

7. Sarokpontok detektálása

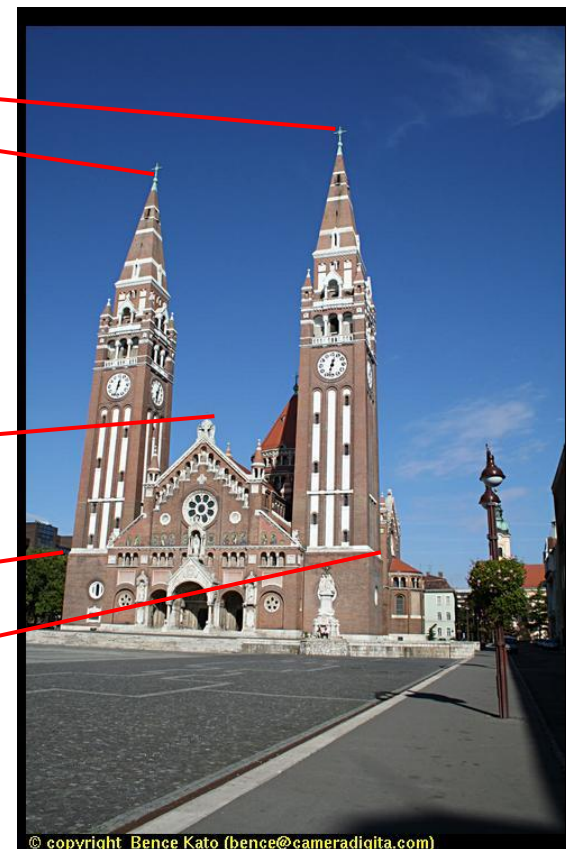
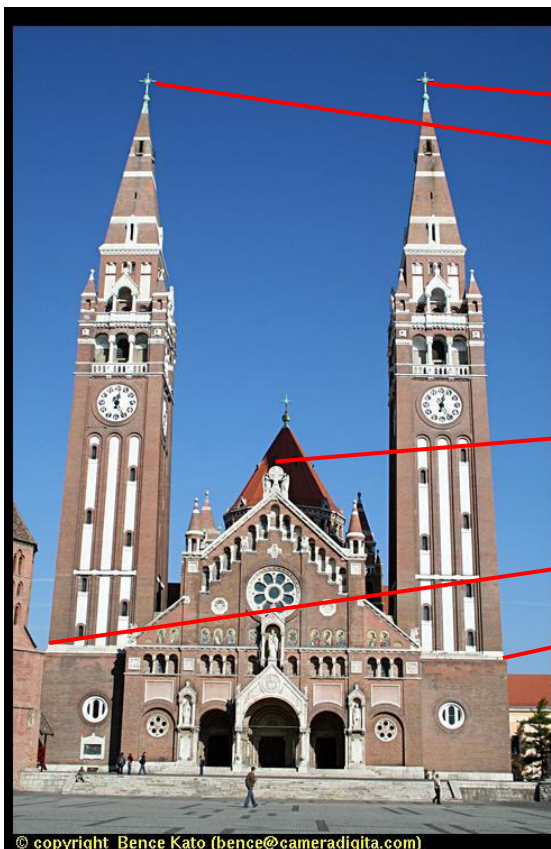
Kató Zoltán

**Képfeldolgozás és Számítógépes Grafika tanszék
SZTE**

(<http://www.inf.u-szeged.hu/~kato/teaching/>)

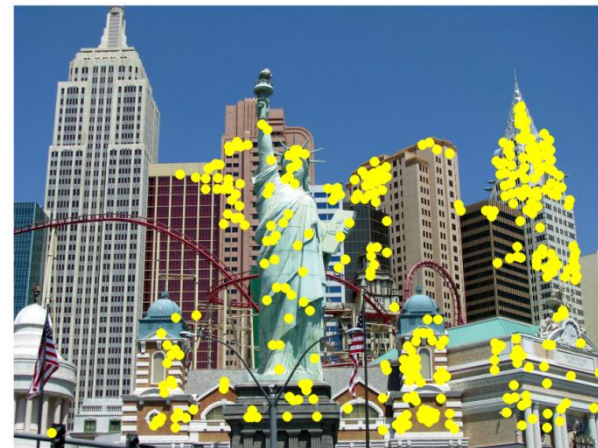
Jellemzők és megfeleltetésük

- A képfeldolgozás, számítógépes látás számos területén felmerülő probléma egy látványról készült képpár közötti pontmegfeleltetések megkeresése.
 - Ehhez megbízható jellemzőket (tipikusan sarokpontokat) nyerünk ki a képekről
 - Majd a kinyert pontokat megfeleltetjük egymásnak.

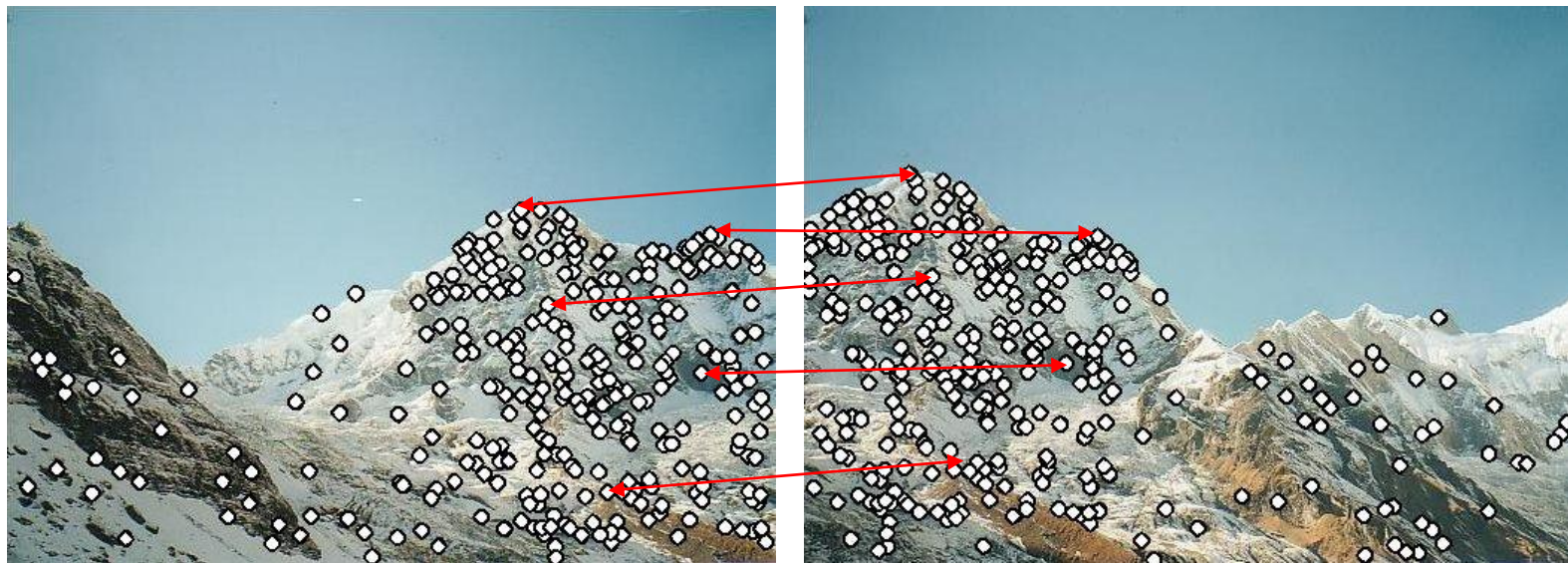


Alkalmazások

- Képek illesztése (regisztráció)
- 3D rekonstrukció (sztereo)
- Mozgáskövetés
- Robot navigáció
- Képi adatbázisok indexelése
- Objektumok azonosítása
- ...



Példa: panoráma kép készítés



1. Jellemzőpontok detektálása mindkét képen
2. Kinyert pontok megfeleltetése
3. Megfeleltetések alapján a képpár illesztése



Megválaszolandó kérdések

- Melyek azok a pontok, amelyeket megbízható módon detektálhatunk a képeken?
 - Sarokpontok
- Hogyan tudjuk leírni/jellemezni a kinyert pontokat?
 - Invariáns jellemzők
- Hogyan feleltessünk meg két képről kinyert pontokat?
 - Jellemzők összehasonlítása, robusztusság

Sarokpontok



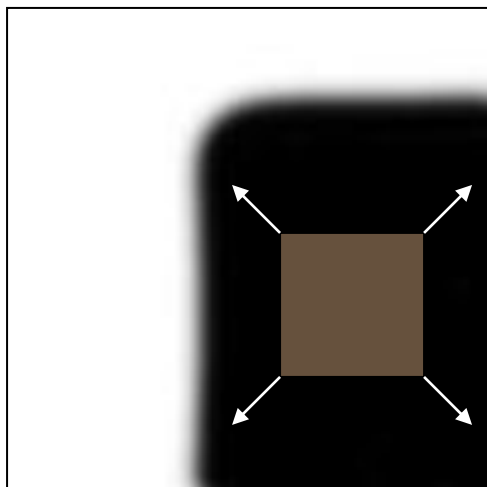
Sarokpont detektálás

- **Sarokpont = élek találkozása**
 - Élek ott vannak, ahol a képfüggvény valamely irányban nagyot változik
 - Két él találkozásánál a képfüggvény mindkét irányban nagyot változik
- **Sarokpontok jól “ellenállnak” a különböző geometriai és fotometriai deformációnak**
 - Megbízhatóan detektálhatóak,
 - Egy objektum a sarokpontjaival stabilan leírható

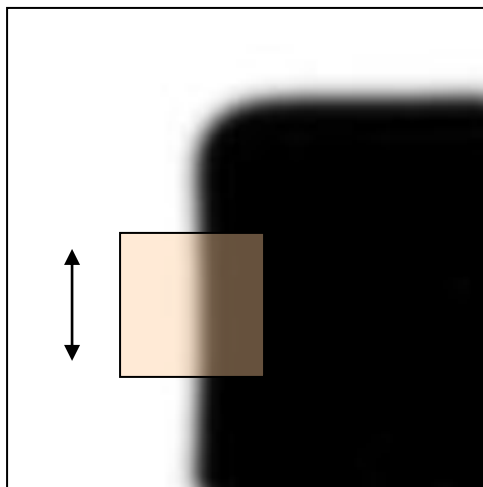
C.Harris and M.Stephens. ["A Combined Corner and Edge Detector."](#)
Proceedings of the 4th Alvey Vision Conference: pages 147--151.

