

Számítógépes képfeldolgozás

Tófalvi Tamás

Eötvös Loránd Tudományegyetem, Informatikai Kar
tofi@inf.elte.hu



ELTE
EÖTVÖS LORÁND
TUDOMÁNYEGYETEM

Tartalom

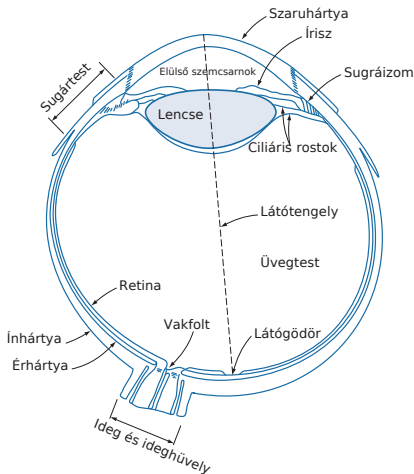
- 1 Az emberi szem
 - A szem részei
 - Fényérzékelés
 - Fényérzékelés
 - A szem tulajdonságai
- 2 Képkalkotás
 - A fény
 - Szenzorok
 - Egyszerű képkalkotási modell
 - Mintavételezés és kvantálás
 - Digitális képreprezentáció
- 3 Képjavítás
 - Intenzitás-transzformációk
 - Hisztogram-alapú módszerek
 - Több képen végzett műveletek

Tartalom

- 1 Az emberi szem
 - A szem részei
 - Fényérzékelés
 - Fényérzékelés
 - A szem tulajdonságai
- 2 Képkalkotás
 - A fény
 - Szensororok
 - Egyszerű képkalkotási modell
 - Mintavételezés és kvantálás
 - Digitális képreprezentáció
- 3 Képjavítás
 - Intenzitás-transzformációk
 - Hisztogram-alapú módszerek
 - Több képen végzett műveletek

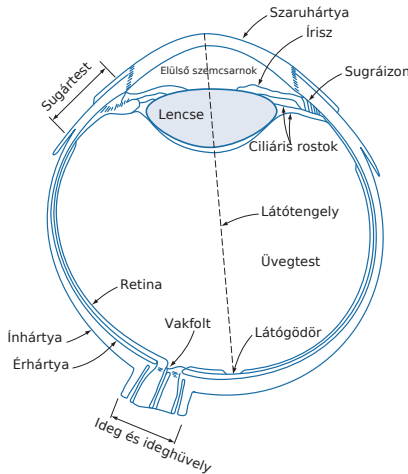
Az emberi szem

- **írisz:** a beeső fény mennyiségét szabályozza, a *pupilla* átmérője 2–8 mm
- **lencse:** rugalmas, a ciliáris izmok alakítják $\Rightarrow$ változó fókusz távolság (14–17 mm)
- **retina:** itt keletkezik a kép, fényérzékelők borítják
- **látógödör (fovea):** $\approx 1,5$ mm átmérőjű bemélyedés, a legélesebb látás helye
- **vakfolt:** a látóideg kilépési pontja, nincs benne receptor



Fényérzékelés

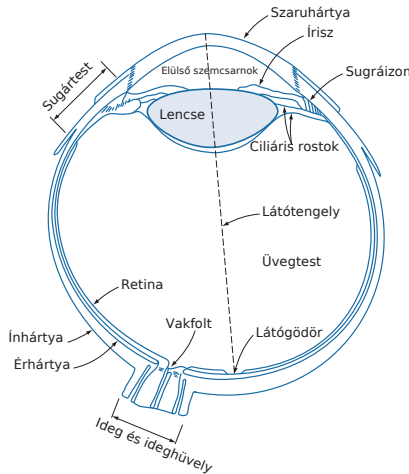
- **Pálcikák:** 75–150 millió, a retina teljes felületén
 - nincs színérzékelés, alacsony fényszintre érzékenyek
 - több pálcika közös idegvégződésre
⇒ kevés részlet
 - *szkotopikus* (szürkületi) látás
- **Csapok:** 6–7 millió, főleg a foveában
 - színérzékelés
 - minden csaphoz saját idegvégződés
⇒ finom részletek
 - *fotopikus* (nappali) látás



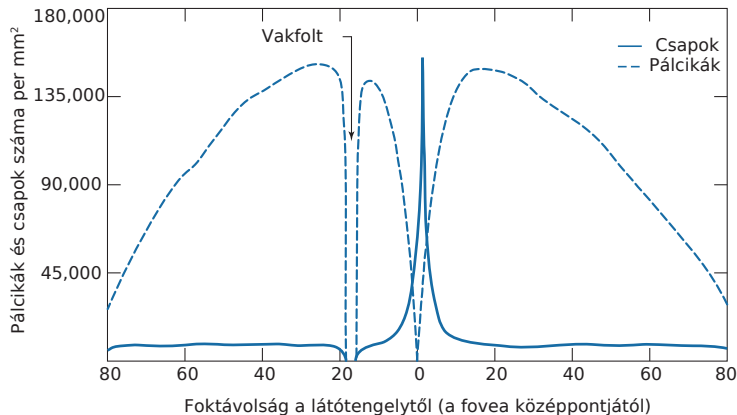
Fényérzékelés

- A szem „felbontása”:

