaverődő vagy azokon áthaladó fény hullámhosszostályait vizsgáljuk, és az így nyert képekből olyan többletinformációkhoz jutunk, amiket egyszerű vizuális megfigyeléssel a látható hullámhossz tartományban nem tudunk értelmezni.

A **másik rész, a számítógépes képelemző** lehetővé teszi a számszerű méréseket, belső összefüggéseket feltáró elemzéseket, a korábbi minták archiválását és a dokumentumok összehasonlító elemzését.

Az **optikai rész** egy nagy érzékenységű a 400-1100 nm-es spektrumtartományban látó monokróm CCD TV kamerából és a köré épített fényforrásokból és színszűrő rendszerből áll. Az okmányokon, bankjegyeken egyéb vizsgálandó objektumokon lévő látható és nem látható jelek különböző keskeny spektrum tartományba eső fényre érzékenyek. A megfelelő hullámhossz-tartományú fény kiválasztását különböző spektrumú lámpák, lámpaszűrők és kamera szűrők kombinációjával biztosítják. A wolfram szálalás halogén izzó az infravörös lumineszcencia tanulmányozásához, a ráeső infravörös megvilágítás abszorpciós és reflexiós vizsgálatokhoz nyújt segítséget. Az átmenő fehér és infravörös megvilágítással vízjelek, elfedett vagy törölt információk deríthetők fel. A ráeső ultraibolya fényben lumineszcencia biztonsági jelzések és a papír fluoreszcenciája vizsgálható.

A halogén fényforrás elé forgatható szűrőváltó tárcsát helyeznek el, ezzel szűkítve a 400-700 nm spektrumszélességű fényforrás spektrumtartományát. A szűrőváltó optimalizálja a kép kontrasztosságát minimalizálva a háttér lumineszcenciáját.

A vizsgált objektumot egy sokoldalú monokróm monitor jeleníti meg. A monitor alkalmas a kép tükrözésére, valamint negált kép megjelenítésére is.

A további felhasználói követelmények kielégítésére nagy erejű villanócsöves megvilágítás is rendelkezésre áll.

A **számítógépes képelemző legfontosabb eleme az integráló egység.** Az integráló egy speciális digitalizáló, amely a digitalizálás mellett arra is alkalmas, hogy maximum 20 másodpercig összegezze a képeket ezzel 500-szorosra fokozva a berendezés érzékenységét. Az integrátor használata nagymértékben megkönnyíti fluoreszkáló tintával írt írásjelek tinta szerinti megkülönböztetését. Az integrátor segítségével lehetőség van arra is, hogy a megjelenítő monitoron megfelezzük a képet akár vízszintesen akár függőlegesen. Az egyik félkép tartalmazhat egy rögzített úgynevezett etalon képet, míg a másik félkép egy aktuális képet. Megfelelő nagyítással és geometriai elhelyezéssel a két különböző objektum egymáshoz illeszthető, és a vizsgált elemeinek - például szálak vagy más azonosító jelek - folytonossága ellenőrizhető. Az integrátor videokimenetére képmagnetofon vagy más video rögzítő eszköz csatlakoztatható. A készüléket számítógéphez illesztő felület és számítógépes kiértékelő programok teszik teljessé.

A rendszer felhasználási területei közül kiemelendő többek között, hogy nagy biztonsággal kimutatható a dokumentumok **hamissága**, illetve **megváltoztatottsága**; a

törlések illetve **elfedések**; **vízjelek**, a **festék-összetétel** különbözősége, a **papír egyenetlensége**. Így a készülék különösen jól használható **útlevelek**, **vízumok** illetve **bankjegyek** vizsgálatára, valamint **általános okiratelemzésre** is.

A vizsgálatok során fontos információt nyújt bizonyos alakzatok megléte vagy hiánya. Egy gyanús okmány (útlevelek, személyi igazolványok, gépjárművezetői engedélyek) vizsgálatánál adatbázisbeli okmányminták etalonként felhasználhatók. Szabad szemmel nem látható festékekkel nyomtatott sorszámok, ábrák megkeresése szintén a vizsgálat tárgya. A papír anyagának spektrális vizsgálata során eldönthető, hogy a kérdéses pénzt megfelelő „gyári” vagy más minőségű papírra nyomták. Érdekes lehet a vízjel, illetőleg a biztonsági szálak vizsgálata. Meghatározott hullámhossz-tartományú átvilágító fényben készített felvételeken jól kirajzolódnak a vízjel ábra elemei, amelyek folytonosság vizsgálata, tartalmi értékelése és geometriai hasonlósága szerint eldönthető eredetisége. A biztonsági szálak anyaga, mérete és elhelyezkedése szintén jó információ az eredetiség megállapításához.

Útlevelek, személyi igazolványok hamisításánál gyakran előfordul, hogy a lopott okmányt preparálják. Anyaga ugyan eredeti, de a preparálás során különböző vegyszerekkel a hamisítók megpróbálják eltüntetni az eredeti bejegyzéseket, majd új adatokkal helyettesítik. A számítógépes képfeldolgozó eljárás kimutatja az eltávolítás tényét. Fényképek, bélyegzőnyomatok cseréje szintén megállapítható.

### **Fantomkép összeállítás**

A rendőrségi munka során gyakran van szükség a sértett vagy a szemtanuk elmondása alapján összeállított fantomkép elkészítésére. Korábban ez kizárólag portré rajzoló bevonásával történt, később megjelentek a speciális diavetítő felhasználásával, diákon tárolt arcképek egymásra vetítésével működő optikai rendszerek. Hátrányuk volt, hogy az arcrészleteket csak vízszintes csíkonként lehetett cserélni és a kezelés lassú és körülményes volt. Az elmúlt tíz évben megjelent számítógépes rendszerek lényegesen rugalmasabbak és gyorsabbak, kezelésük egyszerű. A különböző lehetséges arcelemeket azonnal megjelenítik, így a szemtanú választása gyorsabb és megbízhatóbb lesz. Megfelelő geometriai transzformációk segítségével gyakorlatilag tetszőleges alakú és különböző méretű arcból származó részleteket össze lehet illeszteni. Képjavitási eljárásokkal a különböző bőrszín árnyalatokat lehet korrigálni, grafikus funkciókkal a képet lehet retusálni. A fentiekre is találunk példát az ALKALMAZ/FANTOM könyvtárban.

### **Rendszám felismerés**

Sok tekintetben az ipari képfeldolgozó rendszerekkel azonos elvárásoknak kell ele-

get tennie az automatikus rendszám felismerő rendszereknek. Első pillanatban azt gondolhatnánk, hogy a feladat megoldható valamelyik irodai felhasználására szánt karakterfelismerő program alkalmazásával vagy annak adaptálásával. A felmerülő speciális problémák miatt azonban más megoldást kell keresni. Gondot okoz

- a gépkocsi mozgása,
- a különböző méretű és alakú rendszámtáblák,
- a rendszámtábla változó elhelyezkedése,
- a felvételi helytől függő geometriai torzulás,
- a rossz kontraszt,
- a piszkos vagy sérült rendszámtábla.

A sebesség miatt gyors algoritmusokat kell alkalmazni a rendszámtábla megkeresésére, és a gépkocsi elhaladása alatt rögzíthető néhány kép közül ki kell tudni választani a további feldolgozásra leginkább alkalmasat. A rendszámtáblán belül meg kell találni a rendszámokat és elkülöníteni egyéb információktól (pl. országjelzés). A feldolgozás kétszintes képen történik, a rossz kontraszt azonban nehezíti az optimális vágási küszöb automatikus meghatározását.

A sok probléma mellett szerencsére számos előzetes információnk van (különösen, ha egy országon belüli rendszámtáblákat kell felismerni), ami megkönnyítheti a munkánkat. A magyar rendszámtáblák például egységes fekete kerettel vannak ellátva. Pontosan lehet tudni, hogy a rendszámtáblán belül hol helyezkednek el a számok és betűk, ismert az alkalmazott betű mérete, a típusa.

A feldolgozás során először éldetektálási, képjavítási feladatok jelentkeznek (tábla megtalálása, zaj hatásának csökkentése), majd szegmentálás, végül mintaillesztés és/vagy kontúrelemzés alapján történhet a karakterek azonosítása. A rendszer szerves része a keresett kocsik azonosítóit tartalmazó adatbázis, így a rendőrautóba, vagy háttárakelőhelyekre telepített rendszer automatikus riasztást ad, ha lopott kocsi kerül a kamera elé.

### 14.3. Biztonságtechnikai alkalmazások

A biztonságtechnikai alkalmazások jellemzője, hogy a képfeldolgozást más megoldásokkal kombinálva alkalmazzák. A képfeldolgozás az esetek többségében egyszerűen képtárolást, archiválást jelent, és a számítógép egyéb rendszerek felügyeletét is ellátja. Funkció szerint két nagy csoportba sorolhatók a feladatok, az egyik a beléptetés, a másik a figyelés. Természetesen komplex rendszerek esetén mindkét funkció szerepet kap.

## Beléptető rendszerek

Az automatikus vagy fél-automatikus beléptető rendszerek esetén a **személyekhez kapcsolódó egyedi jeleket** használják fel a biztonság fokozására. Gyakran használt megoldás az ujjlenyomat felhasználása. Ez az alkalmazás annyiban tér el a rendőrségi ujjlenyomat azonosító rendszerektől, hogy itt a beléptetéskor elvárható az együttműködési szándék, így az input kép feltehetően jó minőségű lesz. Ezt elősegíti, hogy optimális megvilágítást, pozicionálást biztosító mintavételi eszközt lehet alkalmazni. Könnyíti a feladatot, hogy korlátozott számú mintán belül (a belépésre jogosultak) kell az azonosítást elvégezni. A nagyobb biztonság kedvéért, ahol őr is van a belépési ponton, az ujjlenyomat alapján az ujjlenyomat tulajdonosának képe meg is jeleníthető a képernyőn. Gyakran alkalmazott megoldás, hogy az egyéb okok miatt is használt azonosító mágnes- vagy chipkártyára írják fel az ujjlenyomatot, ezzel elkerülhető, hogy hamisított vagy ellopott kártyával lépjenek be az őrzött területre. Ebben az esetben a rendszer az aktuális ujjlenyomatot hasonlítja össze a mágneskártyán lévővel. Vannak olyan próbálkozások, hogy a szem íriszének mintázatát használják azonosításra.

Igen gyakran felmerülő probléma a **parkolók, gépkocsi tárolók őrzése, valamint a zárt parkolóknál a belépésre jogosultak automatikus ki és beléptetésének megoldása**. A hagyományosnak mondható kártyás beléptetés mellett egyre gyakrabban alkalmaznak képfeldolgozó rendszereket is erre a feladatra, vagy a kártyás rendszert képfeldolgozással kombinálják. Kézenfekvő a rendszámot használni a gépkocsik azonosítására, a feladat azonban ebben az esetben részben eltérő megoldásokat igényel, mint a bűnügyi alkalmazásnál leírt rendszám felismerés. Könnyítést jelent, hogy a be- és kilépésnél a gépkocsi megáll, ezért a jobb minőségű felvételekre számíthatunk, és több idő van a feldolgozásra. A feldolgozás jellegét megváltoztatja, hogy ebben az esetben nem feltétlenül kell a karaktereket felismerni, hiszen a feladat arra egyszerűsödik, hogy a belépéskor felvett minták között kell megkeresni a kilépő gépkocsiról készített felvételen láthatóval azonosat. Szándékosan nem rendszámot mondtunk, hanem mintát, ugyanis ennél az alkalmazásnál a rendszám ellenőrzése nem mindig elegendő, hanem célszerű a gépkocsi alakját, színét és egyéb azonosítási lehetőségeket is bevonni az ellenőrzésbe, mert ezzel elejét vehetjük, hogy egy értéktelen kocsival beállva a parkolóba, és a rendszámot egy nagy értékű gépkocsira átszerelve kijátsszák a biztonsági rendszert. Az elmondottak miatt, a beléptetési feladatoknál a különböző mintaillesztési megoldások kerülnek előtérbe, és sokszor a rendszám felismerésére nem is kerül sor. Ezt a megoldást az is indokolja, hogy a parkolóházaknál, kempingeknél sok külföldi autóra kell számítani, ami a rendszám felismerését megnehezíti, de minták összehasonlításánál nem okoz gondot. Sokszor ezek a rendszerek kártyás ellenőrzéssel párosulnak, ami például a szolgáltatás díjtételének adminisztrálására is megoldást jelent.



## Figyelő rendszerek

Klasszikusnak számító feladat az objektumok őrzése, ahol számos egyéb megoldás mellett, régóta használnak kamerás figyelést is. Korábban ezt a feladatot kizárólag videotechnikai eszközökkel oldották meg, de napjainkban egyre gyakrabban alkalmaznak számítógéppel vezérelt rendszereket. A számítógép alkalmazása komplett megoldást kínál a kül- és beltéri videó rendszer vezérlésére, a riasztás (tűz, behatolás) és a beléptetés integrált kezelésére, veszélyhelyzetek operatív elhárításának támogatására. Ha a számítógépet csak a rendszer vezérlésére használják, akkor annak nincs köze a képfeldolgozáshoz, de ha a mozgásfigyelés például a digitális képelemzésen alapul, akkor már képfeldolgozásról beszélhetünk. A digitális képen végzett számítógépes mozgásfigyelés lehetőséget ad a **megfigyelt zónák rugalmas programozására**, valamint adaptív (eseményfüggő) figyelő szekvenciák kialakítására. Ezeknél az alkalmazásoknál egyre jelentősebb szerepet kaphatnak az elmúlt években megjelent **intelligens kamerák (kamerával egybeépített számítógép)**, mivel ezek beprogramozhatók előfeldolgozási feladatokra, így a központba, már csak redukált adatmennyiséget kell továbbítani.

A figyelő rendszereknél **fontos szempont a képek tárolása**. Ez részben videókorderekkel, részben digitális úton történhet. Ez utóbbi megoldásnál a különböző képtömörítési eljárásoknak van jelentős szerepe.

## Ellenőrző kérdések

1. Milyen jellegű képfeldolgozási feladatok jelentkeznek a rendőrségi alkalmazásoknál?
2. Miben térnek el az orvosi és a bűnügytechnikai mikroszkópi képkiértékelő rendszerek?
3. Miben közös a távérzékelés és az okmányvizsgálat?
4. Milyen jellegzetességeket alkalmaznak az ujjlenyomat azonosításnál?
5. Milyen képfeldolgozási eljárásokat alkalmaznak a fantomkép összeállító rendszerekben?
6. Milyen speciális problémák jelentkeznek a rendszámfelismerés során?
7. Miben tér el az irodai környezetben és a rendszámfelismerésnél jelentkező karakterfelismerési probléma?
8. Hol jelentkeznek képfeldolgozási problémák a biztonságtechnikában?
9. Milyen speciális problémák jelentkeznek a beléptető rendszereknél?
10. Milyen képfeldolgozási feladatok jelentkeznek a figyelő rendszerekben?

## 15. DIGITÁLIS FÉNYKÉPEZÉS

### 15.1. Bevezetés

Ebben a fejezetben nem egy alkalmazási területről, hanem néhány technikai újdonságról kívánunk szólni, amelyeket tekinthetünk korszerű fényképeszeti megoldásoknak, és ezért az összefoglaló digitális fényképezés címet adtuk a fejezetnek. A képfeldolgozási elméleti témák megértéséhez nem szükséges a digitális kamerák és a fotó CD eszközök ismerete, ezért a 2. fejezetben nem tértünk ki ezekre, de technikai újdonságuk, elterjedésük és alkalmazásuk miatt ebben a fejezetben rövid ismertetést adunk róluk. Ez egyben alkalmas ad néhány egyre általánosabbá váló fogalom megismertetésére, valamint a hagyományos fényképezés és a digitális technika összehasonlítására.