Sarokpont detektálás

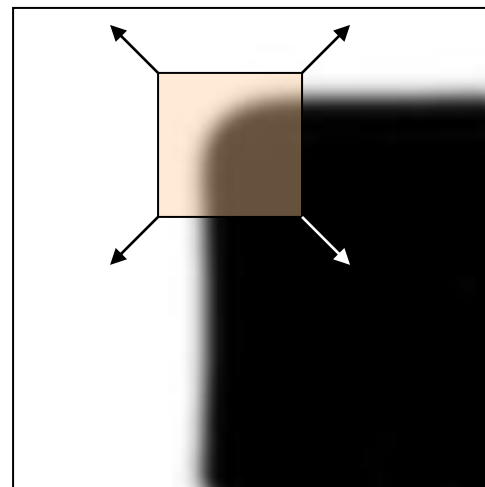
- Egy csúszóablakot vizsgálva ott lesz sarokpont, ahol az ablakot bármilyen irányba mozgatva nagy képfüggvény változást tapasztalunk



“sima” régió:
egyik irányban
sincs változás



“él”:
az él irányában
nincs változás



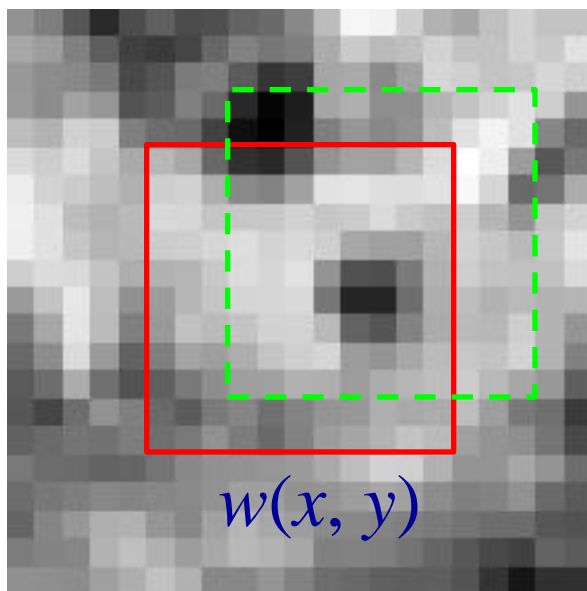
“sarok”:
jelentős változás
minden irányban

Sarokpont detektálás: matematikai modell

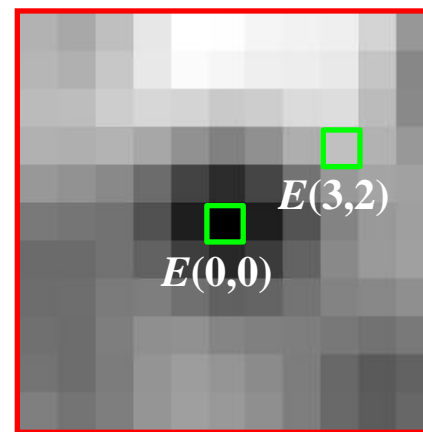
- A $w(x,y)$ ablak tartalmának változása egy $[u,v]$ eltolás hatására:
 - Lényegében auto-korrelációt számolunk

$$E(u, v) = \sum_{x,y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

$I(x, y)$



$E(u, v)$

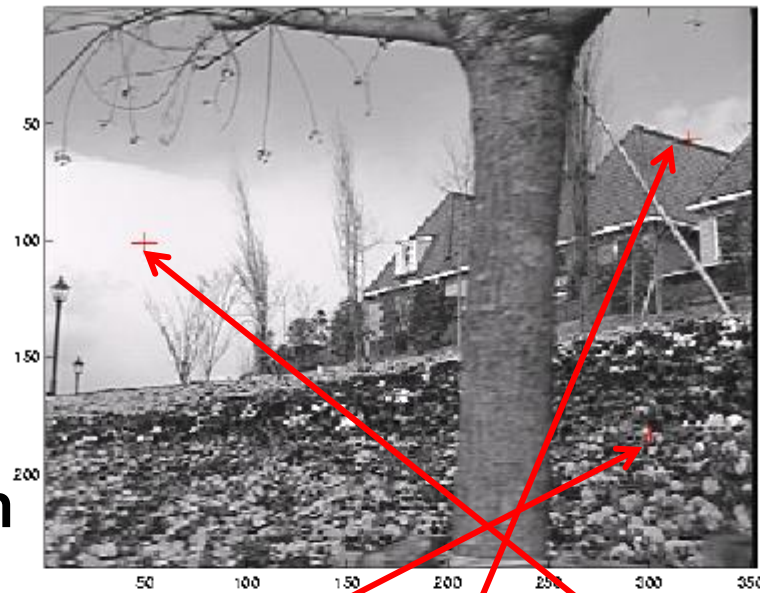


$E(u,v)$ változása

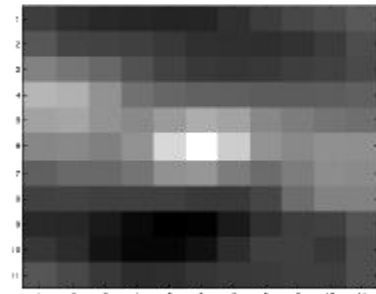
1. Megfelelően változatos intenzitás tartalom
→ jó lokális minimum

2. Élek mentén csak az egyik irányban van lokális minimum

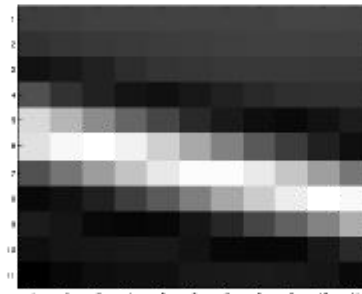
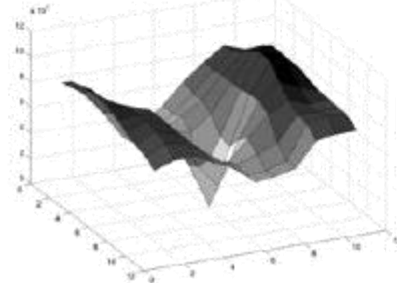
3. Homogén területen $E(u,v)$ alig változik



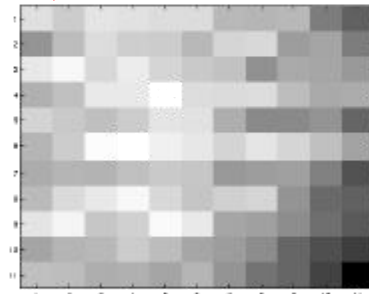
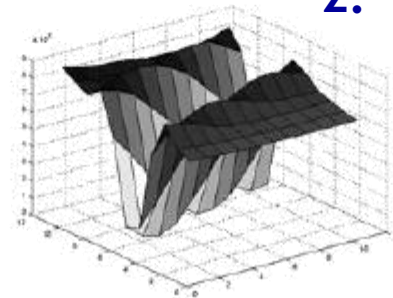
(a)



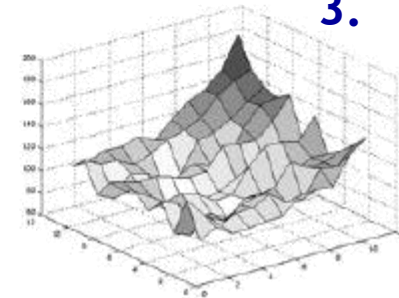
1.



2.



3.



Sarokpont detektálás: matematikai modell

- A $w(x,y)$ ablak tartalmának változása egy $[u,v]$ eltolás hatására:
 - Azt szeretnénk vizsgálni, hogy kis eltolások esetén hogyan viselkedik $E(u, v)$.

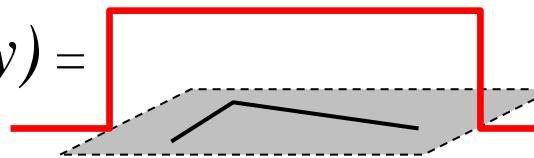
$$E(u, v) = \sum_{x,y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

Ablak
függvény

Eltolt
intenzitás

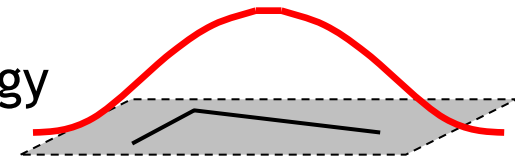
Intenzitás

Ablak (súly)függvény $w(x,y) =$



Doboz: 1 belül, 0 kívül

vagy



Gauss

$E(u, v)$ változásának jellemzése

- Fejtsük Taylor sorba $E(u, v)$ -t a $(0, 0)$ körül a másod rendű (kvadratikus) tagig:

$$E(u, v) \approx E(0, 0) + [u \ v] \begin{bmatrix} E_u(0, 0) \\ E_v(0, 0) \end{bmatrix} + \frac{1}{2} [u \ v] \begin{bmatrix} E_{uu}(0, 0) & E_{uv}(0, 0) \\ E_{uv}(0, 0) & E_{vv}(0, 0) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$E_u(u, v) = \sum_{x, y} 2w(x, y) [I(x+u, y+v) - I(x, y)] I_x(x+u, y+v)$$

$$E_{uu}(u, v) = \sum_{x, y} 2w(x, y) I_x(x+u, y+v) I_x(x+u, y+v) \\ + \sum_{x, y} 2w(x, y) [I(x+u, y+v) - I(x, y)] I_{xx}(x+u, y+v)$$

$$E_{uv}(u, v) = \sum_{x, y} 2w(x, y) I_y(x+u, y+v) I_x(x+u, y+v) \\ + \sum_{x, y} 2w(x, y) [I(x+u, y+v) - I(x, y)] I_{xy}(x+u, y+v)$$