- a foveában $\approx 150\,000$ csap/mm²
- összesen $\approx 265\,000$ csap
⇒ a mai szenzorok ezt jóval meghaladják

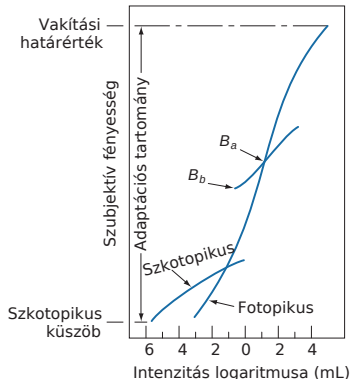


Fényérzékelés Title



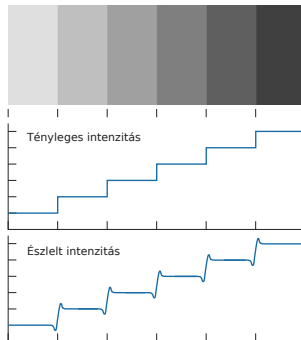
A szem tulajdonságai — fényesség

- **Dinamikatartomány:** a szkotopikus küszöbtől a vakítási határig $\approx 10^{10}$ — de *nem egyszerre!*
- **Szubjektív fényesség:** az érzékelt fényesség az intenzitás *logaritmikus* függvénye
- **Fényességi adaptáció:** a szem az összérzékenységét változtatja; az aktuális beállítás az *adaptációs szint* (B_a)
- Egy adott adaptációs szinten csak 1–2 tucat intenzitásszint különböztethető meg



A szem tulajdonságai — kontraszt

- Az észlelt fényesség *nem* az intenzitás egyszerű függvénye!
- **Mach-sávok:** az eltérő intenzitású tartományok határán a látórendszer alul-, illetve túllövi a valós értéket
⇒ az egyenletes sávok szélén világosabb/sötétebb csíkot látunk



A szem tulajdonságai — kontraszt

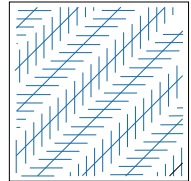
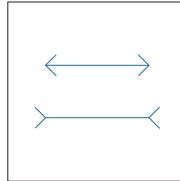
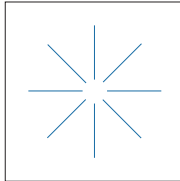
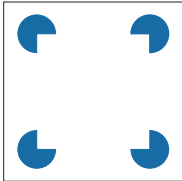
- **Szimultán kontraszt:** azonos intenzitású foltok világosabb háttéren sötétebbnek tűnnek



- **Következmény:** a képminőség megítélése szubjektív, a lokális környezet is számít

A szem tulajdonságai

- **Optikai illúziók:** a látórendszer nem létező részleteket egészít ki, vagy tévesen ítél meg geometriai tulajdonságokat
 - illuzórikus kontúrok (négyzet, kör néhány vonalból)
 - azonos hosszúságú szakaszok eltérőnek látszanak
 - párhuzamos egyenesek dőltnek tűnnek



Tartalom

- 1 Az emberi szem
 - A szem részei
 - Fényérzékelés
 - Fényérzékelés
 - A szem tulajdonságai
- 2 Képpalkotás
 - A fény
 - Szenzorok
 - Egyszerű képpalkotási modell
 - Mintavételezés és kvantálás
 - Digitális képreprezentáció
- 3 Képpjavítás
 - Intenzitás-transzformációk
 - Hisztogram-alapú módszerek
 - Több képen végzett műveletek

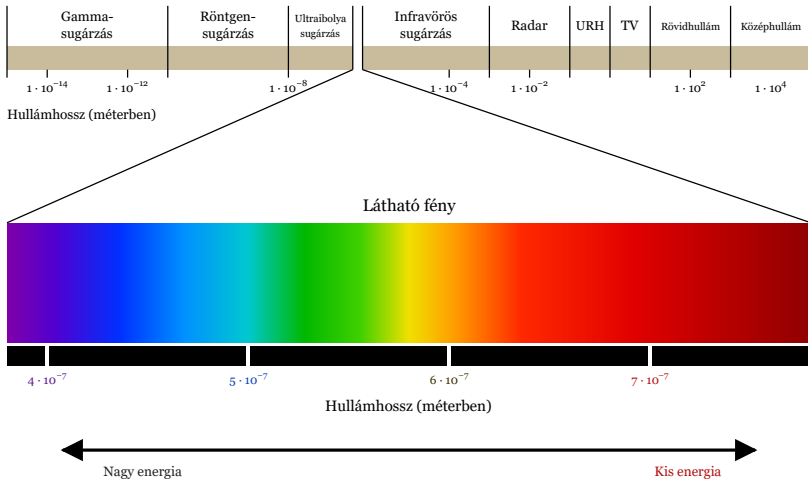
Mi a fény?