### 15.2. Történeti áttekintés

Több mint 150 évvel ezelőtt (1839) **Daguerre** francia festő és színpadtervező felfedezte az első használható fotográfiai eljárást. Az azóta eltelt idő során hihetetlen fejlődésen ment keresztül a fotográfia, gondoljunk csak a mai kamerákra, laboreljáráásokra, a mozgó film megjelenésére és fejlődésére vagy a holográfiára.

A fényképezés ma még az **ezüsthalogenidekre** épül. A kép előállítását tisztán kémiai folyamat. Mindaz, amit a kamera optikája, elektronikája és mechanikája létrehoz a kémiai kép technikai tökéletesítése csupán.

Az elektronikus képalkotás felfedezőjének a tudomány az amerikai **Ch. Carlson**-t tartja (1938). Valójában azonban az irodalmi adatok alapján kimutatható, hogy **Selényi Pál** magyar fizikus, már jóval 1938 előtt elektrosztatikus úton képeket állított elő. Selényi az elvi alapok mellett, gyakorlati eljárást is kidolgozott és megvalósított. Az ipari megvalósítás részben a mai **Xerox** cég nevéhez fűződik és az 1950-es évekre tehető. A rohamos, főleg reprotchnikai elterjedés 1960 után kezdődött. Napjainkban rendkívül széles a használati eszközeink azon palettája, ahol elektronikus képalkotást használnak: fénymásolás, lézernyomtatók, színes nyomtatók, video- és TV technika, telefax, stb. Vegyük észre, hogy az elektronikus képalkotás önmagában még nem eredményez digitális képet! A korábbi fejezetekben (2., és 3. fejezet) láttuk, hogy az elektronikus, de még analóg képből a digitalizálás során különböző eszközök állítják elő a digitális képet. Más kérdés, hogy napjainkban a különböző berendezésekben sokszor egybeépülnek az elektronikus képet előállító és digitalizáló részek.

Az elektronikus és az ezt követő digitális képalkotásnak számos előnye van a ha-

gyománys fotóeljárásokkal szemben, azonban az elmúlt években a modern ezüsthalogén anyagok érzékenységet, szemcsézettségét és színhűségét olyan mértékben sikerült javítani, hogy napjainkban újabb lendületet kapott az ezüsthalogénides fotográfia. A hagyományos fotóeljárásokat a TV rendszerekkel szokták összevetni és felbontás tekintetében a TV rendszer valóban a filmek felbontási lehetősége alatt marad, de halatlan előnye, hogy a kép, sőt a mozgóképek azonnal rendelkezésre áll, ami még a polaroid technikával sem valósítható meg. A nyomdaipar részére sem elegendő a TV rendszerek felbontása, viszont a nyomdai előkészítésben ma már nélkülözhetetlen a digitális technika alkalmazása, ezért a filmeket nagyfelbontású szkennerekkel digitalizálják. A szakadatlan technikai fejlődés újabb és újabb eszközöket és eljárásokat eredményez. az alábbiakban a digitális kamerákat és a fotó CD-t ismertetjük.

### 15.3. Digitális kamerák működése

A képdigitalizálás fotótechnikai gyakorlatában két lehetőség van a kép három szín-csatornájának létrehozására:

1. A képet egymás után háromszor digitalizálják, külön-külön a fény útjába helyezve a három alapszínnek (piros, zöld, kék) megfelelő színszűrőt. Így „dolgozik” pl. a **Hasselblad DB 4000**-es digitális hátfal.
2. A másik eljárás során minden képponthoz három CCD szenzor tartozik úgy, hogy minden szenzor csak a három alapszín egyikét érzékeli. Ilyen eszköz pl. a **Fuji FUJIX DS-H2**.

Jó minőségű képekhez sávonként és képpontonként legalább 1 byte szükséges. Így előállítható színek száma: 16.777.216! Ez nem azt jelenti, hogy minden digitalizált kép 16.7 millió színt is tartalmaz. Valójában nem adható vissza több különböző színárnyalat, mint ahány képpontból a kép áll. Egy foto CD-n tárolt kép pl. 6.3 millió pixelből áll, ami maximálisan 6.3 millió színt enged meg. A színelbontás nem annyira a sok színárnyalat megkülönböztetésére (szemünk közel egymillió színt tud megkülönböztetni), mint az észrevehető árnyalatugrások elkerülésére szolgál.

Elsőként a japán Sony Co. 1981-ben mutatta be a **MAVICA** (Magnetic Video Camera) nevű, elektronikus állóképrögzítő készülékét. Ennek külső formája egy kis-filmes fényképezőgépre hasonlított. A képeket 2 inches mágneslemezen rögzítették, és 50 képet lehetett készíteni egy lemezre.

A nagy világcégek - okulva a különféle videomagnó-rendszerek versengéséből származó nehézségekből - 1984. májusában megállapodtak egy egységes mágneses fotó világszabványban. Ez a **Still Video Floppy (SVF)** egy 47 mm átmérőjű, mágnesezhető korong, amelyre 25 állókép „vehető fel”. Az SVF lemezek kapacitása 740 kbyte és 1.44 Mbyte (hagyományos floppy lemez). Minden egyes új kép felvételekor

a már rögzített kép mágneses jelét az írószerkezet vagy törli, vagy átugorja a floppy egységen. Ilyen kamerákat a SONY cég jelenleg is forgalmaz.

Napjaink populáris digitális kamerai felépítésüket tekintve részben hasonlóak a hagyományos fényképezőgépekhez, részben attól teljesen eltérő megvalósításokat tartalmaznak. A leképezés hasonló a hagyományos fényképezőgépekhez, azaz lencserendszer (általában autófókuszos) vetíti a képet a CCD érzékelőre. Eztán megtörténik a CCD vagy CMOS képpontok kiolvasása, digitalizálása. A kamerát általában saját állandó memóriával ellátott mikroprocesszor vezérli. A processzor legtöbbször elvégzi a kép korrekcióját (leképező rendszer okozta hiba javítása, automatikus fehéregyensúly állítás, gammaérték állítás, stb.), majd tömörítését és tárolását. A tömörítés mértéke (1:1 vagy 10:1-20:1) és típusa (veszteségmentes vagy veszteséges) minőségi faktor beállításával lehetséges (Kiváló, Jó, Elfogadható). A veszteségmentesen tömörített képek általában TIFF formátumúak. A veszteségesen tömörített képek általában szabvány JPEG vagy FLASHPIX, néha egyedi formátumúak (lásd vizuális adatformátumok fejezet). Gyakran lehetséges MPEG-1 formátumú videó részletek mentése is. A kép mentésével a felvételi körülmények is (dátum, expozíció, vakuhasználat, stb.) tárolásra kerülnek.

A CCD vagy CMOS érzékelőhöz a beépített vakú jellemzői illeszkednek, amely eltérő a hagyományos vakuktól. Az elkészített képek egy LCD képernyőn közvetlenül megtekinthetők, módosíthatók esetleg törölhetők. A fényképezőgép közvetlenül csatlakoztatható szabvány felületen (USB, IEEE1394, IrDA, Soros:RS-232) számítógéphez és/vagy TV készülékhez és/vagy nyomtatóhoz. A kamerák felbontása általában 1 - 2,3 Megapixel körüli érték, speciális alkalmazások esetén 6 Megapixel is lehet. Az érzékelő és a beépített szolgáltatások erősen ár-függőek.

Lényeges eltérés, hogy a képek olyan memória egységre kerülnek mentésre, amely írható-olvasható, gyors elérésű, szabvány háttértárolóként is használható:

- CompactFlash (leggyakoribb),
- Floppy lemez (kevésbé használatos),
- SmartMedia (gyakori),
- microdrive (profi eszközökben).

Fényképezési szempontból napjaink legérdekesebb képdigitalizáló eszköze a **kamera-szkenner**. Ez egyfajta keveréke a hagyományos fényképezőgépnek és az elektronikus szkennernek. Kameraként a szériában gyártott közép- vagy nagyformátumú gépvázakat alkalmazzák. A szkenner egy speciális tartórészbe került, ami a kamerára helyezhető a filmtároló kazetta helyére illeszthető. A szkenner közvetlenül a filmsíkban tapogatja le a kamera objektívje által leképzett képet. Ez a képletapogatási folyamat több percig is tarthat. Általában **progresszív** azaz soronkénti kiolvasással történik a digitális kép létrehozása. Az eszközök felbontása figyelemreméltó: akár 30 millió pi-

xel is lehet. Érdekességként jegyezzük meg, hogy a mai korszerű diafilmek területegységre vetített képpontszáma hasonló. A 30 milliós képpontszám kb. A3-as méretig problémamentes képkészítést biztosít.

A kamerák általános működési elveit, valamint a jelenleg használt érzékelők (CCD, CMOS) alapvető jellemzőit az Olvasó a „Képfeldolgozás eszközei” című fejezetben találja.

A 26. ábrán összefoglaltuk a hagyományos és a digitális fényképezőgépek alapvető jellemzőit:

<b>Jellemző</b>	<b>Tükörreflexes kamerák</b>	<b>Hagyományos digitális kamerák</b>
Optikai leképezés	Kiváló minőségű lencserendszerrel	Felbontáshoz illeszkedő, jó minőségű lencserendszerrel
Élességállítás	Autófókusz	Autófókusz
Zoom	Van	Van
Vaku	Beépített, külső	Beépített
Felvételek tárolása	Ezüst-halogenid alapú	Mágneses táruk
Felvételek kapacitása általában	36 felvétel/tekerecs	4-16 MB/tár kb. 15-40 nagyfelbontású tömörített kép
Tárolás formátumai	-	TIFF, BMP, JPEG, PLASHPIX, MPEG-1
Energiaellátás	Szárazelem, akku	Szárazelem, akku
Felvételek közvetlen ellenőrzésének lehetősége	Nincs	LCD kijelzőn
Nem vizuális adatok tárolása	Részben	Teljesen automatikus
Hangadatok tárolásának lehetősége	Nincs	Van
Minőség	Jelenleg a legjobb	Hagyományost közelítő
Csatlakozás személyi számítógépekhez	Nincs	Többféleképpen
Tartozékok	Teljes körű	Szegényes

**26. ábra**

## 15.4. A fotó CD

Az elmúlt években vált ismertté a fotó CD rendszer kidolgozója a Kodak cég volt. Ez egy átmenet a hagyományos és digitális fényképezés között. Lehetőséget ad, hogy olcsón, jó minőségű, nagyfelbontású elektronikus képekhez jussunk anélkül, hogy ehhez drága digitális eszközöket vásárolnánk. Maga a fényképezés továbbra is a szokásos módon, a megszokott kamerákkal, a lehető legmagasabb felbontást biztosító film-anyagra történik. A kép csak ezután kerül digitalizálásra.

A rendszer négy különböző összetevőből áll:

1. **diaszkenner**,
2. nagy teljesítményű **számítógép** (az összes folyamat irányítására és ellenőrzésére),
3. speciális **CD író**,
4. **thermoprinter**, amellyel kiváló minőségű, közel A/4-es színes képek nyomtathatók.

A szkener alatt a film vagy dia soronként kerül letapogatásra, a felbontás általában 2048 sor, soronként 3072 elemmel. A teljes képfelbontás során 6,3 millió képpont keletkezik sávonként. A színmélység sávonként 8 bit. Egy fotó CD kép teljes - nem tömörített - adatmennyisége közel 18 MB. Ez jó közelítéssel egy 24x36 mm-es hagyományos negatív film minőségének felel meg. A szkennelés után az adatokat utófeldolgozzák (pl. színeskorrekció). Speciális, **veszteségmentes** eljárással tömörítik, majd fotó CD-re írják. A **lemezfelíró** eszköz mechanikai felépítésében az ismert lejátszó készülékekre hasonlít. A döntő különbség: a lézersugár, amely itt nem kiolvasásra, hanem az információ beírására szolgál. A fotó CD lényeges ismertetője egy lakkréteggel fedett leheletvékony ezüst- vagy aranszínű filmréteg. Felíráskor a lézersugár apró lyukakat éget a lakkrétegbe, és így az alatta lévő fémréteg szabaddá válik.

A Kodak a fotó CD-re való adatfeldolgozásnál a kép szkennelése után kapott adatokat két információs blokkra bontja. Az első blokk csak a világosságinformációkat tartalmazza. Ez az úgynevezett **luminancia** jelcsomag, tulajdonképpen egy tiszta fekete-fehér képnek felel meg. A második, krominanciának nevezett blokk csak a színértékeket tartalmazza. A **krominancia** jel az alapszíneknek (vörös, zöld és kék) megfelelően három részből áll. Így összesen négy információs csatorna jön létre:

- luminancia (Y)
- krominancia-vörös (R)
- krominancia-zöld (G)
- krominancia-kék (B).

Mivel a luminancia a három krominancia-jel összegét tartalmazza, a négy csatornából egy meghatározható. Gyakorlatilag a következő három csatornát tárolják:

- $Y$  (luminancia) =  $R + G + B$
- $C1$  (krominancia 1) =  $R - Y$
- $C2$  (krominancia 2) =  $B - Y$

Mindhárom krominancia-rész egy egyszerű átrendezéssel rekonstruálható:

- $R = C1 + Y$
- $B = C2 + Y$
- $G = -(C1 + C2) - Y$ .

Ez az YCC-nek nevezett kódrendszer tulajdonképpen a televíziós szintovábbítás rendszerének felel meg. Az **YCC-kódrendszer** döntő előnye a luminancia (fényesség) és a krominancia (szín) teljes szétválasztásából adódik. Ez olyan adatredukciót tesz lehetővé, ami más színrendszereknél nem lehetséges.

A módszer azt a tényt használja ki, hogy az emberi szem a világosságkülönbségeket lényegesen jobban megkülönbözteti egymástól, mint a színárnyalatokat. A luminancia-adatok teljes egészében átvitelre kerülnek, a krominanciaértékeknel azonban a látásunk tulajdonságain alapuló adattömörítést hajtanak végre. Más szóval: szigorúan technikai szempontból a szem egy kis csalás áldozata lesz, de csak annyira, hogy senki ne vegye azt észre.

A tárolóhely jobb kihasználtságához az így kapott adatokat még tovább tömörítik. Az itt alkalmazott módszer megfelel az adatfeldolgozásban elterjedt eljárásoknak. Az **adatredukció** segítségével az eredeti 18 MB-nyi adatmennyiséget mindegy 3,5-6 MB-ra sikerült csökkenteni. Így egy fotó CD lemezen legalább 100 (száz) színes kép helyezhető el.