$E(u, v)$ változásának jellemzése

- Mivel

$$E(0,0) = 0$$

$$E_u(0,0) = 0$$

$$E_v(0,0) = 0$$

$$E_{uu}(0,0) = \sum_{x,y} 2w(x,y)I_x(x,y)I_x(x,y)$$

$$E_{vv}(0,0) = \sum_{x,y} 2w(x,y)I_y(x,y)I_y(x,y)$$

$$E_{uv}(0,0) = \sum_{x,y} 2w(x,y)I_x(x,y)I_y(x,y)$$

- Ezért az alábbi közelítést kapjuk, amelyet közvetlenül a képfüggvény parciális deriváltjaival fejezhetünk ki

$$E(u, v) \approx [u \quad v] \begin{bmatrix} \sum_{x,y} w(x,y)I_x^2(x,y) & \sum_{x,y} w(x,y)I_x(x,y)I_y(x,y) \\ \sum_{x,y} w(x,y)I_x(x,y)I_y(x,y) & \sum_{x,y} w(x,y)I_y^2(x,y) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$E(u, v)$ változásának jellemzése

- A kvadratikus közelítés tehát az alábbi alakban írható:

$$E(u, v) \approx [u \ v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

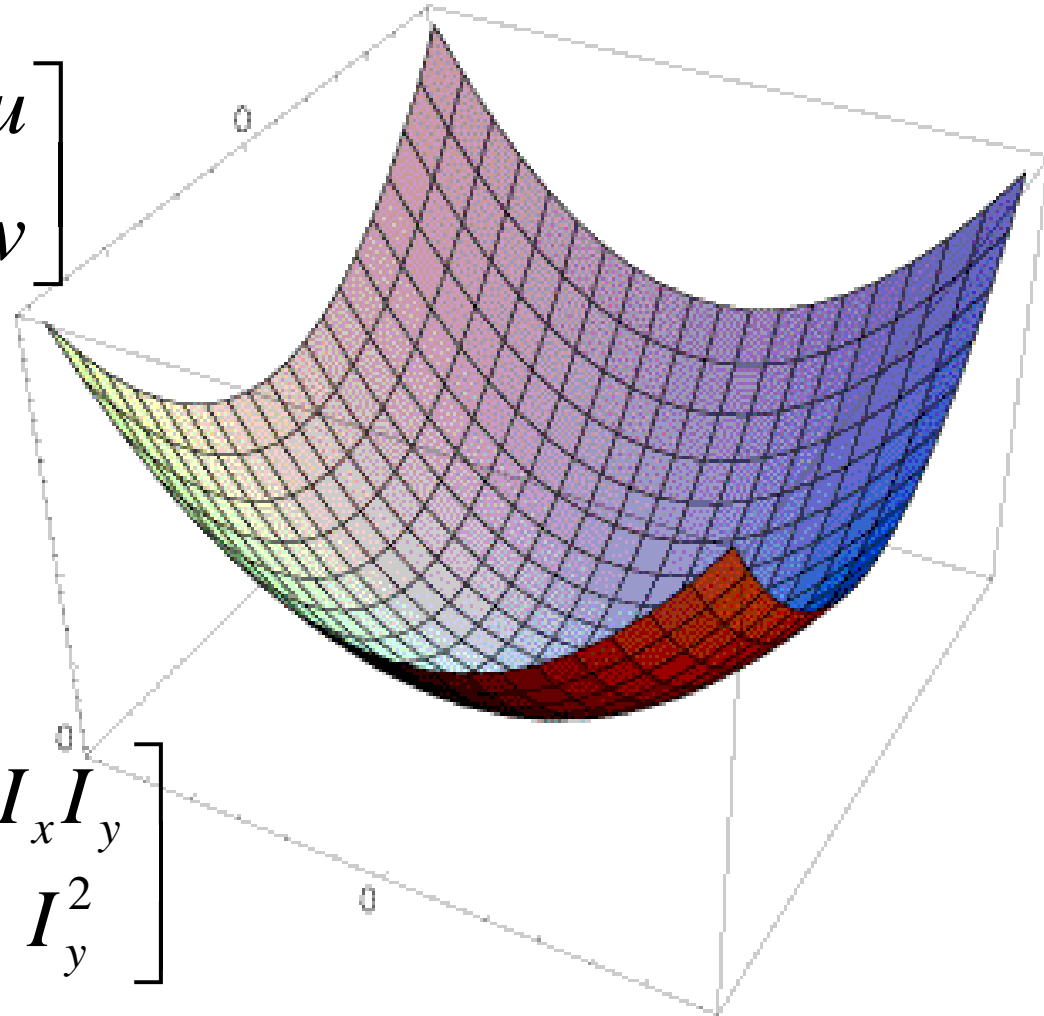
- Ahol M a képfüggvény deriváltjainak második momentum mátrixa:

$$\begin{aligned} M &= \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \\ &= \sum_{x,y} w(x, y) \begin{bmatrix} I_x \\ I_y \end{bmatrix} \begin{bmatrix} I_x & I_y \end{bmatrix} = \sum_{x,y} w(x, y) \nabla I (\nabla I)^T \end{aligned}$$

Az M második momentum mátrix

- Az $E(u, v)$ felszín lokális approximációja kvadratikus alakban

$$E(u, v) \approx [u \ v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

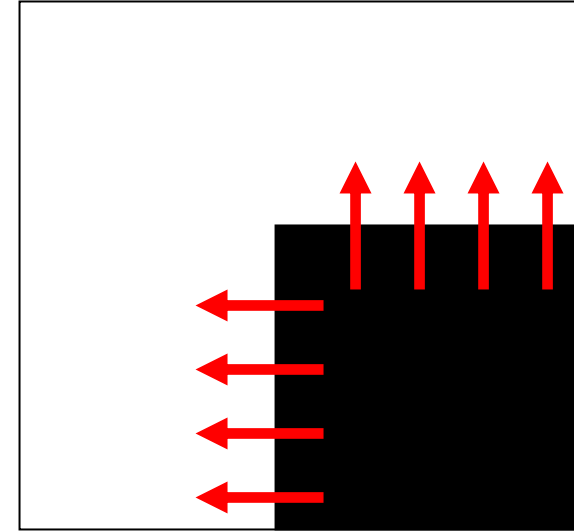


$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Tengely állású sarokpont esetén

- M diagonális lesz, hiszen a domináns gradiens irányok egybeesnek az x és y tengelyekkel

$$M = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$



- Ha valamelyik $\lambda \approx 0$, akkor ez nem sarokpont!

Általános esetben

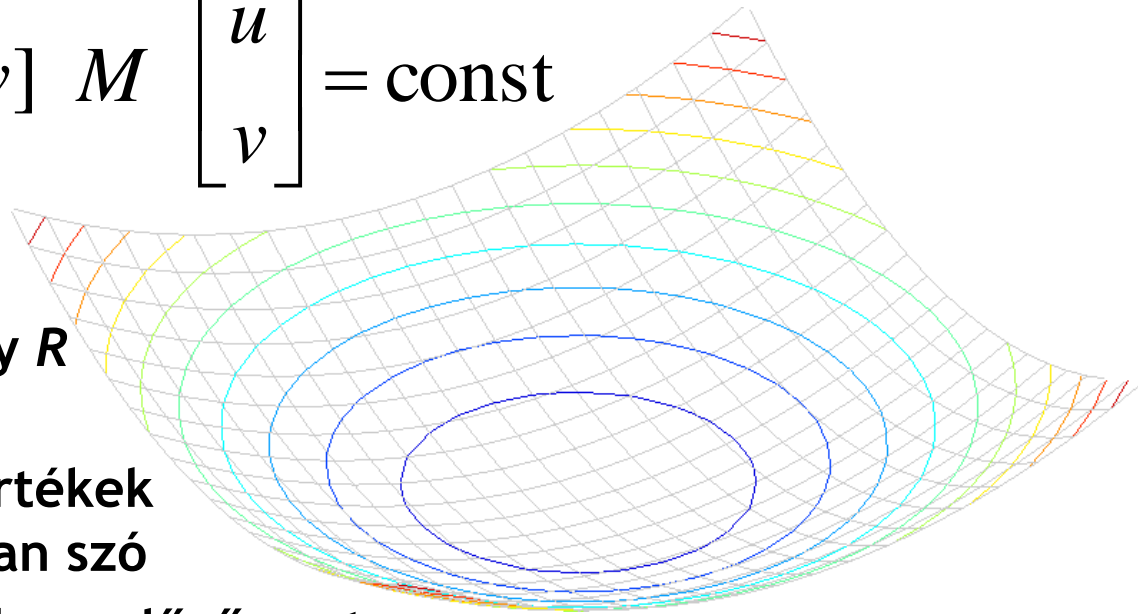
- Vegyük észre, hogy a kvadratikus approximáció következtében a felszín bármely horizontális szelete egy ellipszis lesz