- A fény elektromágneses sugárzás, amelyet a szem érzékelni tud
- Hullámként: hullámhossz λ és frekvencia ν kapcsolata

$$\lambda\nu = c, \quad c = 2,998 \cdot 10^8 \text{ m/s}$$

- Részecskeként (foton): $E = h\nu \Rightarrow$ a rövidebb hullámhossz nagyobb energiát jelent
- **Látható tartomány:** $\approx 0,43 \mu\text{m}$ (ibolya) – $0,79 \mu\text{m}$ (vörös)
— az EM spektrum igen keskeny szelete

Mi a fény?



A fény jellemzői

- **Monokromatikus (akromatikus)** fényforrás: egyetlen jellemzője az intenzitás $\Rightarrow$ *szürkességi szint, szürkeskála*
- **Színes (kromatikus)** fényforrás leírása:
 - **Radiancia** (sugársűrűség, W): a forrásból kiáramló összes energia — *fizikai* mennyiség
 - **Luminancia** (fényssűrűség, lm): az a mennyiség, amit a megfigyelő ténylegesen érzékel a forrásból
 - pl. a távoli infravörös forrás radianciája nagy, luminanciája szinte nulla
 - **Fényesség** (brightness): *szubjektív* leíró, gyakorlatilag nem mérhető
- Egy tárgy „látásához” az EM hullám hullámhosszának a tárgy méretével összemérhetőnek vagy kisebbnek kell lennie (pl. molekulákhoz UV vagy lágy röntgen szükséges).

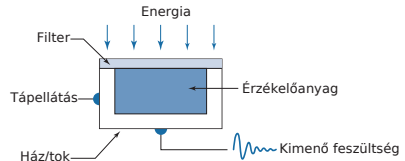
Képpalkotás

- **Mi a kép?** Kétváltozós függvény, $f(x, y)$, amelynek értéke a forrás által kibocsátott energiával arányos:

$$0 \leq f(x, y) < \infty$$

- A kép a *megvilágítás* és a jelenetről *visszavert* (vagy azon *áthaladó*) energia együttese
 - a megvilágítás lehet látható fény, röntgen, radar, ultrahang, ...
 - a jelenet lehet tárgy, molekula, közetréteg, emberi test, ...

- **Szenzor:** a beeső energiát feszültséggé alakítja
 - a kimenő feszültség hullám digitalizálásával kapunk számértéket



Szenzor típusok

- **Egyetlen érzékelő + mechanikus mozgás**

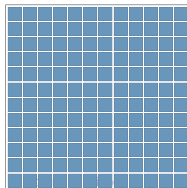
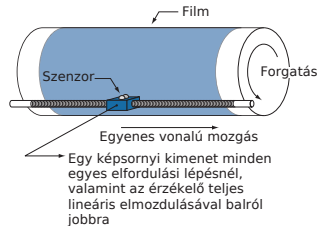
- pl. dobra feszített film, forgatás + lineáris mozgás
- nagy felbontás olcsón, viszont lassú és nem hordozható

- **Sorszenzor (line sensor)**

- a merőleges mozgás adja a másik irányt
- szkennerek, légi/műholdas felvételezés

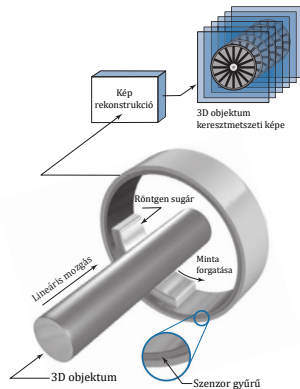
- **Mátrixszenzor (array sensor)**

- kamerák



Szenzor típusok

- Gyűrű alakú sorszenzor: CT (CAT)
 - forgó röntgenforrás + szemközti érzékelők
 - a kép nem közvetlenül áll elő:
rekonstrukciós algoritmus kell
 - azonos elven: MRI, PET $\Rightarrow$ 3D térfogati adat
- Nem EM-alapú képkalkotás:
 - ultrahang, elektronmikroszkóp, szintetikus (számítógéppel generált) képek



Egyszerű képkalkotási modell

- A képfüggvény két komponens szorzata:

$$f(x, y) = i(x, y) \cdot r(x, y)$$

ahol i a **megvilágítás** és r a **visszaverődés**

- Értéktartományok:

$$0 < i(x, y) < \infty, \quad 0 < r(x, y) < 1$$

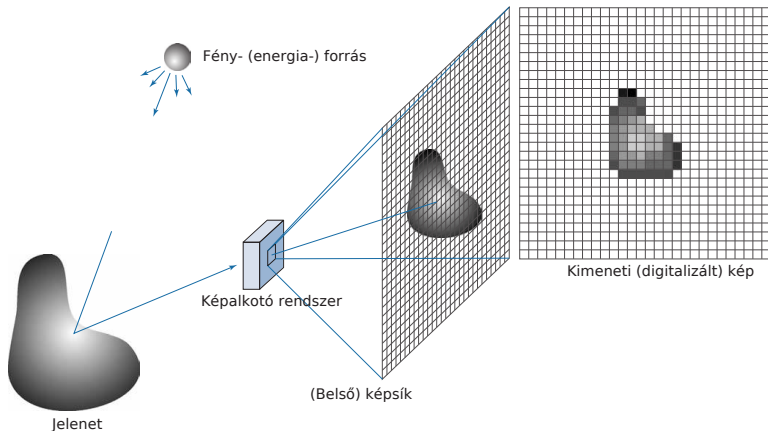
($r = 0$: teljes elnyelés, $r = 1$: teljes visszaverés)