A fotó CD-n minden kép hat különböző változatban kerül tárolásra. Mind a hat képet ugyanabból az adatállományból nyerik, ezek csupán felbontásban különböznek egymástól. Az alapfelbontás - amit a **Kodak Base**-nak (alapnak) nevez - 512 sornak (soronként 768 képpontnak) felel meg, ami egy normál tv-kép minőségével közel azonos. Emellett találunk még 4-szer, illetve 16-szor kisebb, valamint 4-szer, 16-szor (kisfilmnél nagyobb film esetén 64-szer) nagyobb felbontású képváltozatokat is:

- Base/16            128 x 192   képpont
- Base/4            256 x 384   képpont
- Base                512 x 768   képpont
- 4Base              1024 x 1536 képpont
- 16Base            2048 x 3072 képpont
- 64Base            4096 x 6144 képpont

A kisebb felbontások csupán a képek kikeresésére, valamint az úgynevezett **index**-képek (tartalomjegyzék) kinyomtatására szolgálnak. Ezek a képek tömörítés nélkül kerülnek a lemezre, ezáltal gyorsabban jeleníthetők meg a képernyőn. A nagyfelbontású 4 Base, 16 Base és 64 Base változatokat tömörítik. Ezek az adatcsomagok nem tartalmazzák a teljes kép adatait, hanem csak azokat a részeket, amelyek nem található meg az alacsonyabb felbontású adatkészletekben. Ha tehát valaki 2048x3072 pontos felbontású képet szeretne kinyomtatni, akkor ahhoz nemcsak a 16 Base adatait kell leolvasnia a fotó CD-ről, hanem a 4 Base és a Base adatait is.

A fotó CD-n tárolt digitalizált képek bármely, a kereskedelemben kapható televízióon megjeleníthetők. A fotó CD minden jelenleg ismert és használt televíziós normával (PAL, SECAM, NTSC), valamint a megszületés fázisában lévő HDTV normával is kompatibilis. Ehhez szükség van egy fotó CD lejátszóra vagy egy Philips CD-I lejátszóra, amelyekkel egyébként a szokványos audio CD-k is lejátszhatók. Az **asztali fotó CD lejátszók** mellett léteznek **hordozható playerek** is. Ezek a készülékek új prezentációs lehetőségeket rejtenek. A CD-n található képeket a lejátszóval tetszőleges sorrendben állíthatjuk össze és mutathatjuk be. Még a képkivágásokon is változtathatunk, és ezt a „műsort” tárolhatjuk is.

## 15.5. Szkennerek használata

Az alábbiakban ismertetjük a szkennelés gyakorlata során felmerülő legfontosabb fogalmakat, jellemzőket, paramétereket. A fejezet végén hasznos tanácsok szerepelnek, amelyek a gyakorlati munkát hivatottak segíteni.

A kép érzékelő eszközök osztályozásával – ide tartoznak a szkennerek is - „A képfeldolgozás eszközei” fejezetben már részletesen foglalkoztunk. Amennyiben az olvasó önállóan kívánja feldolgozni a szkennerekre vonatkozó ismereteket, kérjük ismétlje át az említett fejezet ide tartozó részeit.

Az asztali vagy síkágys szkennerek elterjedése hihetetlenül megnőtt az elmúlt években, köszönhetően az alacsony árnak. Így bárki hozzájuthat egy otthoni használatra alkalmas eszközhöz, melyek kiváló műszaki paraméterekkel rendelkeznek és a legtöbb feladatnak megfelelnek.

A szkennelés tárgya alapján két nagy csoportra bonthatók az eszközök, bár napjainkban egyre gyakrabban találkozhatunk olyan szkennerekkel, amely mindkét feladatnak megfelel:

1. ránézeti (cél tárgyról visszavert fényt érzékelő) vagy
2. átnézeti (cél tárgy által áteresztett fényt érzékelő) eszközök.



Az érzékelő fej tartalmazza a CCD vagy CMOS érzékelők mellett, a mesterséges megvilágítást adó integrált hidegfényű fényforrásokat is. Ezt a fejet egy fejmozgató mechanika az érzékelők geometriai kialakítására merőlegesen mozgatja (lásd a CD képfeldolgozás eszközei fejezetét). Az alábbiakban ismertetjük azokat a paramétereket, amelyekkel a felhasználó találkozhat egy síkágvas szkennert használata során. Nem foglalkozunk azonban az eszközök üzembe helyezésével.

A szkennert *üzemmódjának beállítása* során a megadhatjuk – amennyiben a szkennert többfunkciós –, hogy visszavert fényt érzékelünk vagy áteresztett fény alapján kívánunk képi adatokat nyerni, így általában három lehetőség közül választhatunk:

1. visszavert (reflective)
2. negatív (negative)
3. dia (slide)

A digitalizált képi adatok céltárgytól és a felhasználó választásától függően adott *színtér*be kerül. Ezen paramétert a színtér vagy színérzékelés esetén választhatjuk ki (RGB, CMYK, Szürke, Színtrevágott).

Az eszköz felbontóképességnek tárgyalásakor a gyártók a technikai adatoknál illetve az eszközt működtető meghajtók esetén alapvetően két paramétert adnak meg. *Optikai vagy fizikai felbontóképesség* (pl. 600x300 dot per inch) esetén az első paraméter a vonali érzékelőbe integrált detektorok számát adja inchenként, míg a második paraméter függ, a detektorok szélességétől valamint a mozgató mechanika lehetőségeitől. A *szoftveres vagy kimenő felbontóképesség* pedig a ténylegesen digitalizált kép felbontóképessége dpi-ben. Itt a gyártók a műszaki paraméterek között is adnak meg általában a fizikainál lényegesen magasabb (1 nagyságrend) értékeket, valószínűleg üzleti fogásként. Ez már egyfajta interpoláció során kapható meg, ahol a fizikai jellemzők mellett lényeges az interpolációs függvény is (erről részletesen az „Újramintavételezés” fejezetben olvashatunk). A kép valós élességének jellemzésére az ún. *MTF* (Modulation Transfer Function) *moduláció átviteli függvényt* használható. Ez mérések alapján, az egyes érzékelési frekvenciák (RGB) esetén, a szkennert által adott felbontást mutatja. Mindez tartalmazza már a teljes optikai rendszer (érzékelési frekvenciánként eltérő, pl. 60 % zöld, 30 % vörös és 10 % kék), a mechanika és az elektronika együttes működését.

A *szkennelési méret* az egy művelettel beszkennelhető maximális terület (lapadagolós modell esetén nincs hosszbeli korlátozás, ekkor csak szélességi adat található) nagysága. Egy A4-es síkágvas szkennert esetén ez az érték 210x297 mm x mm.

A gyakorlatban használt korszerű eszközök egyetlen lépésben tapogattják le mindhárom érzékelési frekvenciának, színek (RGB) megfelelő információt úgy, hogy az érzékelő fejbe épített fehérfényű fényforrás tárgyról visszavert vagy átengedett fényét

egy prizma a három alapszínnek megfelelő színre (vörös, zöld, kék - RGB) bontja, majd egy lencserendszer a három érzékelő sorra vetíti az egyes vonali képeket.

Ehhez paraméterként a *színmélység vagy színfelbontás* tartozik, amely bit-ben kifejezett értéke a maximálisan érzékelt színnek. Általában 8, 10, 12 bit-es eszközökkel találkozhatunk. Így egy 12 bit felbontású eszközzel elméletileg maximálisan  $2^{12 \times 3} = 68\,719\,476\,736$  különböző szín érzékelhető (ilyen kép található a CD lemez images/formatum könyvtárában toll48bit.tif név alatt, a kép 12bit/csatorna kapacitású filmszkennerrel került digitalizálva, majd 16bit/csatorna alapú TIFF formátumban mentésre). A szkennereket működtető programok zöme azonban, egy ilyen nagy színfelbontású képet 24 bitessé alakít. Miért van akkor szükség, ennél nagyobb felbontásra? A gyakorlatban épített érzékelők adatait mindig terheli elektronikus zaj, amely elsősorban az A/D átalakítás során jelentkezik. Ez elsősorban a sötétebb részeken jelentkezik erőteljesebben (digitális kameráknál ez kevés fény esetén kiugró), mivel itt alacsony az érzékelők érzékenysége. Ezért célszerű nagyobb színfelbontást választani, a valóság hű kép készítése érdekében.

A *denzitás* a képen található legvilágosabb és legsötétebb pontok közötti árnyalati különbség tízes alapú logaritmus. A színes diafelvétel denzitása legalább 3,5, egy 10 bites színes szkennerek denzitása maximálisan 3,0, egy sávonként 8 bites digitális kép denzitása maximálisan 2,4 míg egy tintasugaras színes nyomtatásé 1,5-2,0. Mindez jól mutatja, hogy csökkent az árnyalatok átvitele a számítógépes átalakítással.

Természetesen egy 8 bitnél nagyobb színfelbontású szkennerek használatakor – általában előre megadott függvény alapján – az A/D átalakító által szolgáltatott, 8 bitnél magasabb sávonkénti értékek közül - kiválaszthatunk előre megadott módon 8 bitet azaz 256 különböző értéket. Sötét tónusú fénykép szkennelése esetén kapott 10 bit-nyi sávonkénti értékből, többet tartunk meg az alacsonyabbak közül, mint a világosak esetén (Gyakorlatban mindez történhet úgy, hogy az előszkennelt képen egérrel megadjuk a legsötétebb és az általunk legvilágosabbnak vélt pontokat, majd újbóli szkenneléssel az eszköz már ezen tartományokon belül választja ki a tárolásra alkalmas 3x8 bitet, holott mindkét esetben az érzékelők 10 biten dolgoztak!).

A fenti jellemzők gyakran grafikusán egy ún. *tónusgörbe* megadásával változtathatók a szkennelés során. Ezen görbe vízszintes koordinátájára a bemenő értékeket visszük fel (általában színsávonként), a függőleges tengelyre pedig a kimenő értékeket (szintén sávonként). A tankönyv „Képjavítás” fejezetében ezzel részletesen is foglalkoztunk. Itt gyakorlatok is találhatóak, amelyek a már digitalizált képen jól szemléltetik, hogy mindezek a szkennelés művelete előtt is elvégezhetőek - amennyiben a meghajtó szoftver támogatja – ezáltal kihasználva az eszköz fizikai lehetőségeit. Gyakran találkozhatunk a *világosság* vagy a *kontraszt* külön-külön történő megváltoztatásának a lehetőségével is.

Egy másik gyakran használt paraméter, amely még a denzitás fogalmával áll kapcsolatban a **gamma-érték** fogalma. Ez a tónusgörbe (egyenes szakaszának) a meredeksége. Amennyiben ez az érték 1 alatti, a kimenő képem lágyabb tónusú lesz a bemenőnél. Egynél nagyobb érték esetén kontrasztos, kemény tónusú képet kapunk. Monitorokon történő megjelenítés során általában 1-nél nagyobb gamma-értéket használunk, amelynek beállítása – illesztése a felhasználóhoz – egy előzetes monitor kalibrációval történik (pl. Adobe Gamma).

A fenti paraméterek mellett számos egyéb jellemző is megadható:

- a mentett **fájl formátuma** – lásd „Vizuális adatformátumok” fejezet,
- **paletta** – csak palettás képek esetén,
- **dithering** („Képi másolat készítése” alfejezet) – szürke árnyalatos szkennelés esetén,
- **screening** – a ditheringhez hasonló cella megadása,
- **filmtípus** – síkágyas szkennerek esetén filmek típusának (negatív vagy dia) megadása,
- **filmtípus regisztrálás** – mérés alapján meghatározzuk a szkennelendő színes negatív film típusát, stb.

Az alábbiakban összefoglaljuk a szkennelés folyamatának jellemző lépéseit feltételezve, hogy a hardver és szoftver eszközök működőképes állapotban vannak:

1. Első lépésként vizsgáljuk meg a céltárgyat, állapítsuk meg, hogy a rendelkezésre álló eszközzel elvégezhetjük-e a szkennelést (pl. színes negatív szkennelése egyszerű síkágyas lapszkenneléssel nem lehetséges, stb.).
2. Helyezzük el megfelelően a szkennerbe a céltárgyat – pl. tisztítsuk meg a szkennelő üveglapját, használjuk a gyári diatartót, stb!
3. Ellenőrizzük gépünk és a szkennelő közötti megfelelő kapcsolatot - általában szoftveresen!
4. Állítsuk be, hogy visszavert fényel egy objektumot, diát vagy negatívot szkennelünk!
5. Készítsünk egy próba képet, egyszerű, alapbeállítások alkalmazásával! Szükség esetén készítsünk további próba szkenneléseket eltérő beállítások mellett.
6. Jelöljük ki a próba képen a szkennelni kívánt területet!
7. Állítsuk be a geometriai felbontást a kimeneti kép méretének figyelembevételével! Ügyeljünk a fizikai és az interpolált felbontás esetleges különbözőségére.
8. Állítsuk be a színteret és a színfelbontást a szkennelt kép további felhasználásának figyelembevételével (Internet – kis geometriai felbontás, jó színfelbontás, RGB színtér; Kiadványszerkesztés - nagy geometriai felbontás, jó színfelbontás).

- színtér; Multimédia alkalmazások – közepes vagy nagy geometriai felbontás, jó színelbontás, RGB színtér, stb.)!
9. Állítsuk be a denzitással kapcsolatos paramétereket – tónusgörbe, világosság, élesség, gamma-érték, kontraszt, stb!
10. Amennyiben szükséges a „Vizuális adatformátumok” fejezetben ismertetett szempontok alapján válasszuk ki a megfelelő fájlformátumot!
11. Esetleges további paramétereket, amennyiben nincs közvetett módon szükségünk az alkalmazásukra alapbeállítás mellett használjuk vagy kapcsoljuk ki őket (pl. zajszűrők, effekt-szűrők, stb.)!
12. Végezzük el a szkennelést!
13. Majd ellenőrizzük - amennyiben ez lehetséges – részletnagyítást, hisztogrammot alkalmazva, hogy elértük-e a kívánt eredményt.!
14. Szükség esetén ismételjük meg az előző lépéseket!
15. Helyesnek vélt kép esetén végezzük el a háttértárolóra történő mentést, csak ezután folytassuk tovább a munkát!

A mentéssel kapcsolatban az alábbiakat célszerű meggondolni:

- archiválási, munka közbeni (veszteségmentes belső formátum), további fel-
- dolgozás céljából kívánjuk-e állományunkat tárolni,
- milyen tárolási kapacitás áll rendelkezésünkre (szükséges-e adattömörítés,
- ha igen mi az ami megengedett?),
- milyen eszközön kerül megjelenítésre adatunk (nyomtatás, vagy monitor)?

A fentiek, a „Vizuális adatformátumok” valamint a „Képkódolás és tömörítés” fejezetben leírtak alapján célszerű kiválasztani a tárolási formátumot.

Az alábbi táblázatban összefoglaltuk, hogy ismert cél esetén általában milyen jellemzőkkel végezhetünk szkennelést! A táblázat tájékoztató jellegű adatokat tartalmaz,

Cél	Szkenner típusa	Formátum	Geometriai felbontás	Színelbontás csatornánként
DTP	Síkágyas, dob, film	Veszteségmentes	Nagy	10-12 bit
Internet	Tetszőleges	Veszteséges	Alacsony	8 bit
Multimédia	Síkágyas, film	Veszteségmentes	Közepes	8-10 bit
Légifelvétel kutatói célokra	Síkágyas, dob, film	Veszteségmentes	Rendkívül nagy	10-16 bit
OCR	Síkágyas, dob	Veszteségmentes	Közepes	8-10 bit, elég egyetlen szürke csatorna
Alakfelismerés	Síkágyas, dob, film	Veszteségmentes	Nagy	8-10 bit

univerzális megoldás minden esetre nem adható!