$$[u \ v] M \begin{bmatrix} u \\ v \end{bmatrix} = \text{const}$$

- **M diagonalizálható**

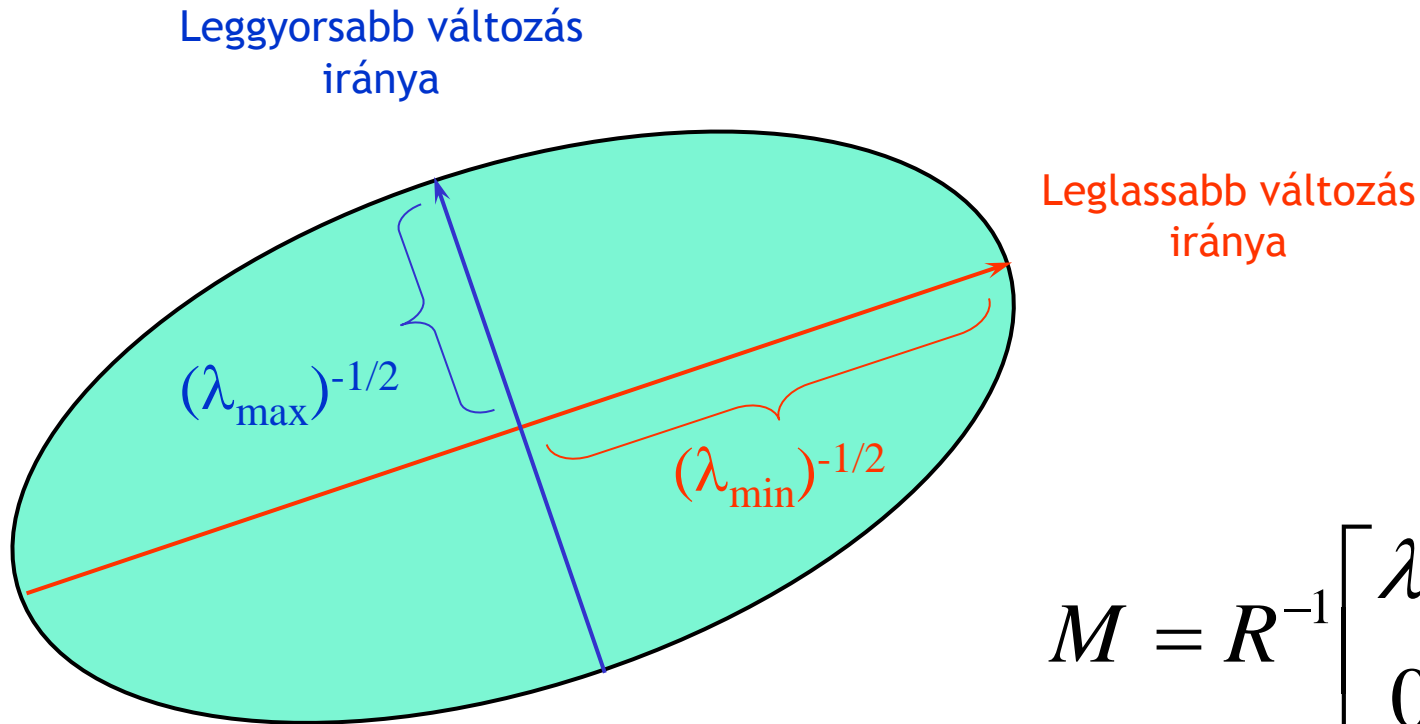
- Ez geometriailag egy R forgatást jelent
- Algebrailag a sajátértékek meghatározásáról van szó
- Ezzel visszavezettük az előző esetre

$$M = R^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R$$



M mint ellipszis jellemzése

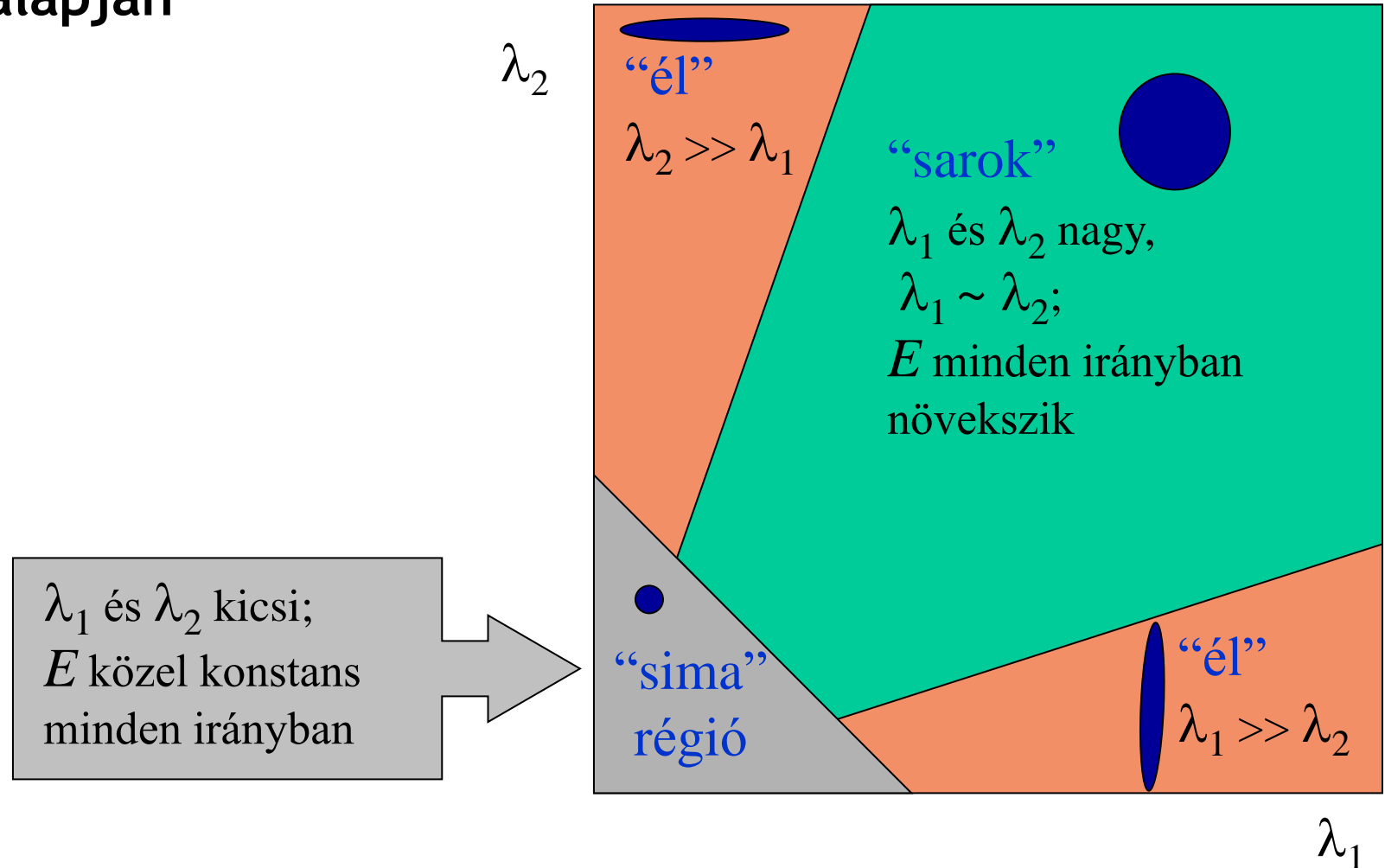
- A tengelyek hosszát a sajátértékek adják
- A tengelyek irányát pedig az R forgatási mátrix adja



$$M = R^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R$$

A sajátértékek jelentése

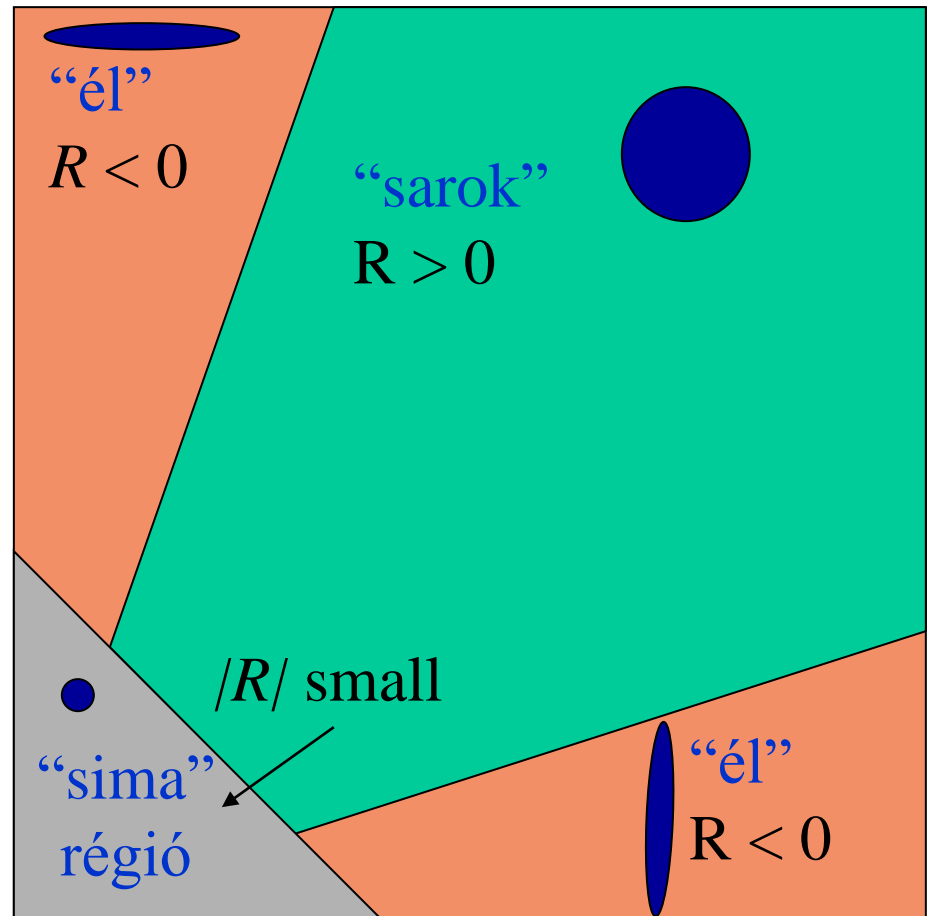
- A képünk pontjait osztályozhatjuk M sajátértékei alapján



Sarkosságot jellemző függvény

$$R = \det(M) - \alpha \text{trace}(M)^2 = \lambda_1 \lambda_2 - \alpha (\lambda_1 + \lambda_2)^2$$