- Bármely $\ell = f(x, y)$ pontban a képen igaz, hogy

$$L_{\min} \leq \ell \leq L_{\max}$$

- Az $[L_{\min}, L_{\max}]$ intervallum az **intenzitás-skála** (grayscale), amelyet $[0, 1]$ -re vagy $[0, L - 1]$ -re szokás normálni:
 0 = fekete, 1 (ill. $L - 1$) = fehér

Egyszerű képpalkotási modell

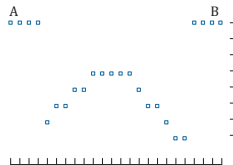
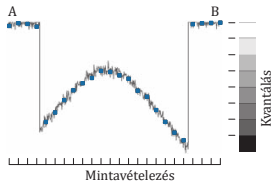
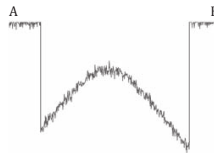
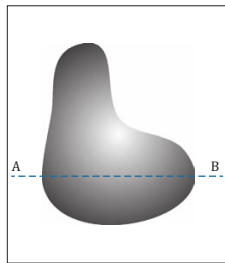


Mintavételezés és kvantálás

A folytonos jelből két lépésben lesz digitális kép:

- **Mintavételezés (sampling):** a *koordináták* diszkrétizálása
- **Kvantálás (quantization):** az *amplitúdó* (intenzitás) diszkrétizálása
- A digitális kép minőségét a minták száma és az intenzitásszintek száma együtt határozza meg (a képtartalom mellett)

Mintavételezés és kvantálás



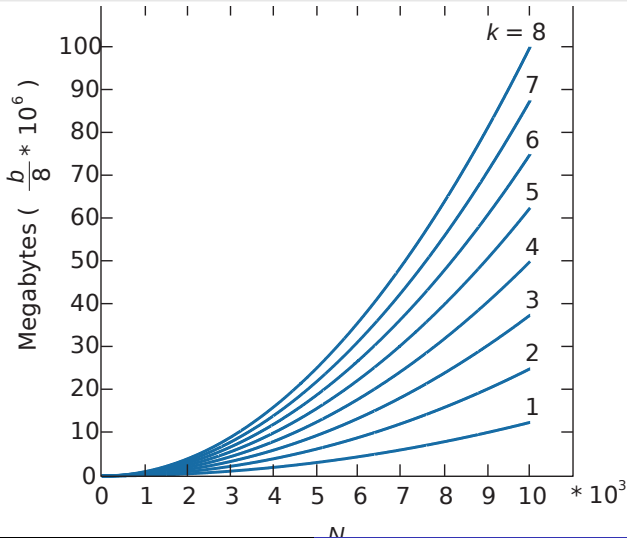
Digitális képreprezentáció

- A folytonos $f(s, t)$ függvényből M sorból és N oszlopból álló $f(x, y)$ digitális kép lesz:

$$x = 0, 1, \dots, M - 1, \quad y = 0, 1, \dots, N - 1$$

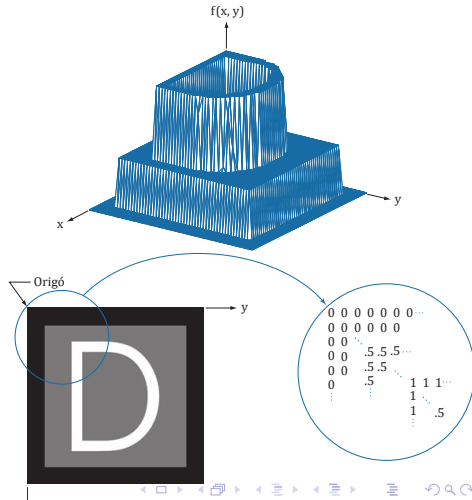
- A koordináták által kifeszített tartomány a **térbeli tartomány** (spatial domain), x és y a térbeli koordináták
- Az intenzitásintékek száma jellemzően kettő hatványa: $L = 2^k$ (pl. $k = 8 \Rightarrow 256$ szint)
- A kép tárolásához szükséges bitek száma: $b = M \cdot N \cdot k$

Kép tárhely



Digitális képreprezentáció

- Három szokásos képreprezentáció:
 - felületdiagramm
 - számmátrix
 - szokásos képi megjelenítés