### 27. ábra

## Ellenőrző kérdések

1. Ki fedezte fel az első használható fotográfiai eljárást?
2. Mi az elektronikus fotográfia lényege?
3. Mi az elektronikus képkészítés előnye és hátránya a hagyományos eljárással szemben?
4. Mire szolgál a fotó CD magas színfelbontása?
5. Hogyan épül fel egy digitális fényképezőgép?
6. Milyen adatformátumokat használnak a digitális kamerák?
7. Milyen eszközöket használnak a digitális kamerák esetén, felvételek tárolására?
8. Milyen előnyökkel rendelkeznek a digitális kamerák?
9. Mit nevezünk kamera szkennereknek?
10. Mi a fotó CD rendszer lényege?
11. Milyen eszközökből áll egy fotó CD rendszer?
12. Mekkora a fotó CD színmélysége?
13. Milyen eljárással tömörítik a fotó CD képet?
14. Mit tartalmaznak a fotó CD luminancia, krominancia jelcsomagok?
15. Mi az YCC színrendszer és hol használják?
16. Mekkora lenne a fotó CD-n, egy tömörítetlen teljes kép tárigénye?
17. Adjuk meg a szkennelés tárgya alapján a szkennerek csoportosítását!
18. Milyen üzemmód beállításokat ismer síkágyas lapszkennerek esetén?
19. Mit rejt egy síkágyas lapszkennerek érzékelő feje?
20. Mit takar az optikai felbontóképesség szkennerek esetén?
21. Mi az MTF?
22. Mi a kimenő vagy szoftveres felbontóképesség?
23. Milyen tipikus színfelbontásokat alkalmaznak szkennereket gyártók?
24. Mi a denzitás?
25. Mit nevezünk tónusgörbének?
26. Mi a gamma-érték?
27. Ismertesse a szkennelés folyamatának jellemző lépéseit!
28. Milyen paraméterekkel készítené eredeti (negatív) légifelvételekről szkennelést?
29. Miért elfogadható az Internetre készített képek zöménél a veszteséges szkennelés?

30. Milyen tipikus színtereket használhatunk a szkennelés során?
31. Ismertesse a mentéssel kapcsolatos alapvető megfontolásokat!
32. Milyen file formátum támogatja a csatornánkénti 8 bit-nél magasabb tárolást?
33. Milyen nehézségekkel találkozhatunk negatívok szkennelése során?

## 16. VIZUÁLIS ADATFORMÁTUMOK

### 16.1. A formátum választása

A digitális képekkel történő munka rendkívül fontos eleme az adatok tárolási módja. A tárolás előtt több szempontot is érdemes figyelembe venni. Megfelelő adatformátum választása fontos lehet a további feldolgozás szempontjából (pl. úrfelvételek veszteséges adattárolása nem célszerű). Amennyiben nyomtatón szeretnénk megjeleníteni képünket célszerű olyan tárolási formátumot választani, amelyik támogatja mindezt (pl. olyan tárolási formátumot választok, amelyik megengedi a nyomtatásra vonatkozó alapvető jellemzők tárolását is). Ha multimédia vagy prezentációs anyagot állítunk elő, olyan tárolási formátumokat célszerű használni, amelyik a megfelelő megjelenítést szolgálja (pl. nem ajánlatos nagy minőségromlással járó, veszteséges adatformátum használata). Az alábbiakban néhány fontosnak ítélt tárolási szempontot sorolunk fel (a legtöbbször az alábbi szempontok összefüggnek):

1. további (esetleg platform független) számítógépes feldolgozás lehetősége,
2. nyomtatásban történő megjelenítés,
3. képernyőn történő megjelenítés,
4. adatkonvertálás - szabványosított tárolási mód,
5. tömörítés - kódolás lehetősége,
6. hírközlésben használt adattovábbítás lehetősége, stb.

Amennyiben nincs még konkrét elképzelésünk a felhasználást illetően, de a lehető legtöbb információt szeretnénk elmenteni a képpel és szerkesztésével, feldolgozásával kapcsolatban célszerű a feldolgozó programok saját, belső formátumát használni (Photoshop - PSD, EPS, PDD; Coreldraw - CDR, CDT; Acrobat - PDF; PowerPoint - PPT, stb.) amennyiben van ilyen.

### 16.2. Fontosabb kétdimenziós adatformátumok

Az alábbiakban igyekeztünk a legáltalánosabb és a gyakorlatban leginkább előforduló kétdimenziós megjelenítést, nyomtatást szolgáló adatformátumokat összefoglalni. Ahol külön nem jeleztük, ott az egyes formátumok fájlkiterjesztése megegyezik az elnevezésben használt három karakterrel.

## TIFF

Az Aldus és a Microsoft által kifejlesztett TIFF (Tagged Image File Format) képfomátum a raszteres képek legelterjedtebb formátuma. Az alábbiak voltak a legfontosabb jellemzők a kifejlesztéskor:

- legyen operációs rendszer független,
- legyen hardver független,
- a file struktúra vizsgálata a teljes beolvasás nélkül megvalósítható legyen,
- és kompatibilis legyen a régebbi és lehetőleg az újabb formátumokkal is.

A TIFF formátumú képek először a DTP területén terjedtek el, majd szinte minden terület átvette a formátum használatát. Így a videó feldolgozásoktól kezdve a műholdak által készített képeken keresztül az orvosi diagnosztikáig szinte minden platformon és operációs rendszeren megtalálható. A fax adatátvitel is ezt a szabványt használja. Alkalmas bináris, vonalas, szürkeségi fokozatokat tartalmazó, RGB és CMYK színtérben készített képek (8-16 bit/csatorna) tárolására. A legújabb TIFF szabvány szinte minden kompatibilitási gondot kiküszöböl. Lehetőség van az adatok tömörítésére is veszteségmentes formában (LZW - Lemple-Ziv-Welch, CCIT G3, CCIT G4, JPEG, stb.). Napjainkban alig található olyan grafikai program, amely ne tudná használni a TIFF formátumot. Amennyiben különböző platformok között akarunk képi adatokat cserélni, mindenképp a TIFF formátum a javasolt adatformátum. A TIFF formátumú fájlok kiterjesztése: **.TIF**.

## JPEG

A JPEG (Joint Photographic Experts Group) szabvány, amely elsősorban veszteséges képtömörítési szabvány, az ISO (International Standards Organization) és a CCITT (Consultation Committee on International Telephon and Telegraph) szervezetek által 1986-ban felállított munkacsoportban születtek. A diszkrét koszinusz transzformáción alapuló eljárással a látvány kismértékű romlása árán kb. 30:1-40:1 tömörítési arány is elérhető. A tömörítési arány illetve az esetleges minőségromlás mértéke a felhasználó által tetszőlegesen állítható. Az eljárás úgy lett megválasztva, hogy hardveres és szoftveres úton is hatékonyan megvalósítható legyen. Veszteségmentes tömörítésre is használható, ekkor közel a felére lehet tömöríteni az adatokat.

Az eljárás elméleti megalapozása a „Képkódolás és tömörítés” c. fejezetben részletesen is megtalálható.

Tipikus alkalmazási területe: képi adatbázisok kezelése, Internet, képek betöltés előtti megtekintése, stb.



Végezetül néhány jó tanács a JPEG formátumban történő mentéshez:

1. Nem valódi színes képek közvetlen mentése a legtöbb programban nem lehetséges JPEG formátumban (a program mentéskor nem kínálja fel a JPEG formátumot). Ezeket a képeket először át kell alakítani - pl. RGB térbe kell konvertálni -, majd utána menthetjük JPEG formátumban.
2. Érdemes a mentés előtt a minőségi faktort (Photoshop 0-10 között, Corel Photo-Paint 0-255 között) beállítani. A jó minőségű kép nagyobb méretet jelent.
3. Kisméretű, részletekben nem gazdag képek mentésére nem célszerű használni a JPEG formátumot.
4. Nagy tömörítési arány választása főleg részlet gazdag képeknél jelentős minőségromlást eredményezhet. Célszerű a képet először tömörítés nélküli, pl. TIFF formátumban menteni, majd többszöri, különböző tömörítési arány mellett JPEG formátumban lementeni az adatokat. Utólag ellenőrizve a megfelelő arány kiválasztható.
5. Veszteségmentes mentésre - ugyan a legtöbb programnál lehetséges - nem ajánlatos a JPEG formátum használata. Sokkal inkább a TIFF javasolt.
6. Nagyobb méretű színes képek Interneten történő eléréséhez célszerű a mentést progresszív módon elvégezni (a legtöbb képfeldolgozó program ezen opciót mentéskor felkínálja).

A JPEG formátumban mentett fájlok kiterjesztése: **.JPG**.

## **BMP**

A DOS és Windows operációs rendszerek alatt egy önálló bittérképes fájlformátum a BMP. Ezen formátumot kezelő (olvasás/írás és alapvető feldolgozások) programot a Microsoft cég minden operációs rendszeréhez csatolta (Windows 3 esetén ez a Paintbrush, Windows 9x és NT-nél a Paint). Mentés során azonban három lényeges jellemzőt adhatunk meg:

1. az alkalmazott operációs rendszer típusát - ez a legtöbb esetben Microsoft-kompatibilis vagy OS/2 lehet,
2. a színmélységet (1, 4, 8, 24 bit), és a
3. futamhossz szerinti veszteségmentes tömörítés lehetőségét (RLE - Run-Length-Encoding - lásd „Képkódolás és tömörítés” c. fejezet).

A fenti operációs rendszerek alatt futó programok mind használják, a többi pedig legalább olvasni tudja, így a BMP formátum nagyon elterjedt. Mivel a BMP formátum nem engedi a képi információtól eltérő, egyéb adatok (pl. nyomtatásra vonatkozó információk, megjegyzések, copyright, stb.) lementését, ezek az adatok a mentéssel elvesznek. A BMP formátumban mentett fájlok kiterjesztése: **.BMP**.

## PCX

A PCX formátumot eredetileg a Z-Soft cég hozta létre, amikor kifejlesztette a Paintbrush nevű festőprogramját. Ez a program elsősorban DOS vagy Windows környezetben fut. Kezdetben csak maximálisan 8 bites képek voltak PCX formátumban menthetők, önálló színpaletták nélkül (version 3 szabvány). A javított PCX formátumot leíró szabvány (version 5 utáni változat) már az RGB 24 bites színmélységet is tudja kezelni. Elterjedt, szinte minden program által támogatott, egyszerű felépítésű formátum. A gyakorlatban ritkán használatos. Elsősorban a DOS vagy Windows alatti gyors megjelenítéseknél, prezentációknál vagy régebbi programváltozatok használata esetén van rá szükség. A CMYK színtérbeli adatokat nem tudja kezelni. A PCX formátumban mentett fájlok kiterjesztése: **.PCX**.

## FIF

A DCT transzformáción alapuló eljárásoktól gyökeresen eltérő, rendkívül hatékony képtömörítési módszer, az ún. fraktál (önhasonló, matematikai alakzatok) transzformáción alapuló eljárás. Az eljárás veszteséges képtömörítést takar. Az elméleti megalapozása a „Képkódolás és tömörítés” c. fejezetben részletesen is megtalálható.

Az eljárással rendkívül nagy tömörítési arány érhető el (akár 100:1 is), kis minőségromlással.

Számos előnye van a széles körben elterjedt JPEG szabvánnyal szemben:

- sokkal nagyobb a tömörítés mértéke közel azonos minőség esetén,
- az emberi agy számára idegen „digitális” hatás nagyításkor kevésbé érezhető,
- jobb a kontúrok visszaadása,
- a színárnyalatok visszaadása során feltűnően jobb FIF formátum esetén.

A legtöbb program ugyan még nem tartalmazza a FIF formátumú képek olvasását jelenleg, azonban az állóképek ki- és betömörítésére szolgáló ún. TWAIN modul vagy önálló program a cég Internetes szerveréről szabadon letölthető és kereskedelmi forgalmat leszámítva szabadon használható is. Az Internetes és multimédiás alkalmazások mellett rendkívül hasznos lehet prezentációk, képi adatbázisok használatakor. A fraktál formátumban mentett fájlok kiterjesztése: **.FIF**.

## PICT

A Macintosh gépeken futó programok kedvelt formátuma a PICT formátum, rasteres képek tárolására. A képernyő megjelenítésére használt Quick Draw utasításkészleten alapul. A formátum támogatja a bittérképes, az árnyalatos és az RGB képek mentését, a CMYK színtérbeli képeket viszont nem. Nagyon hatékony veszteségmentes adattömörítési eljárást (Packbits) használ. Színes képeknél a 16 vagy 32 bites felbon-

tást, míg árnyalatos képek esetén a 2, 4, 8 bites felbontást támogatja. A legújabb változatai már a JPEG formában való tömörítést is támogatják.

A Macintosh gépen futó prezentációk és multimédia fejlesztések hasznos képformátuma, hiszen a formátum kezelése az operációs rendszer szerves része. A legtöbb nem Macintosh típusú gépen futó program által ismert. Használata Macintosh gépre történő képi információk átvitelénél is ajánlatos. A multiplatformos képfeldolgozó programok mindegyike ismeri. A fájlok kiterjesztése: **.PCT** vagy **.PIC**.

## EPS

Az EPS (Encapsulated PostScript File) az Adobe cég által kifejlesztett adatformátum, mind raszteres, mind vektoros adatok tárolására alkalmas. A formátum kidolgozásakor a különböző grafikai elemek, rendszerek egységes kezelését és alkalmazás orientált kezelését tartották fontosnak. A PostScript adatoknak közvetlen a képernyőn nem jeleníthetők meg, ezért a formátum a képet a megjelenítéshez szükséges ún. preview adatokkal együtt menti el. Ennek a felbontása jóval alacsonyabb, mint a nyomtatandó adatoké. A formátumot elsősorban raszteres és vektoros adatok cseréjekor alkalmazzák. A Windows és Macintosh rendszereken futó képfeldolgozó és rajzoló-festő programok ismert formátuma. Napjainkban, a nyomdatechnikában a PostScript nyelvet ismerő levilágítók és nyomtatók által használt szabvány formátum. Az EPS formátumban mentett fájlok kiterjesztése: **.EPS**.

## GIF

Az Interneten történő kalandozások során az olvasó legtöbbször talán a GIF (Graphics Interchange Format) adatformátummal találkozhat. A WEB és más on-line rendszerek HTML leíró nyelvét közvetlenül kihasználó raszteres adatok tárolására alkalmas formátumról van szó. A kifejlesztő a Comuserve cég volt. Kimondottan a telefonos átvitel és a HTML nyelv lehetőségeit szem előtt tartva került kialakításra a GIF formátum. Napjainkban szinte minden program ismeri (legalább olvasni tudja). A képfeldolgozó és grafikai rendszerek pedig a GIF fájlok önálló szerkesztésére is képesek. Bináris (bitmap), árnyalatos és 8 bites színelbontással használatos elsősorban, bár a legújabb szabvány támogatja a 24 bites színelbontást is. Tömörített formában végzi a tárolást. Mentéskor a GIF89a szabvány felkínálja a későbbi megjelenítés normál vagy interlaced lehetőségét. Prezentációk, multimédia anyagok és a WEB-re történő fejlesztéseknél célszerű használni. A GIF formátum egy alternatív, de szinte teljesen azonos változata a PNG formátum. A GIF formátumban mentett fájlok kiterjesztése: **.GIF**.