α : konstans (0.04 - 0.06)



Harris sarokdetektáló algoritmus

- 1) Számítsuk ki az M mátrixot minden egyes képpont feletti ablakban és ebből megkapjuk az R sarkossági jellemzőt
- 2) Keressük meg azokat a pontokat, amelyekre a sarkossági érték elegendően nagy ($R > küszöb$)
- 3) Tartsuk meg ezekből a lokális maximumokat (vagyis nyomjuk el a nem-maximumokat)

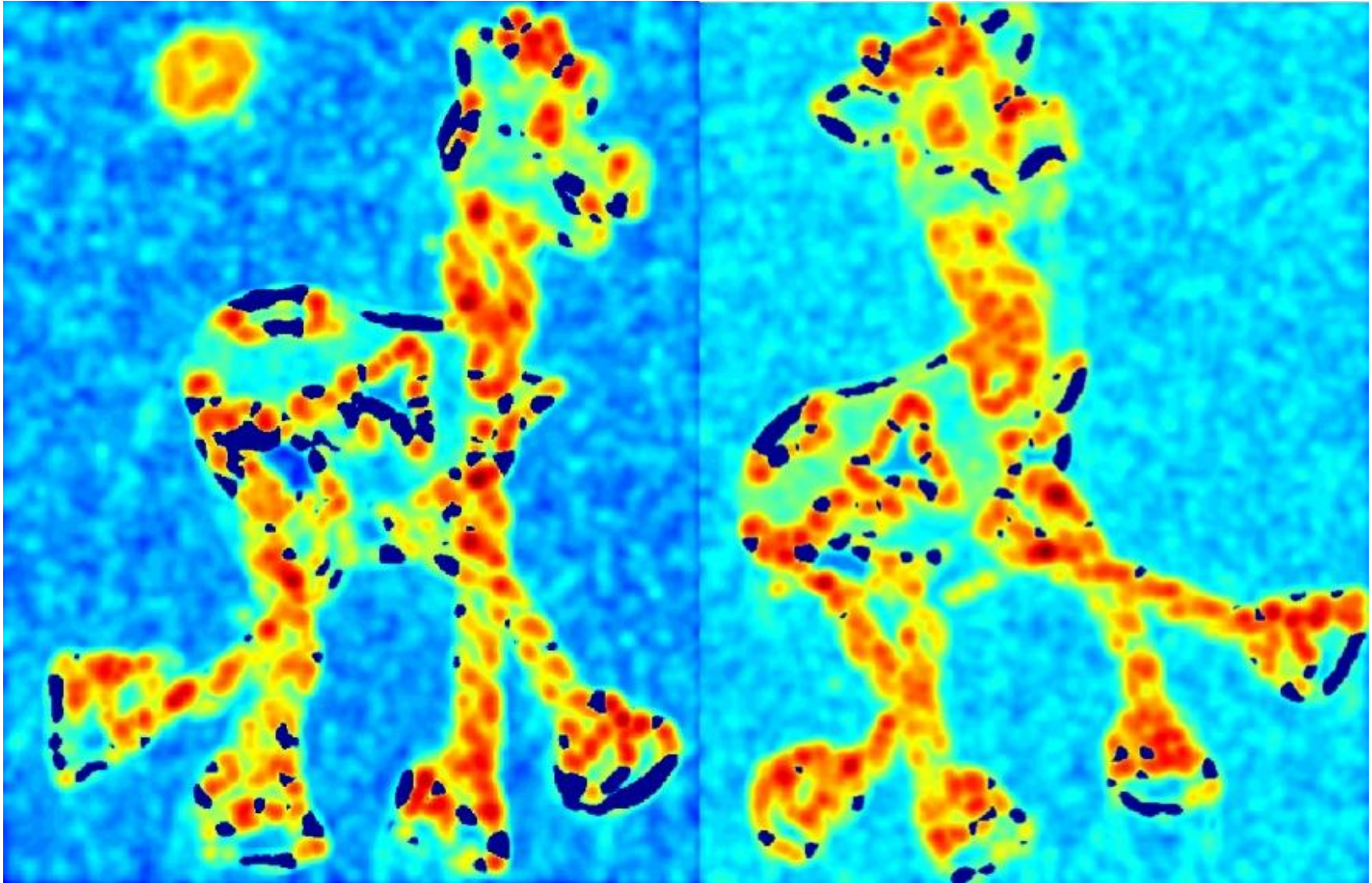
C.Harris and M.Stephens. “A Combined Corner and Edge Detector.”
Proceedings of the 4th Alvey Vision Conference: pages 147–151, 1988.

Harris sarokdetektáló algoritmus működése



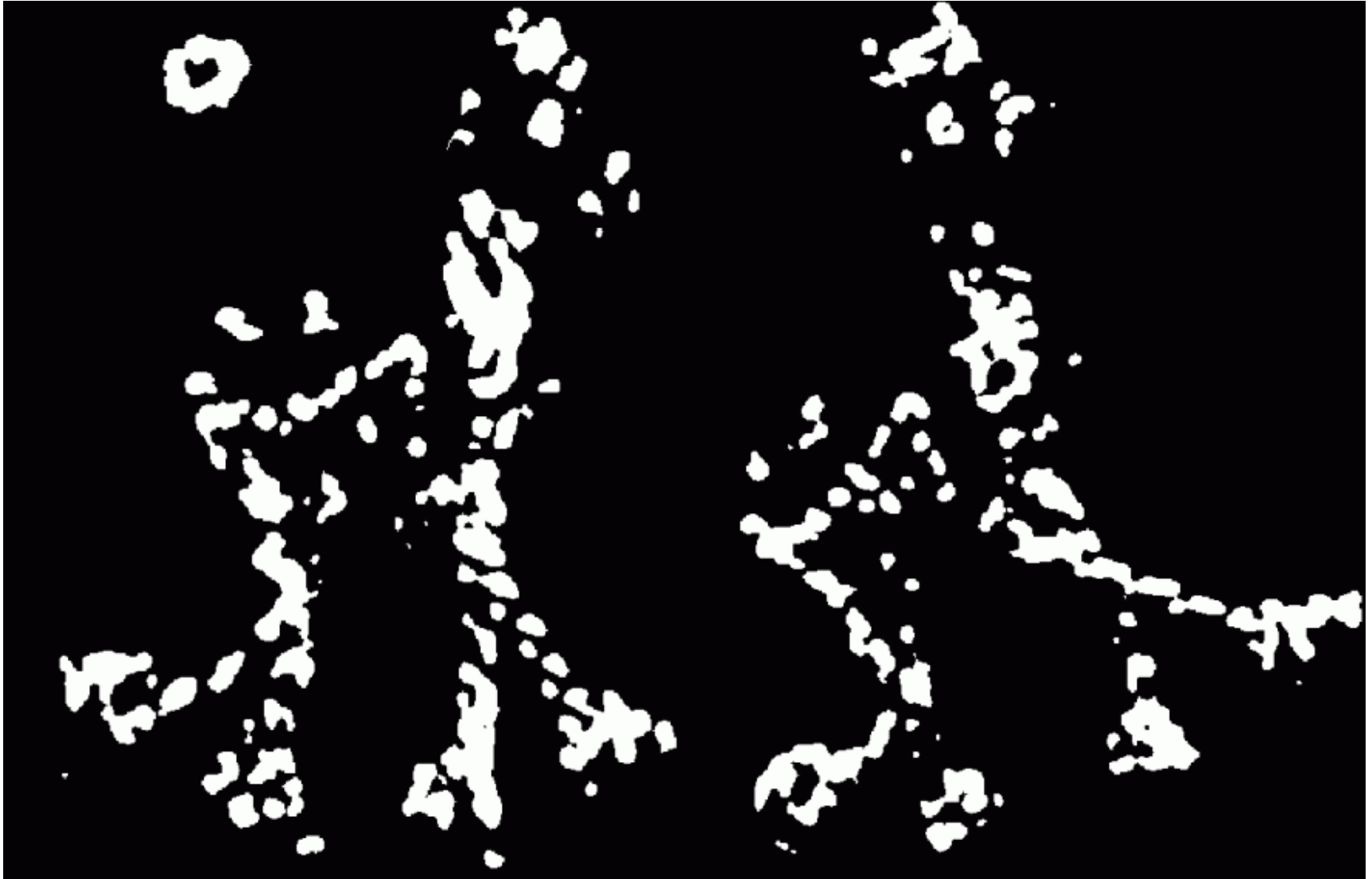
Harris sarokdetektáló algoritmus működése