Tartalom

- 1 Az emberi szem
 - A szem részei
 - Fényérzékelés
 - Fényérzékelés
 - A szem tulajdonságai
- 2 Képkalkotás
 - A fény
 - Szenzororok
 - Egyszerű képkalkotási modell
 - Mintavételezés és kvantálás
 - Digitális képreprezentáció
- 3 Képjavítás
 - Intenzitás-transzformációk
 - Hisztogram-alapú módszerek
 - Több képen végzett műveletek

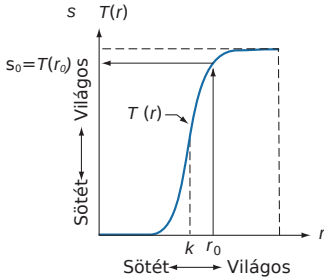
Pontműveletek

- **Képjavítás (image enhancement):** a kép módosítása úgy, hogy egy adott feladatra alkalmasabb legyen — a cél *szubjektív*
- **Pontművelet (intenzitás-transzformáció):** az új pixelérték *kizárólag* a régi pixelértéktől függ, a szomszédoktól nem:

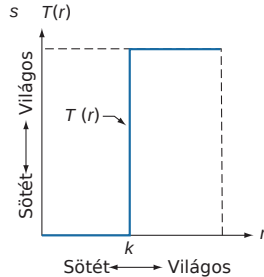
$$s = T(r), \quad r, s \in [0, L - 1]$$

- A transzformáció egyetlen **görbével** (vagy L elemű kikeresési táblával, LUT) megadható $\Rightarrow$ nagyon gyors
- Alap feladatok:
 - sötétítés/világosítás
 - kontrasztnövelés
 - küszöbölés
- Ezzel szemben a *szomszédsági* műveletek (szűrők) egy környezetből számolnak

Pontműveletek



- k alatti/feletti intenzitások sötétebbek/világosabbak lesznek
- nő a *kontraszt*

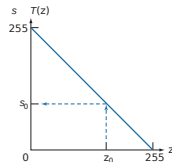


- küszöbölés – k alatti értékek eldobása
- bináris képet kapunk

Negatív

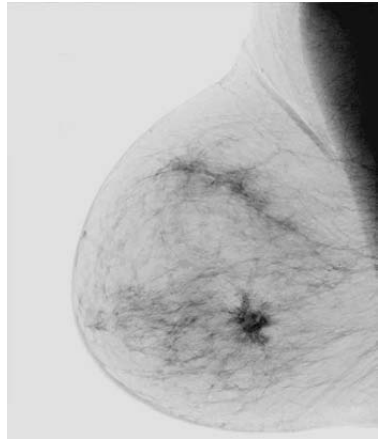
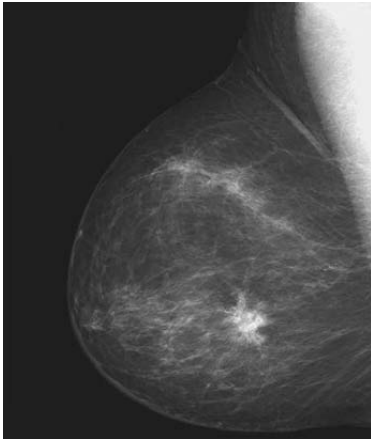
- A fotográfiai negatív digitális megfelelője:

$$s = L - 1 - r \quad (8 \text{ biten: } s = 255 - r)$$



- Az intenzitásskala megfordítása: a sötét részek világosak lesznek és fordítva
- **Mire jó?** Ha a lényegi információ a *sötét* tartományban van, amelyben a szem rosszul különböztet meg szinteket
 - orvosi képek: mammográfia, röntgen
 - nyomtatásnál a világos háttér kevesebb festéket igényel
- Információt nem veszítünk: a művelet *invertálható*

Negatív



Logaritmikus transzformáció

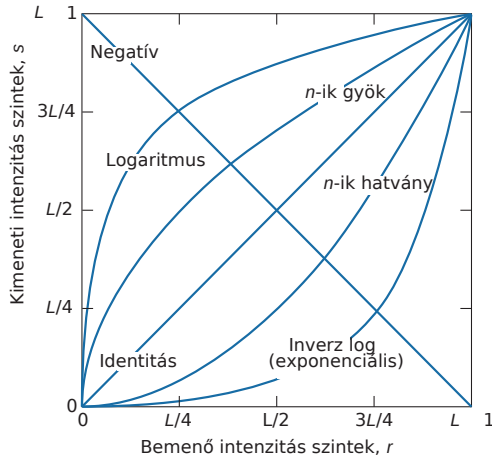
- Alakja:

$$s = c \log(1 + r), \quad c = \frac{L - 1}{\log(L)}$$

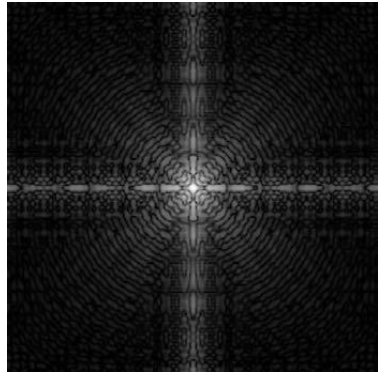
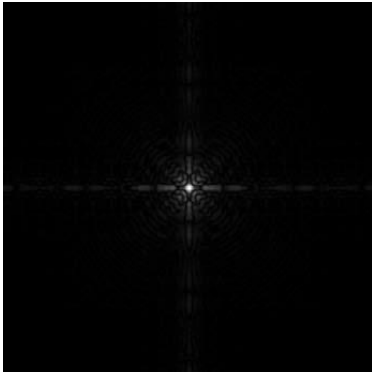
(az $1 + r$ miatt $r = 0$ is értelmezett)