## PDF

Az Adobe cég által kifejlesztett PDF (Portable Document Format) fájl formátum az igényes prezentációk létrehozásának kedvelt formátuma. Kifejlesztésekor az operációs rendszertől független megjelenítés volt az elsődleges cél. Így szinte minden környezetben megtalálható (Macintosh, DOS, Windows, UNIX). PostScript Level 2 leíró nyelv képezi az alapját, ami arra utal, hogy mind a raszteres, mind a vektoros adatokat tudja kezelni. Emellett kiegészítették számos olyan hasznos funkcióval, amelyeknek köszönhetően az egyik legelterjedtebb prezentációs adatformátum. Ezek a funkciók:

- szöveg, grafika, kép, videó és animáció kezelés,
- interaktív keresés,
- navigációs lehetőségek,
- hypertext és
- interaktív tartalomjegyzék.

Mindezek alapján állíthatjuk, hogy a multimédia alapvető elemeit tartalmazza. Ezeket az elemeket tömörített formában tárolja. A PDF formátumú dokumentumok megjelenése egy korszerű könyvre emlékeztet a kivitelezés, viszont teljesen digitális. Szerkesztésére az Adobe Acrobat programrendszer Writer, Distiller vagy Exchange tagja alkalmas, míg olvasására a számos WWW helyről szabadon letölthető és használható Adobe Acrobat Reader szolgál. Az olvasó a CD - PROGRAM könyvtárban az Adobe Acrobat Reader 4.0-ás installálható verzióját találja. Mindez egyben mutatja azon hátrányát is, hogy a legtöbb program nem képes közvetlenül olvasni. A PDF formátumban mentett fájlok kiterjesztése: **.PDF**.

## PhotoCD

A PhotoCD, a Kodak cég által kifejlesztett rendszer és szabvány. Mintegy átmenet a hagyományos és digitális fényképezés között. Lehetőséget ad, hogy olcsón, jó minőségű, nagyfelbontású elektronikus képekhez jussunk anélkül, hogy ehhez drága digitális eszközöket vásárolnánk.

A szabványról részletesen is szoltunk a „**Digitális fényképezés**” c. fejezetben.

A színrendszereknél tárgyalt YCC színtérbe történő átírás után, a kapott adatokat veszteségmentes, önálló szabadalomként használt eljárással tömörítik. Az adatredukció segítségével az eredeti 18 Mbyte adatmennyiség mindegy 4-6 Mbyte-ra csökken. Így egy fotó CD lemezen minimum 100 (száz) színes kép helyezhető el.

A fotó CD-n minden kép öt (hat) különböző változatban kerül tárolásra. Mind az öt (hat) képet ugyanabból az adatállományból nyerik, ezek csupán felbontásban különböznek egymástól. Az alapfelbontás - amit a Kodak Base-nak (alpnak) nevez - 512 sornak (soronként 768 képpontnak) felel meg, ami egy normál tv-kép minőségével közel azonos. Emellett találunk még 4-szer, illetve 16-szor kisebb, valamint 4-szer, illet-

ve 16-szor (kisfilmnél nagyobb méret esetén 64-szer) nagyobb felbontású képváltozatokat is:

- Base/16            128 x 192 képpont
- Base/4            256 x 384 képpont
- Base                512 x 768 képpont
- 4Base              1024 x 1536 képpont
- 16Base            2048 x 3072 képpont
- 64Base            4096 x 6144 képpont (csak kisfilmnél nagyobb méret esetén!)

A PhotoCD-re írt képfájlok kiterjesztése: **.PCD**. Önálló készülékkel televízión vagy megfelelő (a PhotoCD szabványt ismerő) CD-ROM olvasóval megjeleníthető.

A 35 mm-es mérettől a 102\*127 mm-es síkfilm formátumig, film vagy diafelvételek tárolására alkalmas (Master és Pro Master szabvány). Emellett nyomdai előkészítésre, interaktív üzleti- és kereskedelmi bemutatók, fotóalbumok létrehozására (Portfolio szabvány), cégek, ügynökségek, archívumok képeinek tárolására, elektronikus katalogizálásra (Catalog szabvány) és orvosi diagnosztikai célokra is (Diagnostic szabvány) használható a rendszer. Napjainkban szinte minden jelentősebb raszteres adatokat kezelni tudó program használja a szabványt. A cég WWW szerveréről a jelentősebb operációs rendszerekhez illeszkedő TWAIN modul letölthető.

### **MrSID**

Jelenleg az ipari jellegű nagyméretű képeket felhasználó alkalmazások (térinformatika, kiadványszerkesztés) körében vezették be a MultiResolution Seamless Image Database (MrSID) elnevezésű formátumot 1999-ben. A formátum kifejlesztője a Los Alamos National Laboratory, forgalmazója a LizardTech Inc. cég.

A matematikai transzformáción (Discrete Wavelet Transformation - DWT) alapuló tömörítési eljárással jelenleg az egyik legjobb eredmény érhető el állóképek közel veszteségmentes tömörítése esetén a tömörítési arány és a minőség tekintetében. A tömörítés mértéke legalább 15-20:1 fekete-fehér árnyalatos kép, 20-50:1 színes kép esetén veszteségmentesen vagy kismértékű veszteséggel, ami látvány szempontjából veszteségmentesnek tekinthető.

A tömörítés algoritmusá védett, nem került nyilvánosságra. Annyit tudunk, hogy a teljes képet felhasználják tömörítéskor, nem kerül részekre bontásra, mint például JPEG esetén. A cég kifejlesztett az FBI részére is egy formátumot (WSQ) ugyanezen elvek alapján, elsősorban biztonságtechnikai alkalmazásokhoz.

Elterjedése a következő években várható. A fájlok kiterjesztése: **.SID**.

## FLASHPIX

Az 1995-ös év kezdetén a Kodak, a Microsoft, a HP és a Live Picture cégek közösen elkezdtek egy kimondottan digitális kamerához tervezett képformátum kidolgozását, amelynek a bevezetése 1996 júniusában történt. A formátum alkalmas több eltérő felbontás (Multi Resolution) hierarchikus tárolására tömörítetlen vagy JPEG tömörítést alkalmazva. Emellett a fejlesztők gondoskodtak sok olyan információ tárolásának a lehetőségéről is, amelyek nem képi adatok, hanem a képpel történő műveletekre, a felvételezés és megjelenítés körülményére utal.

Ilyen adat:

- a 64x64 pixel<sup>2</sup>-es címke-kép,
- az alapvető transzformációs műveletek (forgatás, eltolás, élesítés, stb.),
- a színtérre vonatkozó információk (RGB, YCC, RGB kalibrált, monokróm, stb.),
- a fájlstruktúrára, a felvételezés körülményeire és a felvételt készítő eszközre (digitális kamera, szkennel, analóg kamera, film) vonatkozó paraméterek és
- a tömörítésre vonatkozó jellemzők (tömörítetlen, JPEG tömörített).

Elterjedése a fenti cégek digitális kameráinak térhódításával napjainkban jelentősnek mondható. A fájlok kiterjesztése: **.FPX**.

## 16.3. Háromdimenziós adatformátumok

### VRML

A WWW (World Wide Web) HTML (HiperText Markup Language) hipertext-es programnyelvezetének kiegészítésére, továbbfejlesztésére született a VRML (Virtual Reality Modelling Language). Nem egy formátum, inkább egy leíró nyelv. A „virtuális világ” lehetőségeivel, előnyeivel és hátrányaival teszi lehetővé az interaktív kommunikációt. A VRML nyelv az OpenGL könyvtárára épülő OpenInventor modellező nyelvből alakult ki. A nyelv kifejlesztője a Silicon Graphics cég volt. A modellek mozgását JAVA nyelvű kódok adják. A VRML forráskód a HTML-hez hasonlóan ASCII formátumú. Az Internetre tervezett nyelv létrehozásánál fontos szempont volt a tömörség. A rendszer nem korlátozódik csupán a hálózatra. Jelentős szerepet kapott a multimédia alapú fejlesztésekben és alkalmazásokban is.

A VRML működése: a VRML-kód letöltése után a gépen futó VRML-browser megjeleníti a 3D objektumokból álló virtuális világot. A browser lehet önálló alkalmazás vagy egy pl. Netscape vagy Explorer ablakban futó ún. Plug-in modul. Ezután természetesen mozoghatunk a 3D virtuális világokban, amelyek lehetnek statikusak vagy

mozgók. A VRML 2.0 szabvány már valóság, interaktív animációk megjelenítésére is alkalmas. A VRML teljes lehetőségeinek kihasználásához azonban nagyteljesítményű, 3D grafikus gyorsítóval és digitális médiával felszerelt munkaállomásra van szükség. Egy MMX alapú személyi számítógép is csak szegényes megjelenítésre alkalmas.

A 3D térben való mozgás az egér segítségével történik, ami a felhasználótól kis gyakorlatot igényel.

A VRML alkalmazásokat ma két csoportba sorolhatók:

1. Az eddigi számítógépes gyakorlatban használt 3D alkalmazások (CAD/CAM, molekulamodellzés, 3D adatábrázolás, térinformatika, orvosi alkalmazások, oktatás, stb.)
2. Bár az első felhasználók a játékipar területén jelentek meg, az interaktív virtuális valóság rövidesen komoly vetélytársa lehet a hagyományos filmnek az ismeretterjesztés, a művészet és a szórakoztatás egyes területein.

A VRML formátumú fájlok kiterjesztése: **.WRL**.

Az olvasó a CD lemezen az DEMO/VRML könyvtárban találja a Windows 95/98 alatti, VRML 2.0 szabványnak megfelelő programcsomagot és számos érdekes példát.

## **DXF**

Az Autodesk Inc. által kifejlesztett formátum elsősorban két- és háromdimenziós modellek tárolására alkalmas. A CAD (Computer Added Design) programok szabványának tekinthető adattárolási formája. A DXF (Drawing Interchange File) kifejlesztésekor a környezetfüggetlen, vektoros környezetben előforduló adatok tárolása volt a cél. Általában a személyi számítógép alapú képfeldolgozó programok nem képesek önállóan kezelni a formátumot vagy gyakran jelentős információk elhagyásával, képesek csak olvasni. Az AutoCAD, ArchiCAD és a legtöbb tervező programrendszer által használt tárolási formátum eredetileg az ASCII alapú adatcserét hivatott elősegíteni. Jelenleg háromdimenziós animációs lehetőségeket is tartalmaz. Szintén az Autodesk cég által kifejlesztett, szabvány, tárolási mód, az FLI és FLC formátumok, amelyek kimondottan animációk tárolását célozzák. A DXF formátumú fájlok kiterjesztése: **.DXF**.

## **16.4. Digitális mozgókép formátumok**

### **MPEG**

Az MPEG szabvány ismertetésekor részletesen szoltunk az eljárás elvéről (8. fejezet). Most a gyakorlatban előforduló MPG kiterjesztésű adatfájlokat jellemezzük. Az MPEG formátumú fájlok leglényegesebb vonása, hogy az időben szomszédos képek

közi változást kódolja. Biztonsági okokból azonban időről-időre „újrakezdi” az egymásra épülő kódolást, s ezzel meggátolja az esetleges adatátviteli hibák továbbterjedését. Leegyszerűsítve, az adatfolyam három, különböző szerepet betöltő képtípus egymásutánjából áll:

1. **Különálló kép.** Átlagosan fél másodpercenként egy teljes kép tömörítését jelenti, JPEG formában. Miután ez a kép a következő adatok számára mintegy referenciaként szolgál, viszonylag alacsony tömörítési arányú (12:1), viszont jó minőségű.
2. **Előrebecsült kép.** Az előző képhez viszonyított változást kódolja, s ezzel magas tömörítési arányt biztosít. Minél több előrebecsült kép következik egymás után, annál inkább eltérhet a helyreállított képsorozat az eredetitől.
3. **Kétirányú kép.** Itt a megelőző és a következő kép átlagát használják fel. Ezáltal nagyon magas tömörítési arány érhető el.

Annak meghatározása, hogy mikor melyik képtípust kell kódolni, a képsorozat tartalma alapján az MPEG kódoló egység (általában speciális hardver) feladata. Ez alapvetően befolyásolja a kialakuló videó minőségét.

Az MMX utasításkészletet hatékonyan megvalósító processzorok az MPEG-1 fájlok kitömörítésével, megfelelő szoftver használata esetén megbirkóznak. A CD lemezen az olvasó talál ilyen driver-t a PROGRAM könyvtárban. Az MPEG-1 formátumú fájlok kiterjesztése: **.MPG**.

## AVI

A Microsoft, amely az AVI (Audio Video Interleaved) formátum kifejlesztője, videó és hanginformáció tárolására dolgozta ki a formátumot. A videó forrás jelét digitalizálva a formátum külön kezeli a mozgóképet és a hanginformációt. A hanganyagot Wave alakban, míg a képet DIB formátumban (Device Independent Bitmap) kezeli. A képinformáció tárolásakor egy képből kiindulva, a következő kép csak azon részleteit tárolják, amelyek megváltoztak. Ezeket a részképeket nevezik deltakeretnek (delta-frame). Tároláskor csak a jelentős változást elszenvedett deltakereteket veszik figyelembe és mentik el. A felvétel minőségének javítása érdekében a tömörítést végző szoftver a teljes képet is rögzíti. A deltakeretek nem alkalmasak a felvételen történő gyors változások megfelelő minőségű rögzítésére, hiszen így rendkívül nagy adatállományt kapnánk. Ennek következtében a lassan változó, nyugodt felvételek egyenletesen, jó minőségben lejátszhatók, míg a gyors változásoknál a felvétel „ugrásokat” szenved. Az ugrások elkerülését a teljes megjelenítő keret kicsinyítésével csökkenthetjük. Lejátszáskor a hangadatok közvetlenül a hangkártyára, míg a képinformáció a megjelenítő eszközre kerül.

A személyi számítógépes környezetben futó kép- és videó feldolgozó szoftverek könnyedén kezelik a formátumot. Még a legegyszerűbb programok is felkínálnak



alapvető szerkesztési műveleteket. A Windows alapú rendszerek kedvelt mozgókép adatformátuma, mivel nem csupán a winchesterről, hanem egy CD-ROM olvasóról is lejátszhatók, minimális hardver igény mellett. Az AVI formátumú fájlok kiterjesztése: **.AVI**.

## QUICKTIME

A felhasználók számára talán a leginkább szerencsés, mozgókép tárolására alkalmas formátum a QuickTime. A formátum kidolgozója az Apple cég. Hang, animáció és videó anyag tárolására alkalmas. A teljes rendszer négy alapvető elemből épül fel:

1. szoftver rendszer,
2. tömörítő algoritmus,
3. szabvány mozgókép formátum és a
4. szabvány felhasználói felület.

A szoftver menedzseli a teljes digitalizáláshoz, szerkesztéshez, tömörítéshez, tároláshoz és lejátszáshoz szükséges elemeket. A tömörítő és kitömörítő algoritmusok együttesét - amelyet codec-nek neveznek - alapvetően ötféle algoritmus végzi:

1. Apple Photo-JPEG- fotók JPEG alapú tömörítése,
2. Apple Video Codec - videó tömörítés,
3. Apple Animation Codec - futamhossz alapú animáció tömörítés,
4. Apple Graphics - 8 bites álló és mozgóképek tömörítése,
5. Apple None - egyéb adattömörítés.

Rendszer rugalmasságát mutatja, hogy szinte minden jelentősebb codec eljárás (PhotoCD, Video Disc, SuperMac, Intel Indeo, MPEG, stb.) beépíthető a QuickTime formátumba.