R sarkossági jellemző értékei



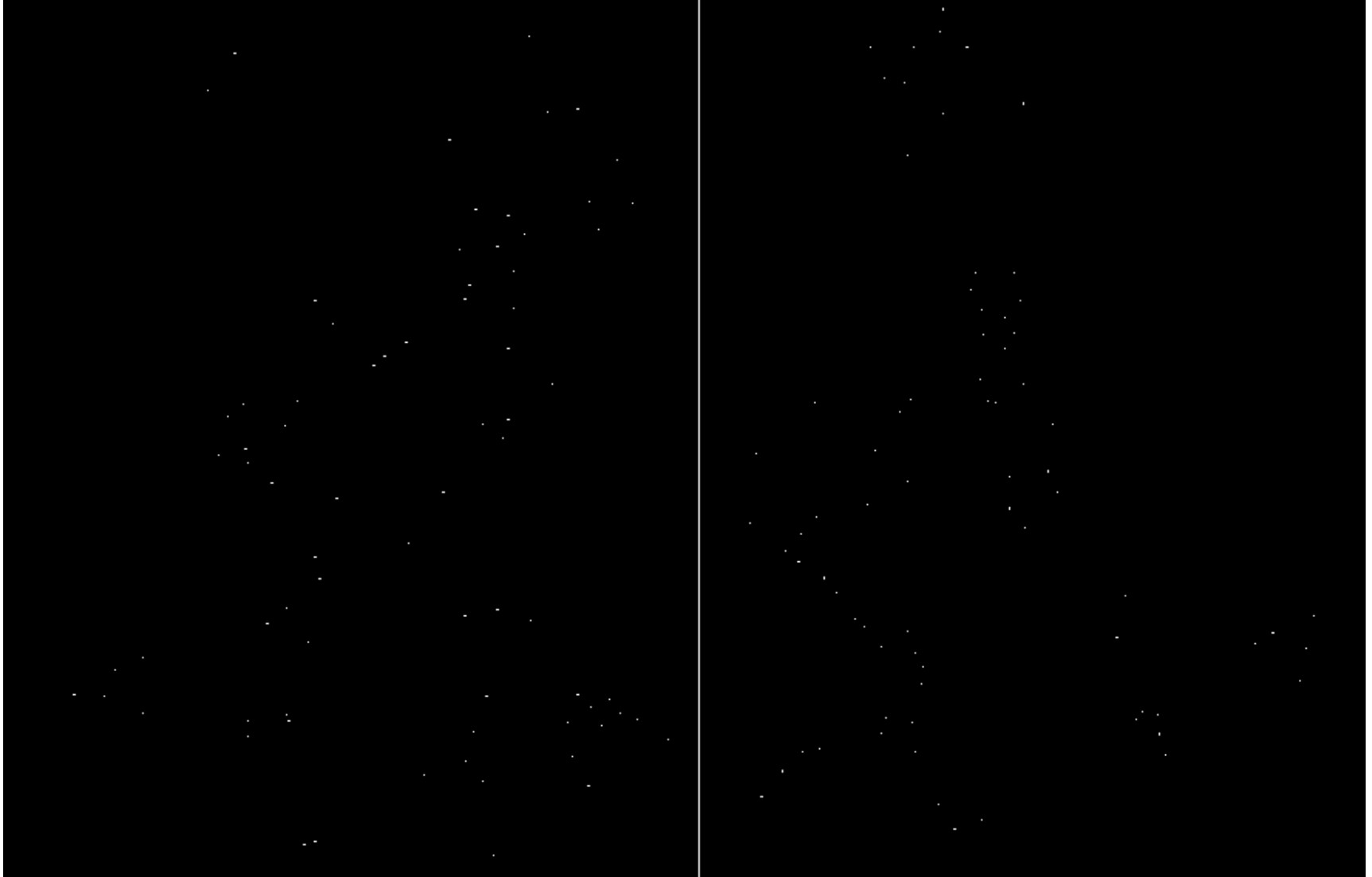
Harris sarokdetektáló algoritmus működése

küszöbölés eredménye: $R > küszöb$



Harris sarokdetektáló algoritmus működése

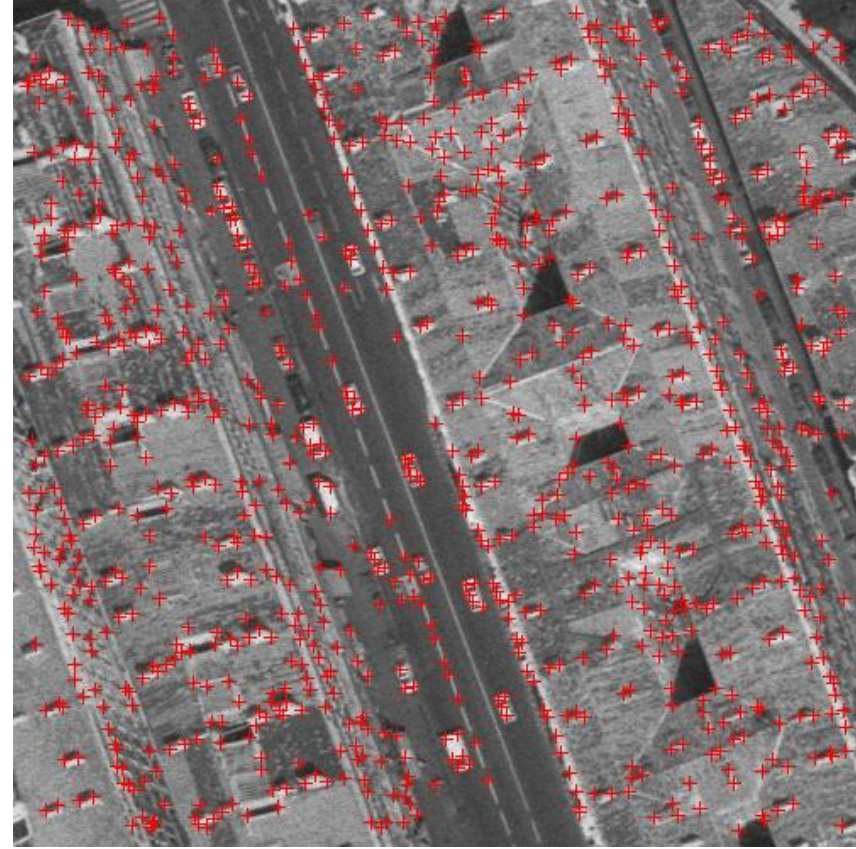
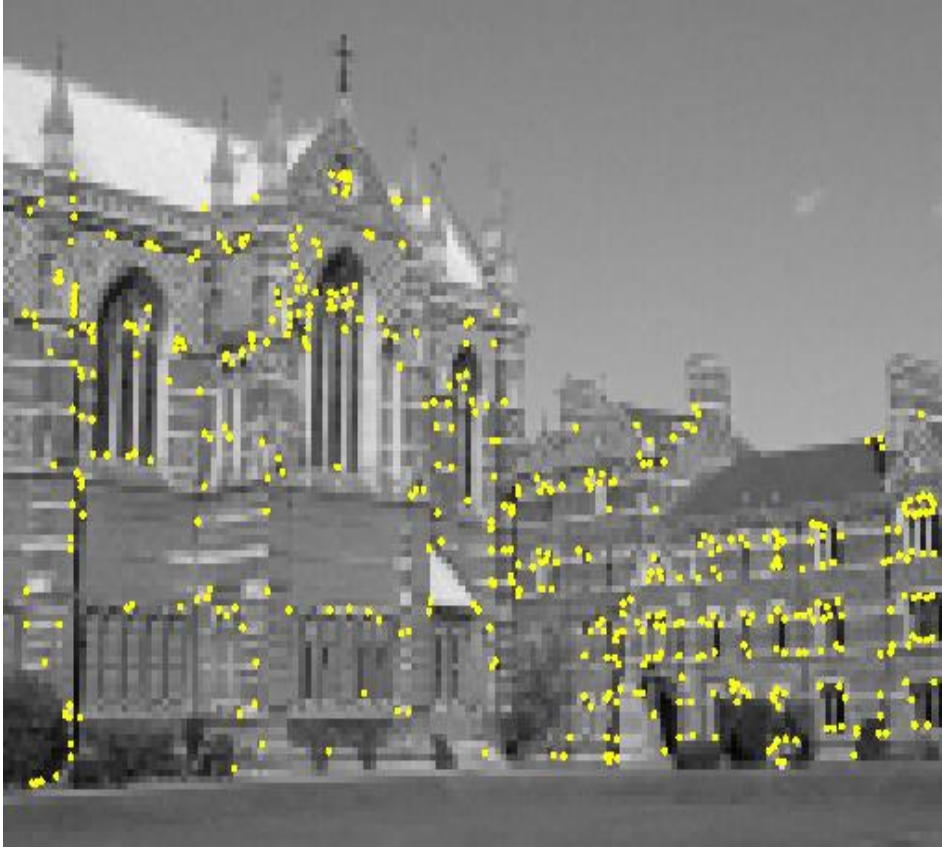
Lokális maximumok meghagyása



Harris sarokdetektáló algoritmus működése



Harris sarokdetektor



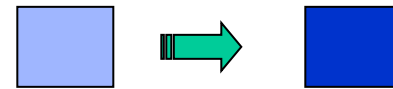
Invariancia

- A sarokpontok detekciójának *invariánsnak* kell lennie a fotometriai és geometriai transzformációkra.
- Vagyis egy sarokpontot akkor is megbízhatóan kell detektálnunk, ha
 - a két kép fotometriai tulajdonságai különböznek (pl. más a megvilágítás)
 - vagy a két kép között geometriai eltérés van (pl. forgatás, skálázás)



Harris: Fotometriai részleges invariancia

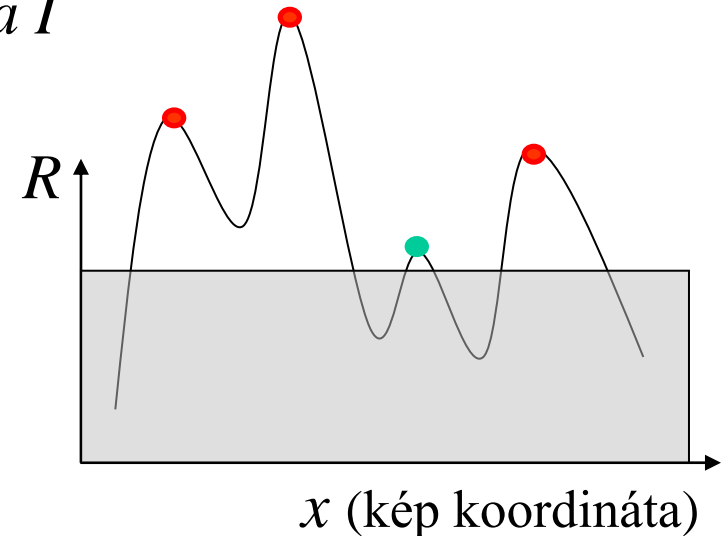
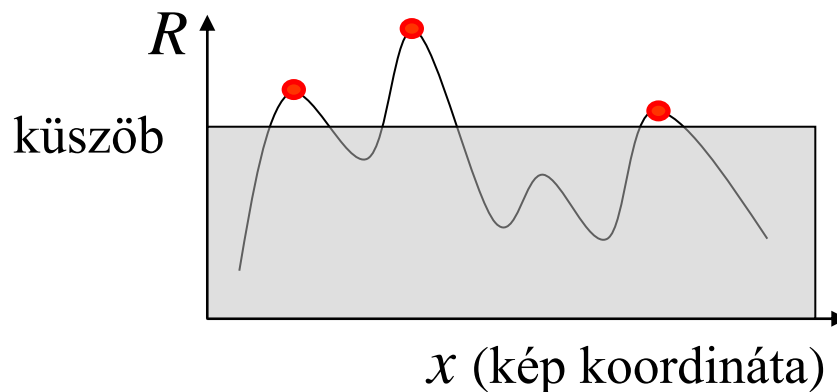
- Mivel csak deriváltakat használunk, ezért az intenzitáseltolásra invariáns lesz