- A *sötét* tartományt széthúzza, a *világosat* összenyomja $\Rightarrow$ előhozza a sötét részletek szerkezetét
- Fő alkalmazás:** nagy dinamikatartományú adatok megjelenítése 8 biten
 - pl. Fourier-spektrum: az értékek 0 és 10^6 között szórnak, lineárisan megjelenítve csak néhány világos pont látszik
- Kapcsolódik a szem szubjektív fényesség-görbéjéhez (logaritmikus!)
- Inverz logaritmus:** a világos tartományt húzza szét

Logaritmikus transzformáció



Logaritmikus transzformáció



Hatványfüggvény (gamma) transzformáció

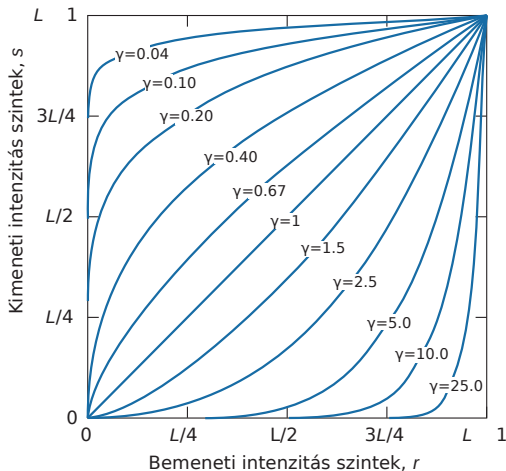
- Alakja:

$$s = c r^{\gamma}$$

(gyakran a $[0, 1]$ -re normált intenzitásokon)

- Egyetlen paraméterrel hangolható görbecsalád:
 - $\gamma < 1$: világosít, a sötét részleteket húzza szét (a log-transzformációhoz hasonlóan)
 - $\gamma = 1$: azonosság
 - $\gamma > 1$: sötétít, a világos részleteket húzza szét
- **Gamma-korrekción:** számos, képkalkotásra, nyomtatásra és megjelenítésre használt eszköz válaszfüggvénye hatványfüggvényt követ (power-law)

Hatványfüggvény (gamma) transzformáció

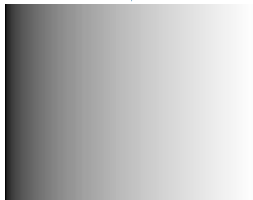


Gamma-korrekció

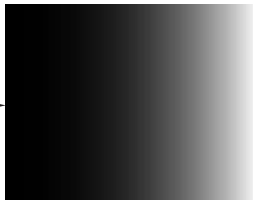


Eredeti kép

Gamma korrekció



Gamma-korrigált kép



Eredeti kép egy monitoron megjelenítve
2,5-ös gammával



Gamma-korrigált kép ugyanazon a
monitoron megjelenítve

Gamma-korrekció

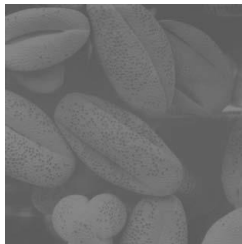
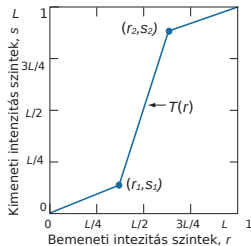


Kontrasztnyújtás és küszöbölés

- **Szakaszonként lineáris** transzformáció két törésponttal, (r_1, s_1) és (r_2, s_2) :
 - a köztes tartományt széthúzza
- az alacsony kontrasztú kép okai: rossz megvilágítás, szűk szenzor-dinamika, rossz rekeszéérték
- Szokásos választás: $(r_1, s_1) = (r_{min}, 0)$ és $(r_2, s_2) = (r_{max}, L - 1)$ („min-max nyújtás”)
- Határeset: ha $r_1 = r_2 = k$ és $s_1 = 0, s_2 = L - 1$, akkor **küszöbölés** (bináris kép):

$$s = \begin{cases} 0, & r \leq k \\ L - 1, & r > k \end{cases}$$

Kontrasztnyújtás és küszöbölés



A hisztogram

- A hisztogram az r_k intenzitásszintek előfordulási gyakorisága:

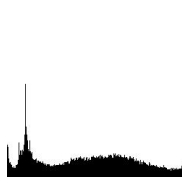
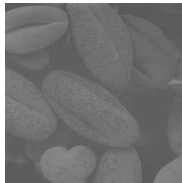
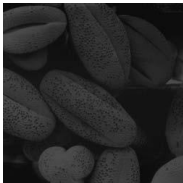
$$h(r_k) = n_k, \quad k = 0, 1, \dots, L-1$$

- **Normalizált hisztogram** — az intenzitás tapasztalati valószínűség-eloszlása:

$$p(r_k) = \frac{n_k}{MN}, \quad \sum_{k=0}^{L-1} p(r_k) = 1$$

- Nem tartalmaz *térbeli* információt: két teljesen különböző képnek lehet azonos a hisztogramja
- A hisztogram alakja árulkodó:
 - sötét kép: a bal oldalra tömörül; világos kép: a jobb oldalra
 - alacsony kontraszt: keskeny csúcs középen
 - jó kontraszt: az egész skálán, közel egyenletesen elterül

A hisztogram



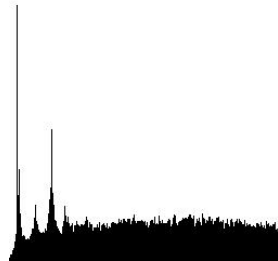
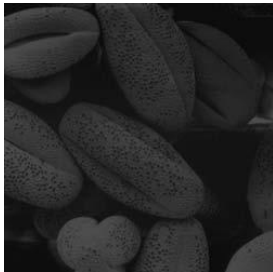
Hisztogram-kiegyenlítés

- **Cél:** automatikusan olyan transzformációt találni, amely után a hisztogram közel *egyenletes* $\Rightarrow$ a teljes skálát kihasználjuk
- **Ötlet:** a transzformáció legyen a kumulatív eloszlásfüggvény (CDF):

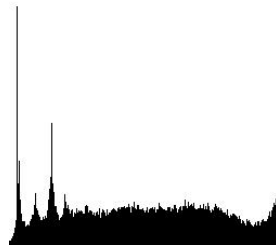
$$s_k = T(r_k) = (L - 1) \sum_{j=0}^k p(r_j)$$

- Tulajdonságok:
 - monoton növekvő $\Rightarrow$ a sorrend (sötétebb–világosabb) megmarad
 - paramétermentes, *automatikus* — a képből számolódik
 - diszkrét esetben csak közelítőleg egyenletes (a szintek összeolvadnak)
- **Veszély:** a zajt és a háttérszemcsét is felerősíti, túl „agresszív” lehet

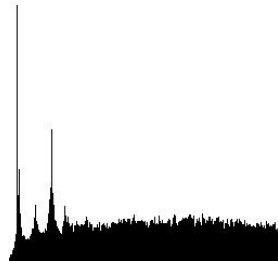
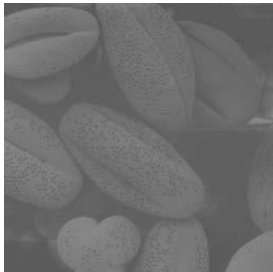
Hisztogram-kiegyenlítés



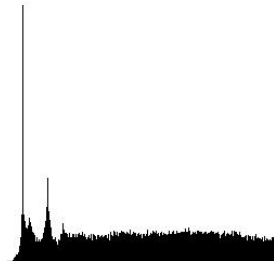
Hisztogram-kiegyenlítés



Hisztogram-kiegyenlítés



Hisztogram-kiegyenlítés



Hisztogram-illesztés és lokális kiegyenlítés

- **Hisztogram-illesztés (specification/matching):** nem egyenletes, hanem *előírt* alakú hisztogramot célzunk meg
 - a kép és a cél CDF-jét is kiszámoljuk, majd $s \mapsto G^{-1}(T(r))$
 - akkor hasznos, ha a kiegyenlítés eredménye nem tetszetős, vagy ha képeket akarunk egymáshoz igazítani
- **Lokális hisztogram-kiegyenlítés:** a hisztogramot egy csúszó ablakban számoljuk
 - előhozza a lokális részleteket, amelyeket a globális módszer elnyom
 - viszont drágább és erősen zajérzékeny
 - elterjedt változat: **CLAHE** (kontraszt-korlátozott, csempékre bontott adaptív kiegyenlítés)

Átlagolás zajcsökkentésre

- Additív, korrelálatlan, nulla várható értékű zaj modellje:

$$g_i(x, y) = f(x, y) + \eta_i(x, y)$$

- K darab felvétel átlaga:

$$\bar{g}(x, y) = \frac{1}{K} \sum_{i=1}^K g_i(x, y), \quad E\{\bar{g}\} = f$$

- A zaj szórása $\sqrt{K}$ -ad részére csökken:

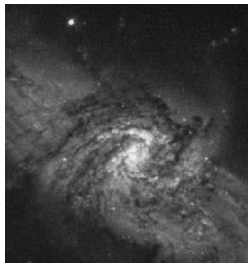
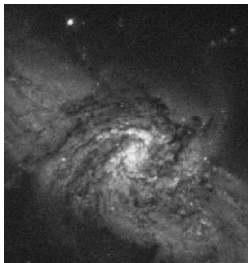
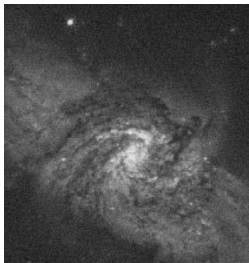
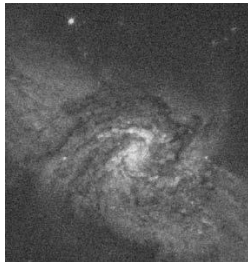
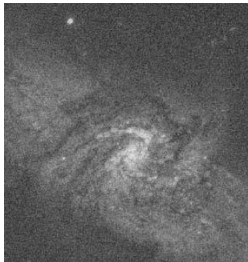
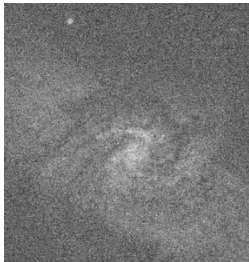
$$\sigma_{\bar{g}} = \frac{1}{\sqrt{K}} \sigma_{\eta}$$

(pl. $K = 100 \Rightarrow$ tizedére)

Átlagolás zajcsökkentésre

- **Feltétel:** a képeknek *illesztettnek* (regisztráltnak) kell lenniük, különben elmosódás lesz az eredmény
- Alkalmazás:
 - csillagászati felvételek
 - gyenge fényviszonyok
 - mobiltelefonos éjszakai mód

Átlagolás zajcsökkentésre

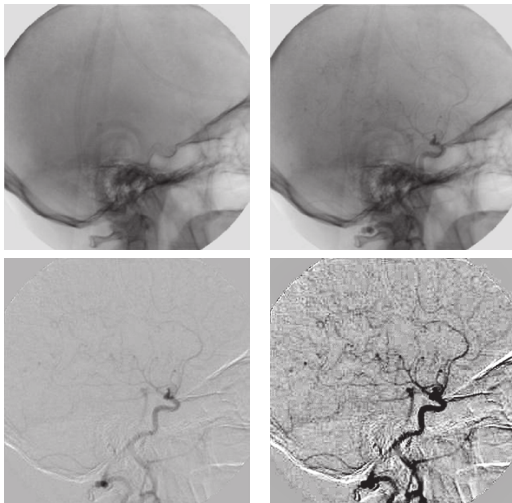


Képek kivonása, háttérkivonás

- Elemenkénti kivonás: $g(x, y) = f(x, y) - h(x, y)$
- **Háttérkivonás:** h a referencia („üres”) felvétel $\Rightarrow$ csak a *változás* marad
 - mozgásdetektálás, minőségellenőrzés, változáselemzés
 - nehézség: a háttér maga is változik (fény, árnyék, kamera-elmozdulás) $\Rightarrow$ a háttérmodellt frissíteni kell
- **Mask mode radiography** (digitális szubtrakciós angiográfia):
kontrasztanyag beadása előtti maszkkép kivonása a későbbi felvételekből $\Rightarrow$ csak az erek látszanak
- **Tartományprobléma:** a különbség $[-255, 255]$ közé esik $\Rightarrow$ eltolás és skálázás kell a megjelenítéshez:

$$g_m = g - \min(g), \quad g_s = K \frac{g_m}{\max(g_m)}$$

Háttérkivonás: Mask mode radiography



Háttérkivonás: mozgásdetektálás



Szorzás és osztás: árnyékkorrekción, maszkolás

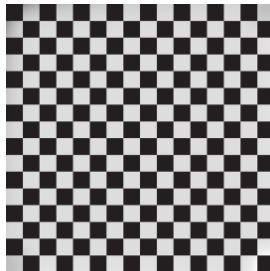
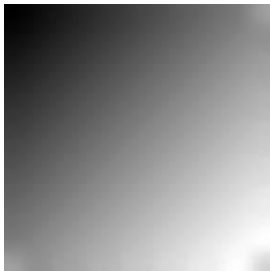
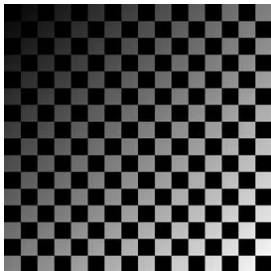
- Egyenetlen megvilágítás / vignettálás modellje:

$$g(x, y) = f(x, y) \cdot h(x, y)$$

ahol h az árnyékolási függvény

- Ha h ismert vagy becsülhető, elemenkénti osztással korrigálható:
 $f \approx g/h$
 - h mérhető: egyenletes fehér felület lefényképezésével („flat field”)
 - a nullával osztás elkerülésére kis konstans adunk a nevezőhöz
- **Maszkolás (ROI):** a képet egy 0/1 értékű maszkkal szorozzuk $\Rightarrow$ csak az érdekes régió marad meg

Szorzás és osztás: árnyékkorrekció



Szorzás és osztás: maszkolás