A QuickTime formátumot tartalmazó fájlok kiterjesztése: **.QTM** vagy **.MOV**. A Windows vagy UNIX alapú környezet önálló lejátszó szoftvert igényel, ami szabadon használható (a legtöbb szoftverhez vagy QuickTime formátumú adathoz mellékel). Multimédia anyagok és bemutatók alkalmával nagyon jól használható, mivel a sokféle algoritmus alapján megtalálható a megjelenítéshez szükséges optimális minőség/méret arány.

Végezetül a következő oldalon található táblázatban összefoglaltuk az alapvető vizuális adatformátumokat az alkalmazási területtel, a kidolgozó cégeket és a formátumhoz kapcsolódó kiterjesztéseket. A tárgyalt formátumokat (normál módon szedett) kiegészítettük olyan adatformátumokkal is, amelyekkel az olvasó igényesebb szoftverek, munkaállomások esetén gyakran találkozhat (dőlt elnevezés).

Formátum megnevezése	Adatfájl kiterjesztése	Kidolgozó cég vagy társaság	Alkalmazási terület
Tagged Image File Format	TIF	Aldus, Microsoft	Raszteres képek
Joint Photographic Experts Group	JPG	ISO, CCITT	Raszteres képek veszteséges tárolása
Windows Bitmap	BMP	Microsoft	Raszteres képek
<i>Windows Bitmap</i>	<i>RLE</i>	<i>Microsoft</i>	<i>Raszteres képek</i>
<i>Amiga Interchange File Format</i>	<i>IFF</i>	<i>Commodore</i>	Vizuális adatesere Amiga és egyéb gépek között
PC Paintbrush PCX	PCX	Z-soft	Raszteres képek
Fractal Image Format	FIF	Iterated	Raszteres képek veszteséges tárolása
Multiresolution Seamless Image Database	SID	LizardTech Inc.	Raszteres nagymértékű képek közel veszteségmentes tömörítése
Macintosh PICT	PCT, PIC	Apple	Raszteres képek
Encapsulated PostScript	EPS	Adobe	Raszteres és vektoros képek, nyomdai munkákhoz
Graphics Interchange Format	GIF	Compuserve	Raszteres képek
<i>Portable Network Graphics</i>	<i>PNG</i>	<i>Compuserve</i>	<i>Raszteres képek hálózatos felhasználása</i>
Portable Document Format	PDF	Adobe	Digitális könyvek készítése, prezentáció
Kodak PhotoCD	PCD	Kodak	Nagyfelbontású raszteres képek
Virtual Reality Modelling Language	VRL	Silicon Graphics	Animáció, virtuális valóság
Drawing Interchange File	DXF	Autodesk	Számítógépes tervezés
<i>Autodesk Animation</i>	<i>FLI, FLC</i>	<i>Autodesk</i>	<i>Animáció</i>
MPEG-1	MPG	Motion Picture Experts Group	VHS minőségű mozgóképek
Audio Video Interleaved	AVI	Microsoft	Videó
QuickTime VR	QTM, MOV	Apple	Hang, kép animáció és videó
<i>Device Independent Bitmap</i>	<i>DIB</i>	<i>Microsoft</i>	<i>Raszteres képek</i>
<i>Raw</i>	<i>RAW</i>		<i>Vizuális adatok cseréje</i>
<i>Targa Sequence</i>	<i>TGA, VDA, ICB, VST</i>	<i>Targa</i>	<i>Vizuális adatok, TGA - nagyfelbontású képek</i>
<i>Pixar PXF</i>	<i>PXF</i>	<i>Pixar</i>	<i>3d képek, animáció</i>
<i>Scitex Continuous Tone</i>	<i>SCT</i>	<i>Scitex</i>	<i>Raszteres képek</i>
<i>Windows Metafile</i>	<i>WMF</i>	<i>Microsoft</i>	<i>Raszteres és vektoros adatok</i>
<i>Computer Graphics Metafile</i>	<i>CGM</i>		<i>Raszteres és vektoros adatok</i>
<i>Micrografx Designer</i>	<i>DRW</i>	<i>Micrografx</i>	<i>Vektoros és raszteres adatok</i>
<i>CorelDraw</i>	<i>CDR</i>	<i>Corel</i>	<i>Vektoros, raszteres adatok</i>
<i>Enhanced Metafile</i>	<i>EMF</i>		<i>Vektoros, raszteres adatok</i>
<i>WordPerfect Graphics</i>	<i>WPG</i>		<i>Vektoros adatok</i>
<i>Filmstrip</i>	<i>FLM</i>		<i>Animáció, videó</i>
Windows Waveform	WAV	Microsoft	Hang
<i>Initial Graphics Exchange Standard</i>	<i>IGES</i>	<i>U.S. Department of Commerce</i>	<i>Vektor adatok</i>
<i>Digital Line Graphics</i>	<i>DLG</i>	<i>U.S. Geological Survey</i>	<i>Topológiai jellegű vektoros adatok</i>
<i>Erdas Imagine Format</i>	<i>IMG</i>	<i>Erdas</i>	<i>Raszteres, rétegelt adatok</i>
<i>Digital Elevation Model</i>	<i>DEM</i>	<i>U.S. Geological Survey</i>	<i>Topográfiai 3D adatok</i>
<i>Topologically Integrated Geographic Encoding and Referencing System - TIGER</i>	<i>GBF/DIME</i>	<i>U.S. Census Bureau</i>	<i>GIS adatok</i>

28. ábra

## Ellenőrző kérdések

1. Mit takar a TIFF formátum?
2. Mikor nem használhatunk JPEG tömörítést?
3. Mi a kvantáló?
4. Mi a progresszív képtömörítés lényege?
5. Milyen jellemzőket adhatunk meg BMP formátum mentésekor?
6. Mit tud a PCX formátumról?
7. Mi a fraktál alapú tömörítés lényege?
8. Mit tud a MrSID tömörítésről?
9. Milyen előnyei mutatkoznak a fractál formátumnak a JPEG szabvánnyal szemben?
10. Milyen színtérben dolgozik a Macintosh PICT formátum?
11. Hol és milyen célból használnak EPS formátumot?
12. Mi volt a GIF formátum kidolgozásának a célja?
13. Ismertesse a PDF funkcióit!
14. Mit takar a Kodak PhotoCD rendszere?
15. Mire használhatók az egyes PhotoCD szabványok?
16. Mit jelent a 16Base elnevezés?
17. Milyen adatok tárolhatók a képi adatok mellett egy FlashPix formátumú kép esetén?
18. Hogyan működik a VRML?
19. Melyek napjaink VRML alkalmazásai?
20. Mit takar a DXF?
21. Melyek az MPEG-1 szabvány alapvető képtípusai?
22. Alkalmas-e a hang tárolására az AVI formátum?
23. Mi a deltakeret?
24. Mi a QuickTime codec?
25. Melyek azok a vizuális adatformátumok, amelyek képi adatok mellett, más információk tárolását is támogatják?

# AJÁNLOTT IRODALOM

## Szakkönyvek

- ÁLLÓ, G - HEGEDŰS, GY. CS. - KELEMEN, D. - SZABÓ, J. (1989): A digitális képfeldolgozás alapproblémái. Műszaki Tudományok, Az elektronika újabb eredményei. Akadémiai kiadó, Budapest.
- ÁLLÓ, G. - FŐGLEIN, J. - HEGEDŰS, GY. CS. - SZABÓ, J. (1993): Bevezetés a számítógépes képfeldolgozásba. BME Mérnöktovábbképző Intézet, Budapest.
- BARNESLEY, M. F. (1988): Fractals everywhere, Academic Press, Inc., London.
- BARNESLEY, M. F. - HURD, L.P. (1993): Fractal image compression, AK Peters Ltd., Wellesley.
- BATES, R.H.T. - McDONNELL, M.J. (1986): Image Restoration and Reconstruction, Oxford University Press, New York.
- BATTY, M. – LONGLEY, P. (1994): Fractal cities, Academic Press, Inc., London.
- BÄSSMAN, H. - BESSLICH, P. W. (1995): Ad Oculos Digital Image Processing, International Thomson Publishing, London.
- BERKE, J. - HEGEDŰS, GY. CS. - KELEMEN, D. - SZABÓ, J. (1996): Digitális képfeldolgozás és alkalmazásai, Keszthelyi Akadémia Alapítvány, ISBN 963 04 7466 2.
- BERKE, J. - HEGEDŰS, GY. CS. - KELEMEN, D. - SZABÓ, J. (1998): Digitális képfeldolgozás és alkalmazásai, Keszthelyi Akadémia Alapítvány, ISBN 963 03 5116 1.
- BERKE, J. - VIRÁG, M. (1998): Számítógépes grafika és prezentáció, Keszthelyi Akadémia Alapítvány - Talentum Kft., ISBN 963 03 4599 4.
- BJALIK, G.I. (1963): Színes televízió. Műszaki Könyvkiadó.
- BUZÁS, F. (1995): Nyomdaipari elektronikus képfeldolgozás. Nyomdász Kiadó, Budapest.
- CORNSWEET, T.N. (1970): Visual Perception. Academic Press, New York.
- DAY, J.B. (1997): Color Scanning Handbook, Prentice Hall International, Inc. USA.
- DAVIES, A. - FENNESSY, P. (1994): Introduction to Electronic Imaging for Photographers, Focal Press, Oxford.
- DUDA, R.O. - HART, P.E. (1973): Pattern classification and scene analysis. Wiley-Interscience Publication, New York.
- DUFF, M.J.B. - LEVIALDI, S. (1981): Languages and Architectures for Image Processing, Academic Press, New York.
- EARNSHAW, R.A. - GIGANTE, M.A. - JONES, H. (1993): Virtual Reality Systems, Academic Press, London.
- ENCARNACAO, J.L. – PEITGEN, H.-O. – SAKAS, G. – ENGLERT, G. editors (1992): Fractal Geometry and Computer Graphics, Springer-Verlag, Berlin Heidelberg.

- ERDAS Field Guide. (1994), Erdas Inc., Atlanta, USA, Third Edition.
- ERDAS Field Guide. (1997), Erdas Inc., Atlanta, USA, Fourth Edition.
- FARACE, J. (1996): The Photographer's Digital Studio, Peachpit Press, USA.
- FU, K.S. (1982): Syntactic pattern recognition and applications. Prentice-Hall, New York.
- FREEMAN, M. (1993): Collins Photographer's Handbook. HarperCollins Publishers.
- GONZALEZ, R.C. - THOMASON, M.G. (1978): Syntactic pattern recognition. Addison-Wesley, New York.
- GONZALEZ, R.C., and WOODS, R.E. (1992): Digital Image Processing, ADDISON-WESLEY Publishing Company, Inc..
- HORD, R. M. (1986): Remote Sensing Methods and Applications. A Wiley-Interscience Publication John Wiley & Sons, New York.
- JAIN, A. K. (1989): Fundamentals of Digital Image Processing, Prentice Hall Information and System Sciences Series, London.
- JAKAB, ZS. (1998): Adobe Photoshop 4.0 magyar változat, ComputerBooks, Budapest.
- KASAI, A. – SPARKMAN, R. (1997): Essentials of Digital Photography, New Riders Publishing, USA.
- LARISH, J. (1992): Digital Photography: Pictures of Tomorrow. Micro Publishing Press.
- LEEUW, B. (1997): Digital Cinematography, Academic Press Professional, London.
- LEVINE, M.D. (1985): Vision in Man and Machine, McGraw-Hill, New York.
- MANDELBROT, B.B. (1982): The Fractal Geometry of Nature, W. H. Freeman, New York.
- MATHER, P. M. (1991): Computer Applications in Geography. John Wiley & Sons.
- MCCLELLAND, D. (1999): PhotoShop 5 Biblia I.-II., IDG Books Worldwide Inc., Kiskapu Kft. Budapest.
- MULLER, J. P. (1988): Digital Image Processing in Remote Sensing. Taylor & Francis, London and Philadelphia.
- NAGY, S. - PERJÉS L. (1996): A számítógépes grafika, LSI Oktatóközpont, Budapest.
- PEITGEN, H.-O. – RICHTER, P.H. editors (1986): The Beauty of Fractals, Springer-Verlag, Berlin Heidelberg.
- PEITGEN, H.-O. – SAUPE, D. editors (1988): The Science of Fractal Images, Springer-Verlag, Berlin Heidelberg.
- PRATT, W.K. (1978): Digital image processing. Wiley-Interscience, New-York.
- ROSENFELD, A. (1979): Picture languages, formal models for picture recognition. Academic Press, New York.
- SABINS, F. F. (1987): Remote Sensing Principles and Interpretation. W. H. Freeman and Company, New York.
- SERRA, J. (1982): Image Analysis and Mathematical Morphology. Academic Press.

- SHANNON, C.E. - WEAVER, W. (1986): A kommunikáció matematikai elmélete, OMIKK, Budapest.
- SZÉKELY, V. (1994): Képkorrektció, hanganalízis, térszámítás PC-n, ComputerBooks, Budapest.
- SZÉKELY, V. - POPPE, A. (1992): A számítógépes grafika alapjai IBM PC-n. ComputerBooks, Budapest.
- SZIRÁNYI, T. - BERKE, J. (1997): Magyar képfeldolgozók és alakfelismerők országos konferenciája kiadvány, PATE Georgikon, Keszthely.
- THEODORIDIS, S. – KOUTROUMBAS, K. (1999): Pattern Recognition, Academic Press, Inc., London.
- TOU, J.T. - GONZALEZ, R.C. (1974): Pattern recognition principles. Addison-Wesley, New York.
- TURNER, M.J. – BLACKLEDGE, J.M. – ANDREWS, P.R. (1998): Fractal Geometry in Digital Imaging, Academic Press, Inc., London.
- UHR, L. (1973): Pattern Recognition, Learning, and Thought. Prentice-Hall, USA.
- UMBAUGH, S.E. (1998): Computer Vision and Image Processing, Prentice Hall International, Inc. London.
- VAUGHAN, T. (1993): Multimedia: Making It Work. Osborne McGraw-Hill.
- YOUNG, T.Y. - FU, K.-S. (1986): Handbook of Pattern Recognition and Image Processing. Academic Press, Inc., London.