$$I \rightarrow I + b$$

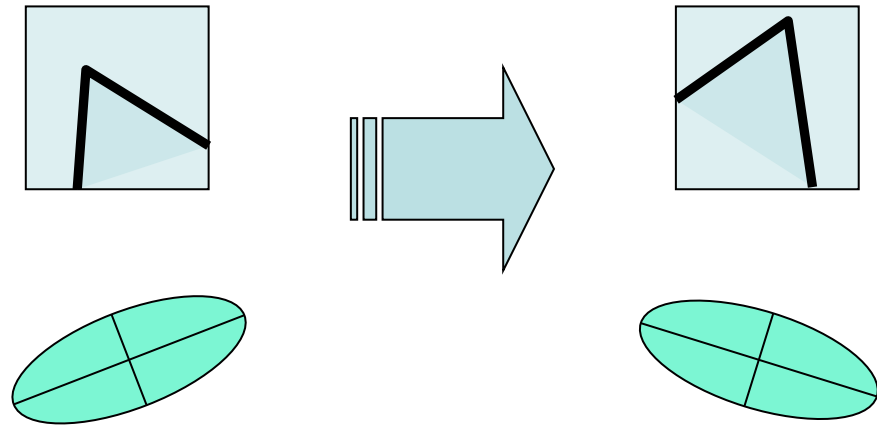
- A skálázásra már csak részlegesen invariáns (küszöbértéktől is függ)

$$I \rightarrow a I$$

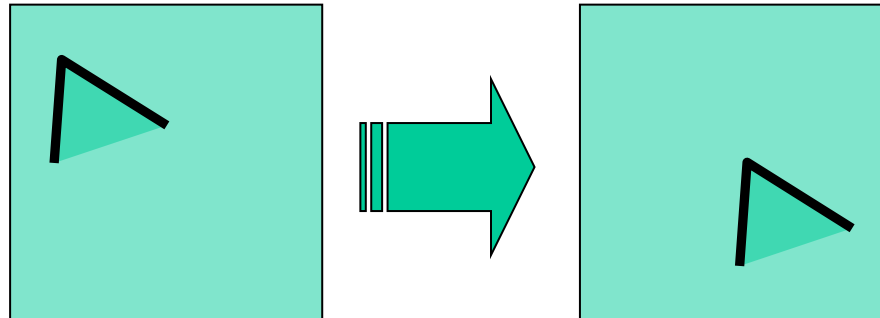


Harris: eltolásra és forgatásra invariáns

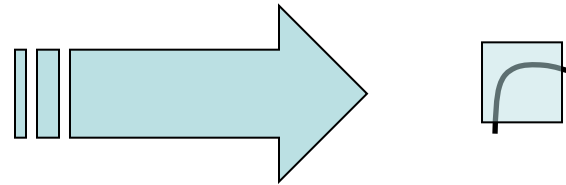
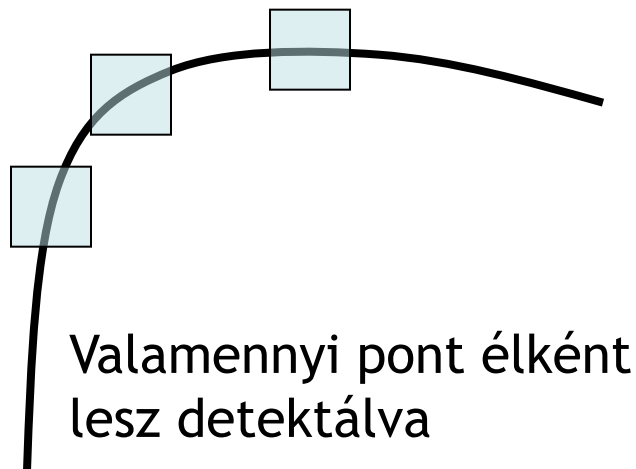
- Az ellipszisek elfordulnak, de az alakjuk (vagyis a sajátértékük) nem változik.
 - ➔ Az R sarkosság-jellemző invariáns a kép elforgatására



- A deriválás és az ablakozás invariáns az eltolásra
 - ➔ Az R sarkosság-jellemző invariáns a kép eltolására



Harris: skálázásra nem invariáns

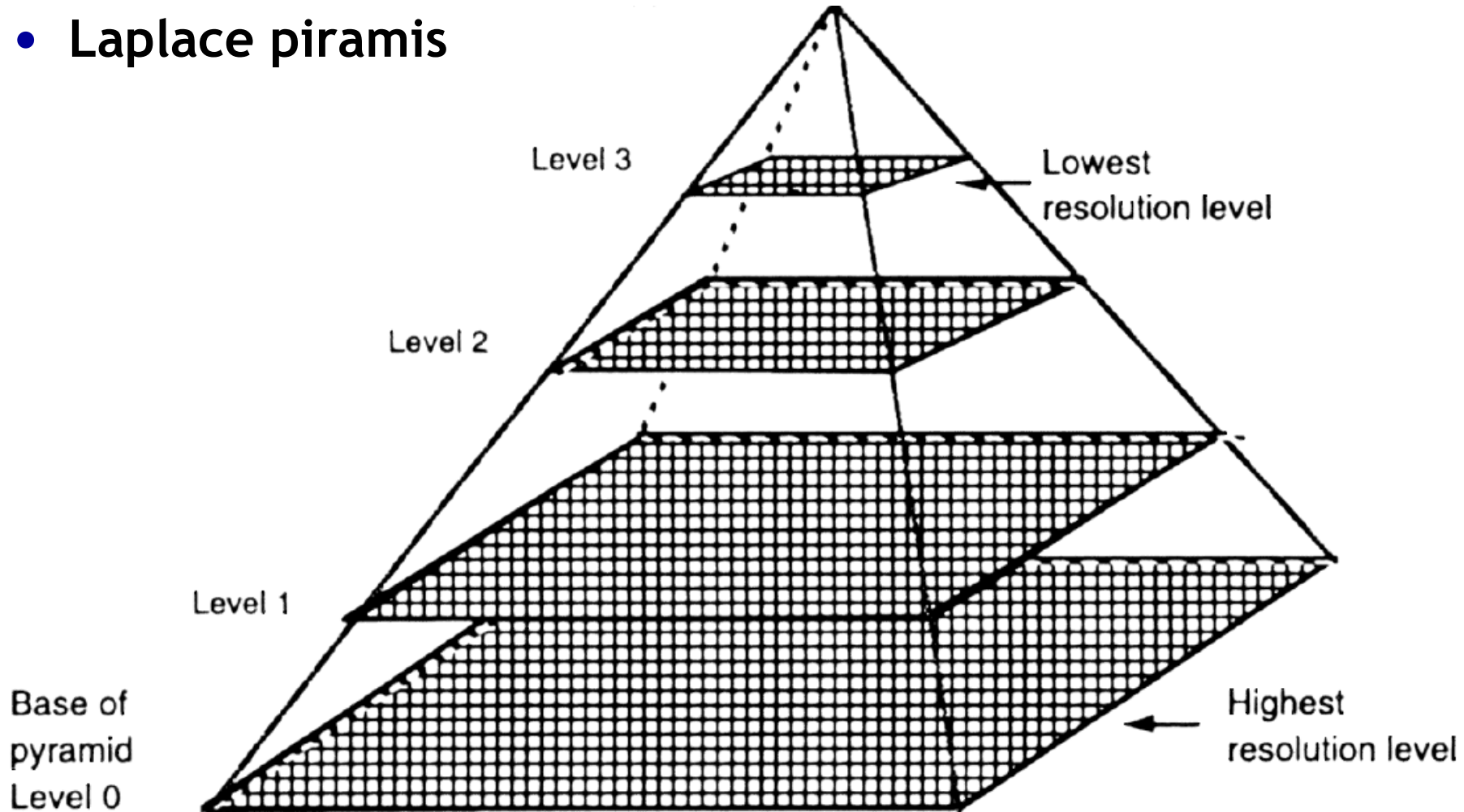


Sarokpont !

- **Hogyan detektálhatunk skála-invariáns, egymásnak megfeleltethető sarokpontokat?**

Többfelbontású képpiramis

- Gauss piramis
- Laplace piramis



Gauss piramis

- Kép kicsinyítése = alul mintavételezés → simítás szükséges!
- Gauss szűrővel simítva Gauss piramist kapunk
 - Vegyük észre, hogy $\text{gauss} * \text{gauss} = \text{újabb gauss}$
- Mivel a Gauss szűrő aluláteresztő, ezért a reprezentáció redundáns.