## Cikkek

- BERKE, J. – CSETVERIKOV, D. – FAZEKAS, A. – GÁCSI, Z. – SZABÓ, J. – SZIRÁNYI, T. (2000): Képfeldolgozás alkalmazásának eredményei Magyarországon, Magyar Informatikusok II. Világtalálkozója, Budapest, 2000. június 5-8.
- BERKE, J. - KOCSIS, P. - KOVÁCS, J. (1997): Psychovisual Comparison of DCT and Fractal Based Image Compressing Methods, KEPAF Conference on Image Analysis and Pattern Recognition, Keszthely, 131-136.
- BERKE, J. – PAIS, CS. – HOFFMAN, P. (2000): Strukturális és morfológiai jellemzők vizsgálata faipari képfeldolgozó rendszerekben, (Structural and Morphological Features in Wood Industrial Image Processing Systems), Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II:125-130.
- BERKE, J. – PALLÉR, N. (2000): ERDAS IMAGINE gyakorlati alkalmazása a Georgikon Karon, Második ERDAS Felhasználói Konferencia, Budapest, 2000. 10. 12.
- BERKE, J. – SIMON, SZ. (2000): VRML modellezés valós vizuális adatok alapján, (VRML Modelling based on Visual Information), Magyar Képfeldolgozók és

- Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II:131-134.
- BERKE, J. - VARGÁNÉ DUGONICS, R. - TÖRÖCSIK, C. (1998): Légi- és űrfelvételeken történő osztályozási eljárások alkalmazása a mezőgazdaságban. XL. Georgikon Tudományos Napok, 1998., szeptember 24-25., Keszthely.
  - BERKE, J. (1999): Comparison and Application Possibilities of JPEG and Fractal-based Image Compressing Methods in the Development of Multimedia Based Material, IEEE Data Compression Conference - DCC '99, USA.
  - BERKE, J. (1999): Vizuális informatikai tantárgyak oktatási tapasztalatai. GEORGIKON-MÉDIA '99 konferencia, Keszthely, 1999. Május 27-28., Computer Panoráma 10:99/7.
  - BERKE, J. (2000): Környezetfigyelő rendszerek, Új Alaplap XVIII/11:18-19.
  - CHEN, W.-H. - PRATT, W.K. (1984): Scene Adaptive Coder, IEEE Transactions on Communications. COM-32:225-232.
  - CZÚNI, L. (2000): A fraktáltömörítés hatékonyságának javítása textúrák kódolásában, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája (KEPAF Conference on Image Analysis and Pattern Recognition), Noszvaj, 2000. Január 20-22, II: 3-8.
  - CSETVERIKOV, D. – NAGY, M. – VERESTÓY, J. (2000): Comparison of tracking techniques applied to digital PIV, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 57-62.
  - CSETVERIKOV, D. (2000): Registration of Shapes for Defect detection, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 17-22.
  - CSOMAI, I. - SZABÓ, J. (2000): Képfeldolgozás alkalmazása in vivo kísérletek kiértékelésében, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II:75-76.
  - CSOMAI, I. - SZABÓ, J. (2000): Színes kép osztályozó program és alkalmazásai, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II:121-124.
  - FAZEKAS, A. – HAJDÚ, A. – HAJDÚ, L. (2000): Véges dimenziós szomszédsági szekvenciák strukturális vizsgálata, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 29-32.
  - FAZEKAS, A. – HAJDÚ, A. – HAJDÚ, L. (2000): Végtelen dimenziós szomszédsági szekvenciák strukturális vizsgálata, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 33-36.
  - FEJES, S. (2000): Automatic 3D Computer Model Generation at Your Fingertips, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 43.
  - GÁCSI, Z. – SÁRKÖZI, G. (2000): Anyaginformatikai ágazat az Anyag- és Kohómérnöki

- Karon. Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22. II: 109-114.
- GYŐRI, E. – FIRTHA, F. (1997): Látórendszer alkalmazása gyümölcs belső hibáinak detektálására. Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Pannon Agrártudományi Egyetem, Georgikon Mezőgazdaságtudományi Kar, Keszthely 1997. Október 9-11, I. 151-155.
  - ISTENES, P.(1997): License plate identification by real-time image processing, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, PATE Georgikon Mezőgazdaságtudományi Kar, Keszthely 1997. Október 9-11, I. 249-252.
  - KELEMEN, D. - SZABÓ, J. (2000): Automatikus mikrobiológiai-válasz kiértékelő rendszer. Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II:155-156.
  - KÉPFELDOLGOZÓK ÉS ALAKFELISMERŐK SZAKOSZTÁLYA,  
<http://silicon.terra.vein.hu/~kepaf>.
  - LADÁNYI, Z. - LADÁNYI, P.- MÁTHÉ, J. – NAGY, F. – NAGY, T. – PATKÓ, T. (2000): A HV-PC16 Videodigitalizáló kártya és alkalmazásai, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II:143-148.
  - SZABÓ, J. (1999): Elektrodaellenőrzés osztott rendszerrel, Műszaki Magazin, 1999. Május. 36-38
  - SZABÓ, J. (1999): Minőség-ellenőrzés képfeldolgozással, Műszaki Magazin, 1999. Március. 44-45
  - SZABÓ, J. (2000): Mérés számítógépes képfeldolgozási módszerekkel, Műszerügyi és Méréstechnikai Közlemények. MTA MMSZ Kft, Budapest, 36 évf. 65 szám, 55-60.
  - SZIRANYI, T. – BERKE, J. ed. (1997): Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája. Pannon Agrártudományi Egyetem, Georgikon Mezőgazdaságtudományi Kar, Keszthely.
  - SZIRÁNYI, T. – TÓTH, Z. (2000): Sztochasztikus festmény generálás, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 45-50.
  - SZIRANYI, T. ed. (2000): Képfeldolgozók és Alakfelismerők (NJTSZ-KEPAF) 2. Konferenciája. Noszvaj, 2000 január 20-22., MTA SZTAKI, Budapest.
  - TANÁCS, A. – PALÁGYI, K. - KUBA, A. (2000): Képek kölcsönös információtartalmán alapuló automatikus módszer orvosi képek regisztrációjára, Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Noszvaj, 2000. Január 20-22, II: 101.
  - ZALÁNYI, S. - BERKE, J. (1997): Digitális morfometriai különbségek az endometrium proliferatív és szekréciós fázisában (Digital Morphometric Differences between the Proliferative and Secretory Phase of the Endometrium), Magyar Képfeldolgozók és Alakfelismerők Országos konferenciája, Keszthely, 1997. Október 9-11, I. 31-35.



## Tárgymutató

## A

A/D átalakító. 28, 46  
 adaptív. 112, 113, 118  
 additív színkeverés. 19  
 AGP. 30  
 alakfelismerés. 10, 95, 96, 101, 105, 143, 161  
 aluláteresztő szűrő. 59, 66, 67  
 aritmetikai kódolás. 112, 127  
 AVI. 199, 200, 201, 202

## B

Bayes-féle döntési módszer. 96  
 bilineáris interpoláció. 48, 82  
 bit-sík kódolás. 113  
 BMP. 180, 192  
 box-módszer. 99  
 Butterworth szűrő. 60

## C

CCD. 25, 26, 27, 152, 156, 157, 158, 162, 163, 171, 172, 178, 179, 180, 184  
 CCITT. 37, 39, 108, 114, 118, 191, 201  
 CD-ROM. 31, 32, 39, 107, 196, 200  
 CMOS. 26, 179, 180, 184  
 CMYK. 19, 184, 186, 188, 191, 193  
 csapocskák. 15, 16, 17, 18, 19

## D

DCT. 117, 118, 119, 120, 122, 125, 193  
 denzitás. 185, 186  
 differencia operátor. 69, 70  
 digitalizálás. 24, 27, 42, 45, 47, 77, 123, 144, 172, 177, 179, 181, 200  
 digitalizáló kártya. 49, 152, 153  
 Diszkrimináns függvények. 98  
 dithering. 37, 38, 126  
 döntésfüggvény. 85, 86, 98, 100  
 DXF. 198, 202

## E

elektromágneses sugárzás. 14, 42  
 EPS. 190, 194, 202  
 exponenciális szűrő. 60, 61, 67

## É

él. 16, 20, 52, 59, 60, 62, 65, 86, 90, 91, 92  
 élkiemelés. 52, 65, 66, 68, 70

## F

fantomkép. 173  
 felbontás. 7, 17, 24, 27, 40, 47, 113, 122, 136, 144, 152, 155, 158, 178, 193, 194, 195, 196  
 fény. 14, 16, 23, 27, 38, 42, 74, 110, 129, 172, 178  
 fényforrás. 14, 19, 20, 23, 24, 42, 172  
 FIF. 123, 193, 201  
 filmíró. 37  
 filmtípus. 186  
 filmtípus regisztrálás. 186  
 fizikai felbontóképesség. 184  
 FlashPix. 179, 197  
 foltkeresés. 87  
 foltok. 15, 20, 60, 62, 84, 86, 90, 101, 113, 146, 157, 164  
 Fourier transzformáció. 51, 58, 59  
 fraktál. 122, 123, 193  
 frekvenciatartomány. 50, 58, 59, 61, 66, 73  
 futamhossz. 103, 108, 114, 192, 200

## G

gamma-érték. 186, 187  
 geometriai korrekció. 10, 76, 77, 78, 80, 81  
 GIF. 194, 202  
 gradiens. 69, 70, 71, 85, 91  
 gyártásvezérlő. 155

## H

HDTV. 107, 108, 124, 183  
 HSI. 37  
 Huffman. 112, 113, 125

## I

inverz. 47, 52, 54, 59, 60, 61, 80, 120  
 iránykód. 85, 109, 114, 115  
 ISODATA. 100

**J**

JPEG. 37, 39, 118, 119, 120, 121, 122, 123, 124, 125, 179, 180, 191, 192, 193, 194, 196, 197, 199, 200

**K**

kalibráció. 147, 153  
 kamera. 25, 43, 49, 130, 138, 143, 148, 152, 156, 158, 162, 172, 197  
 Karhunen-Loeve. 98, 116, 117, 119  
 képfeldolgozás. 1, 3, 5, 9, 14, 19, 23, 27, 29, 40, 42, 49, 50, 58, 70, 76, 78, 95, 107, 134, 140, 14, 151, 160, 166, 170, 176  
 képfelismerés. 10  
 képfokozás. 50, 53  
 képhelyreállítás. 50, 52, 53, 76, 119  
 képkorrekció. 10, 76  
 képosztályozás. 10  
 képpont. 9, 20, 23, 26, 33, 37, 40, 44, 51, 53, 62, 68, 74, 78, 80, 84, 86, 88, 95, 100, 111, 144, 147, 162, 178, 195, 196  
 képtömörítés. 37, 107, 11, 115, 118, 122, 144, 166, 176, 191, 193  
 kimenő felbontóképesség. 184  
 klaszterezés. 96, 99, 100, 101  
 kódolás. 29, 39, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 121, 125, 171, 187, 190, 199  
 kontraszt. 52, 143, 174  
 kontrasztfokozó. 52, 143  
 konvolúció. 37, 51, 61, 68, 70  
 koordináta transzformáció. 79  
 küszöb. 37, 40, 55, 60, 63, 70, 91, 100, 116, 148, 153, 162  
 kvantálás. 42, 43, 45, 46, 47, 48, 49, 50, 54, 77, 119, 120

**L**

Laplace operátor. 70  
 látás. 6, 9, 11, 14, 20, 109, 115, 118, 124, 182  
 lateral inhibition. 16  
 látógödör. 17  
 legközelebbi szomszéd módszer. 82, 98  
 leképezés. 9, 15, 25, 42, 43, 50, 54, 56, 81, 122, 135, 179

lényegkiemelés. 74, 96, 98

LUT. 55

**M**

medián, 63, 103  
 minőség ellenőrzése. 157  
 mintaillesztés. 21, 90, 154, 174, 175  
 mintavételezés. 10, 25, 42, 43, 44, 46, 48, 49, 50, 77, 78  
 MPEG. 37, 39, 124, 125, 179, 198, 199, 200  
 multimédia. 12, 28, 35, 108, 124, 162, 187, 190, 194, 195, 197, 200  
 multispektrális. 44, 47, 74, 95, 131, 135, 138, 139, 140  
 multitemporális. 74, 95, 139, 140  
 moduláció átviteli függvény. 184  
 MrSID. 39, 190  
 MTF. 184

**O**

OCR. 90, 167, 168, 169, 187  
 okmányhamisítás. 171  
 optikai felbontóképesség. 184  
 optikai karakterfelismerés. 12, 90, 167  
 optikai lemez. 32, 39, 107  
 osztályozás. 74, 84, 85, 86, 90, 95, 96, 98, 100, 101, 102, 103, 104, 140, 141, 143, 147, 158, 168

**Ö**

összefűzés. 86

**P**

pálcikák. 15, 17  
 PCI. 29, 30  
 PCX. 7, 36, 193  
 PDF. 190, 195  
 PhotoCD. 195, 196, 200  
 PICT. 193  
 plotter. 37  
 pupilla. 15, 16, 17

**Q**

QUICKTIME. 200

## R

real-time. 27, 39, 151  
 receptor. 15, 16, 17  
 redundancia. 28, 39, 109, 111, 114, 115, 124  
 RGB. 18, 19, 28, 37, 118, 147, 162, 184, 185, 186, 187, 191, 192, 193, 197  
 robot. 9, 154, 164

## S

sajátság. 9, 39, 84, 90, 95, 98  
 sajátságvektor. 85, 86, 87, 88, 90, 96, 97, 98, 99, 100, 101, 103  
 screening. 186  
 simaság. 15, 87  
 Sobel operátor. 70  
 spektrális reflektancia. 20, 140  
 szegmentálás. 10, 54, 84, 143, 145, 154, 163, 174  
 szemlencse. 14, 15, 16  
 szín. 14, 18, 28, 38, 109, 118, 122, 132, 143, 147, 162, 178, 182  
 színérzet. 16, 19, 20  
 színezet. 19  
 színelbontás. 17, 27, 178, 185, 186, 187, 194  
 színlátás. 17, 18  
 színmélység. 104, 181, 185, 192, 193  
 szintaktikus. 10, 95, 96, 101, 102  
 szintre vágás. 55  
 szkennel. 23, 29, 37, 111, 131, 135, 166, 178, 181  
 szoftveres felbontóképesség. 184  
 szubtraktív színkeverés. 19, 37  
 szűrés. 16, 33, 37, 51, 58, 61, 62

## T

tanító. 86, 96, 99  
 távolságfüggvény. 59, 85, 104  
 telítettség. 19  
 textúra. 16, 21, 95, 103  
 TIFF. 7, 118, 179, 180, 185, 191, 192  
 tónusgörbe. 185, 186, 187  
 többsávos kép. 73, 74, 88, 95, 110  
 töltéskép. 25, 38, 43

tömörítés. 25, 39, 40, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 118, 119, 120, 122, 123, 124, 125, 143, 144, 179, 187, 190, 191, 192, 193, 194, 196, 197, 199, 200  
 trapéz szűrő. 61, 68

## U

ujjlenyomat. 164, 170, 171, 175  
 UMA. 31, 35

## Ú

újrámintavételezés. 77, 81, 82

## V

vakfolt. 16  
 veszteséges. 39, 108, 115, 118, 122, 124, 166, 179, 187, 190, 191, 193, 201  
 veszteségmentes. 111, 113, 118, 122, 179, 181, 187, 191, 192, 193, 195, 196  
 világosság. 9, 14, 19, 20, 21  
 világosságkód. 42, 46, 50, 62, 65, 68, 77, 88, 91, 95, 103  
 VLC. 111, 121  
 vonalvékonyítás. 92  
 VRML. 197, 198

## X

Xerox eljárás. 37  
 Xerox. 177

## Y

YUV. 28, 45, 118, 162

## Z

zaj. 52, 55, 57, 61, 70, 74, 87, 107, 110, 134, 143, 167, 174  
 zajlenyomás. 52, 57, 58, 59, 61, 66, 68  
 zajszűrés. 26, 55, 57, 58, 61, 63, 66, 110

# Tartalomjegyzék

<b>ELŐSZÓ</b> .....	<b>5</b>
<b>BEVEZETÉS (SZ.J., B.J.)</b> .....	<b>6</b>
Hogyan használjuk a könyvet? .....	6
Hogyan használjuk a CD-t? .....	6
A TULIPP gyakorló program kezelése .....	7
A digitális képfeldolgozás fogalma .....	9
Történeti áttekintés .....	10
Ellenőrző kérdések .....	12
<b>1. AZ EMBERI LÁTÁS (H.CS., B. J.)</b> .....	<b>14</b>
1.1. Bevezetés .....	14
1.2. A látás alapjai .....	14
1.3. A látás biológiája .....	15
1.4. A látás sajátosságai .....	16
1.5. Színlátás .....	18
1.6. Alakzatlátás .....	20
Ellenőrző kérdések .....	22
<b>2. A KÉPFELDOLGOZÁS ESZKÖZEI (H.CS., SZ.J., B.J.)</b> .....	<b>23</b>
2.1. Bevezetés .....	23
2.2. Képbevitel .....	23
2.3. Képtárolás .....	29
2.4. Képfeldolgozás .....	32
2.5. Képi másolat készítése .....	37
2.6. Képkódolás, képtömörítés .....	38
2.7. Képatvitel .....	40
Ellenőrző kérdések .....	40
<b>3. DIGITÁLIS KÉPALKOTÁS (SZ.J.)</b> .....	<b>42</b>
3.1. Bevezetés .....	42
3.2. Leképezés .....	42
3.3. Mintavételezés .....	43
<i>Gyakorlatok a mintavételezéshez</i> .....	45

3.4. Kvantálás .....	45
<i>Gyakorlatok a kvantáláshoz</i> .....	47
3.5. Folytonos kép helyreállítása .....	48
3.6. A digitális képalkotás és a kép minősége .....	49
Ellenőrző kérdések .....	49
<b>4. KÉPJAVÍTÁS (SZ.J.)</b> .....	<b>50</b>
4.1. Bevezetés .....	50
4.2. A képjavítással kapcsolatos alapfogalmak .....	50
4.3. Képhelyreállítás .....	52
4.4. Világosságkód-transzformációk .....	53
<i>Gyakorlatok a világosságkód-transzformációkhoz</i> .....	56
4.5. Zajelnyomás .....	57
<i>Gyakorlatok a zajelnyomáshoz</i> .....	64
4.6. Élkiemelés .....	65
<i>Gyakorlatok az élkiemeléshez</i> .....	72
4.7. Többsávós képek javítása .....	73
<i>Gyakorlatok a többsávós képek transzformációjához</i> .....	74
Ellenőrző kérdések .....	74
<b>5. GEOMETRIAI KORREKCIÓ (H.CS.)</b> .....	<b>76</b>
5.1. Bevezetés .....	76
5.2. A geometriai korrekció modellje .....	77
5.3. Koordinátatranszformációk .....	78
5.4. Gyakorlati megvalósítás .....	80
5.5. Újramintavételezés .....	81
<i>Gyakorlatok a Geometriai korrekcióhoz</i> .....	82
Ellenőrző kérdések .....	83
<b>6. SZEGMENTÁLÁS (K.D.)</b> .....	<b>84</b>
6.1. Bevezetés .....	84
6.2. A szegmentálás modellje .....	85
6.3. Foltkeresés .....	87
6.4. Foltelemzés .....	90
6.5. Élkeresés .....	90
6.6. Élkorrekció .....	92
<i>Gyakorlatok a szegmentáláshoz</i> .....	93
Ellenőrző kérdések .....	94

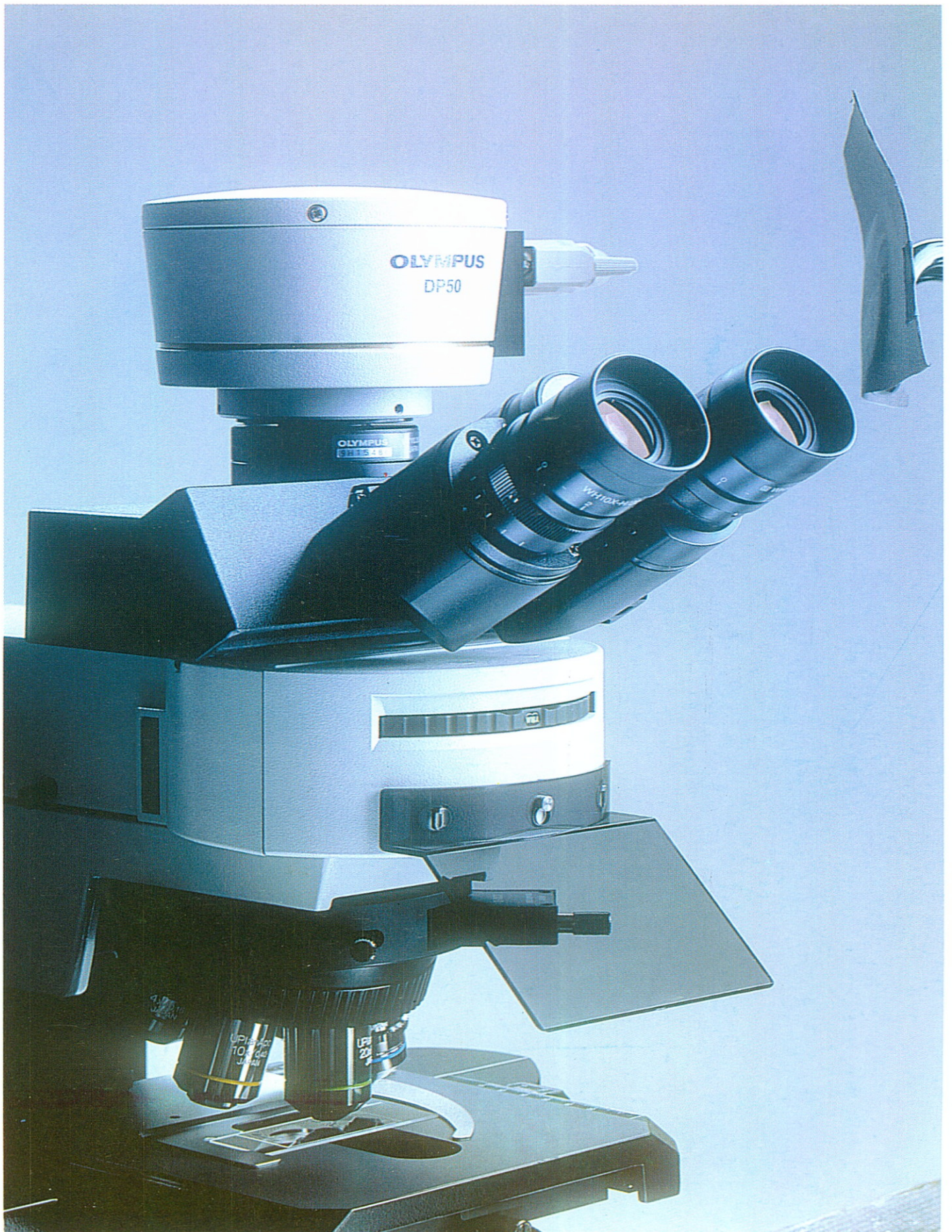
<b>7. OSZTÁLYOZÁS (K.D.)</b> .....	<b>95</b>
7.1. Bevezetés .....	95
7.2. Az osztályozás modellje .....	95
7.3. Statisztikus alakfelismerés .....	96
7.4. Klaszterezés .....	99
7.5. Környezet figyelembevétele .....	101
7.6. Szintaktikus alakfelismerés .....	101
7.7. Textúraelemzés .....	103
<i>Gyakorlatok az osztályozáshoz</i> .....	103
Ellenőrző kérdések .....	105
<b>8. KÉPKÓDOLÁS ÉS TÖMÖRÍTÉS (H.CS., B.J.)</b> .....	<b>107</b>
8.1. Bevezetés .....	107
8.2. Redundanciák .....	108
8.3. Elvárások .....	110
8.4. Veszteségmentes tömörítés .....	111
8.5. Veszteséges tömörítés .....	115
8.6. JPEG szabvány .....	118
8.7. Fraktál tömörítés .....	122
8.8. Mozgó képek tömörítése .....	123
<i>Gyakorlatok a képkódoláshoz és tömörítéshez</i> .....	125
Ellenőrző kérdések .....	127
<b>9. TÁVÉRZÉKELÉS (B.J., SZ.J., H.CS.)</b> .....	<b>128</b>
9.1. Bevezetés .....	128
9.2. Távérzékeléssel kapcsolatos alapfogalmak .....	129
9.3. Az EM sugárzást detektáló berendezések .....	134
9.4. Távérzékeléssel kapcsolatos alkalmazások .....	139
Ellenőrző kérdések .....	142
<b>10. ORVOSI ALKALMAZÁSOK (SZ.J.)</b> <sup>145</sup> <del>143</del> .....	<b>143</b>
10.1. Bevezetés .....	143
10.2. Képkalkotó berendezések képeinek javítása .....	143
10.3. Képkéértékelési feladatok .....	144
Ellenőrző kérdések .....	150
<b>11. IPARI KÉPFELDOLGOZÓ RENDSZEREK (SZ.J., B. J.)</b> .....	<b>151</b>
11.1. Bevezetés .....	151

11.2. Ipari képfeldolgozó rendszerek jellemzői .....	151
11.3. Alkalmazási példák .....	154
Ellenőrző kérdések .....	158
<b>12. MEZŐGAZDASÁGI ALKALMAZÁSOK (B.J., SZ.J.) .....</b>	<b>160</b>
12.1. Bevezetés .....	160
12.2. Képelemzési feladatok .....	162
12.3. Ipari jellegű alkalmazások .....	164
Ellenőrző kérdések .....	165
<b>13. IRODAI ALKALMAZÁS (B.J.) .....</b>	<b>166</b>
13.1. Bevezetés .....	166
13.2. Dokumentumarchiválás .....	166
13.3. Karakterfelismerés .....	167
Ellenőrző kérdések .....	169
<b>14. BŰNÜGY- ÉS BIZTONSÁGTECHNIKA (SZ.J.) .....</b>	<b>170</b>
14.1. Bevezetés .....	170
14.2. Bűnügyi alkalmazások .....	170
14.3. Biztonságtechnikai alkalmazások .....	174
Ellenőrző kérdések .....	176
<b>15. DIGITÁLIS FÉNYKÉPEZÉS (B.J.) .....</b>	<b>177</b>
15.1. Bevezetés .....	177
15.2. Történeti áttekintés .....	177
15.3. Digitális kamerák .....	178
15.4. A fotó CD .....	181
15.5. Szkennerek használata .....	183
Ellenőrző kérdések .....	188
<b>16. VIZUÁLIS ADATFORMÁTUMOK (B.J.) .....</b>	<b>190</b>
16.1 A formátum választása .....	190
16.2 Fontosabb kétdimenziós adatformátumok .....	190
<i>TIFF</i> .....	191
<i>JPEG</i> .....	191
<i>BMP</i> .....	192
<i>PCX</i> .....	193
<i>FIF</i> .....	193

<i>PICT</i> .....	193
<i>EPS</i> .....	194
<i>GIF</i> .....	194
<i>PDF</i> .....	195
<i>PhotoCD</i> .....	195
<i>MrSID</i> .....	196
<i>FLASHPIX</i> .....	197
16.3 Háromdimenziós adatformátumok .....	197
<i>VRML</i> .....	197
<i>DXF</i> .....	198
16.4 Digitális mozgókép formátumok .....	198
<i>MPEG</i> .....	198
<i>AVI</i> .....	199
<i>QUICKTIME</i> .....	200
Ellenőrző kérdések .....	202
<b>IRODALOM</b> .....	<b>203</b>



# OLYMPUS



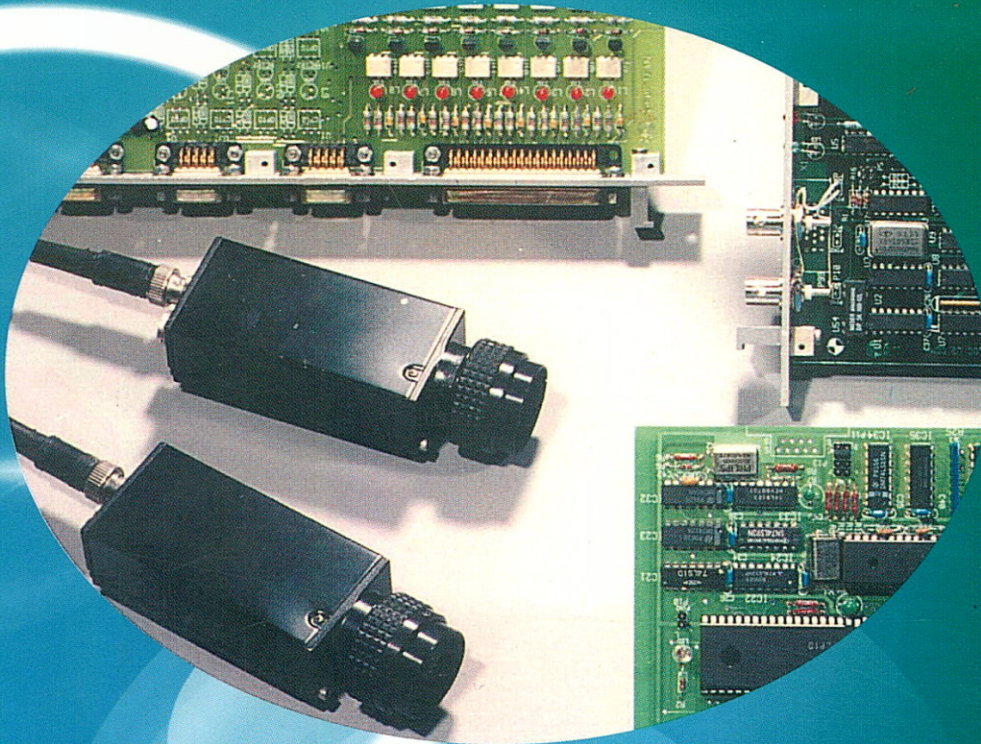
Olympus Hungary Kft. 1089 Budapest, Rezső utca 5-7.  
Tel: 459-7065, 459-7066, 459-7067, Fax: 303-6811, 303-6813  
E-mail: [OHU-Micro@olympus-europa.com](mailto:OHU-Micro@olympus-europa.com)

# Képfeldolgozó rendszerek a Falcon-Vision Kft-től

*100%-os minőség-ellenőrzés nagy sebességű gyártósorokon*



Számítógépes képfeldolgozáson alapuló, ipari környezetben, tömeggyártásban folyamatosan működtethető automata, moduláris minőség-ellenőrző és gyártásvezérlő rendszerek.



**FALCON-VISION Műszaki Fejlesztő és Szolgáltató Kft.**  
**H-1043 Budapest, Dugonics u. 11. Tel.: +36 (1) 370-0010, Fax: +36 (1) 370-0005**  
**E-mail: office@falcon.hu, home page: www.falcon.hu**

## A TARTALOMBÓL:

### A KÉPFELDOLGOZÁS ALAPJAI:

az emberi látás, képfeldolgozás eszközei, digitális képalkotás, képjavítás, geometriai korrekció, szegmentálás, osztályozás, képkódolás és tömörítés.

### KÉPFELDOLGOZÁS ALKALMAZÁSAI:

mezőgazdasági alkalmazások, orvosi alkalmazások, távérzékelés, digitális fényképezés és szkennelés, ipari képfeldolgozó rendszerek, optikai karakterfelismerés, vizuális adatformátumok, bűnügy- és biztonságtechnikai alkalmazások.

Gyakorlóprogram az egyes fejezetekhez /TULIPP/

Önellenőrzés

Interaktív keretrendszer

Egyéb hasznos segédprogramok...

<http://www.georgikon.hu/digkep.htm>