512

256

128

64

32

16

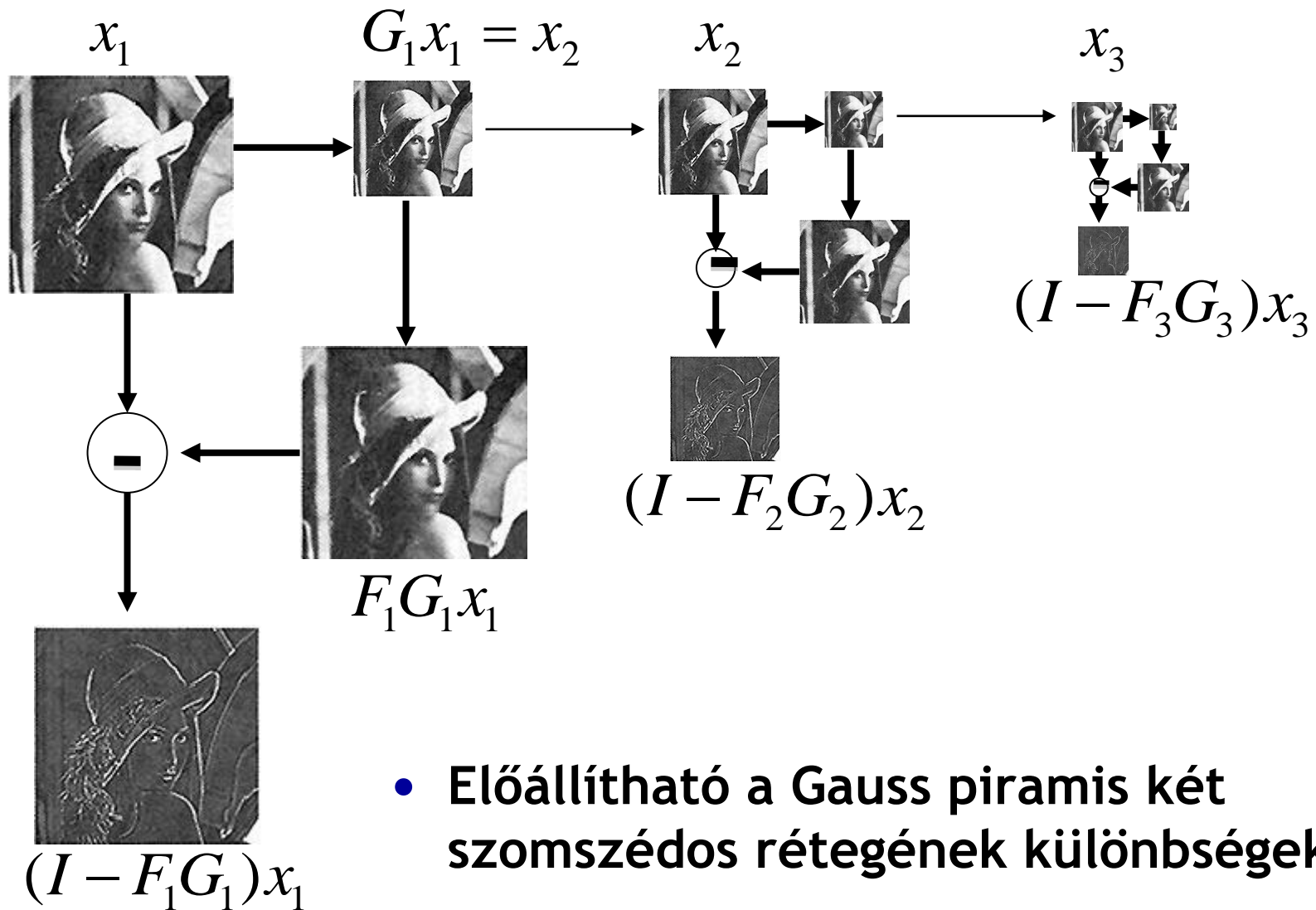
8



Gauss piramis

Aluláteresztő szűrő:
minden szinten
jelen vannak az
alacsony
frekvenciák (vagyis
redundáns)

Laplace piramis



- Előállítható a Gauss piramis két szomszédos rétegének különbségeként



512

256

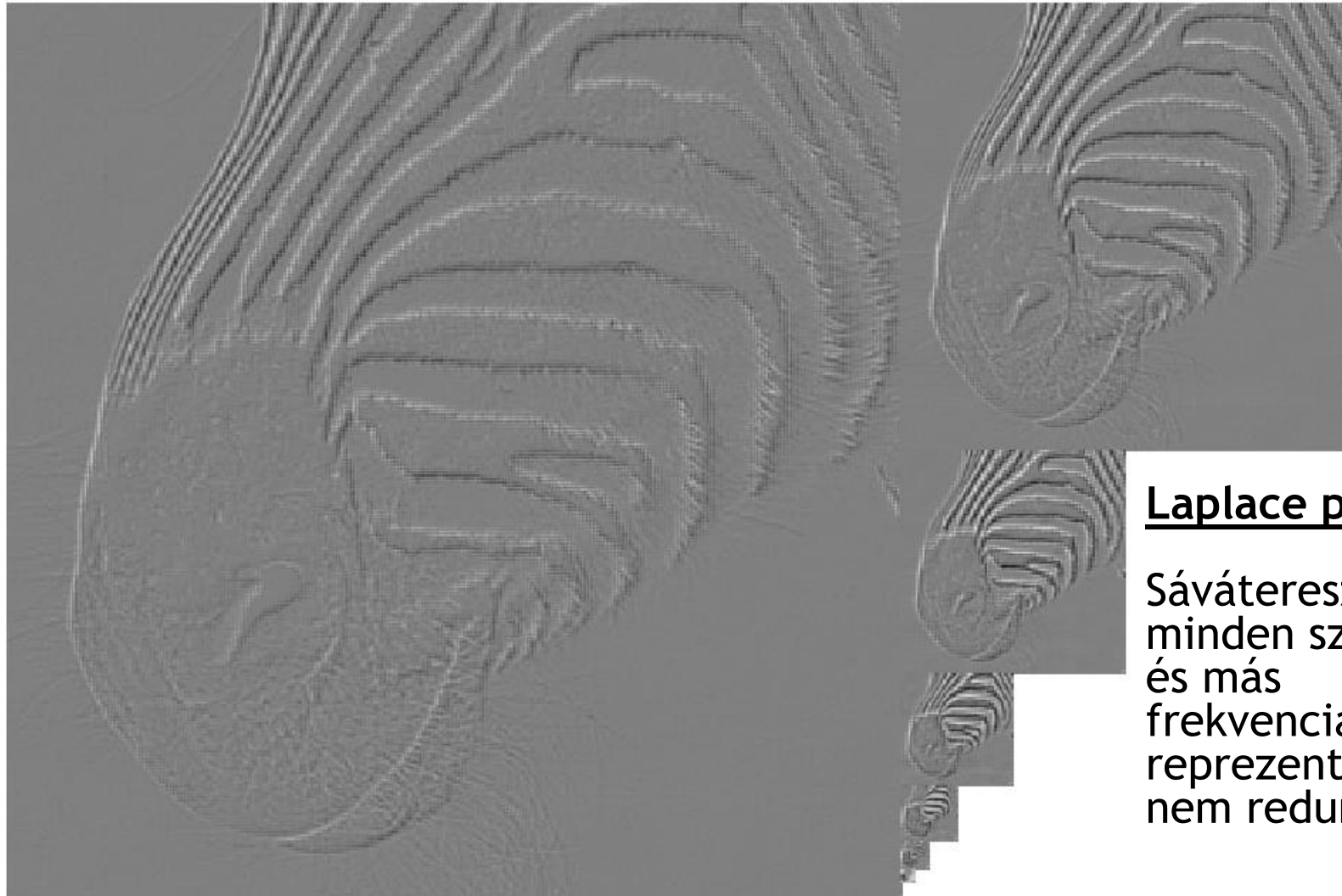
128

64

32

16

8

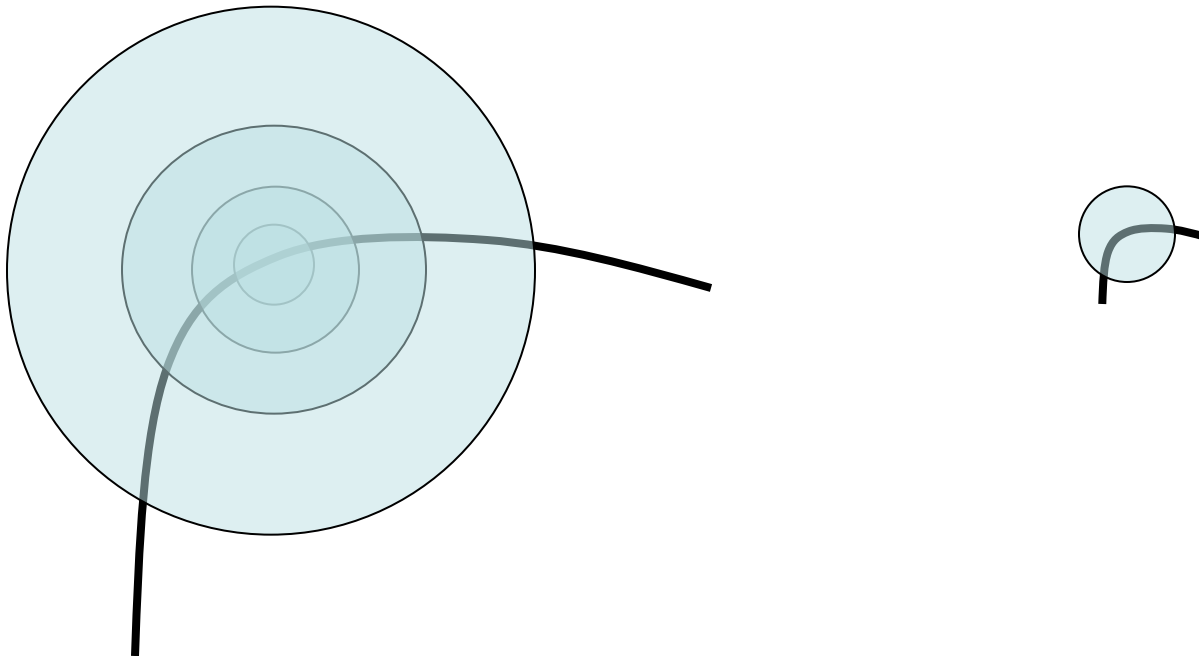


Laplace piramis

Sáváteresztő szűrő:
minden szinten más
és más
frekvenciákat
reprezentál (vagyis
nem redundáns)

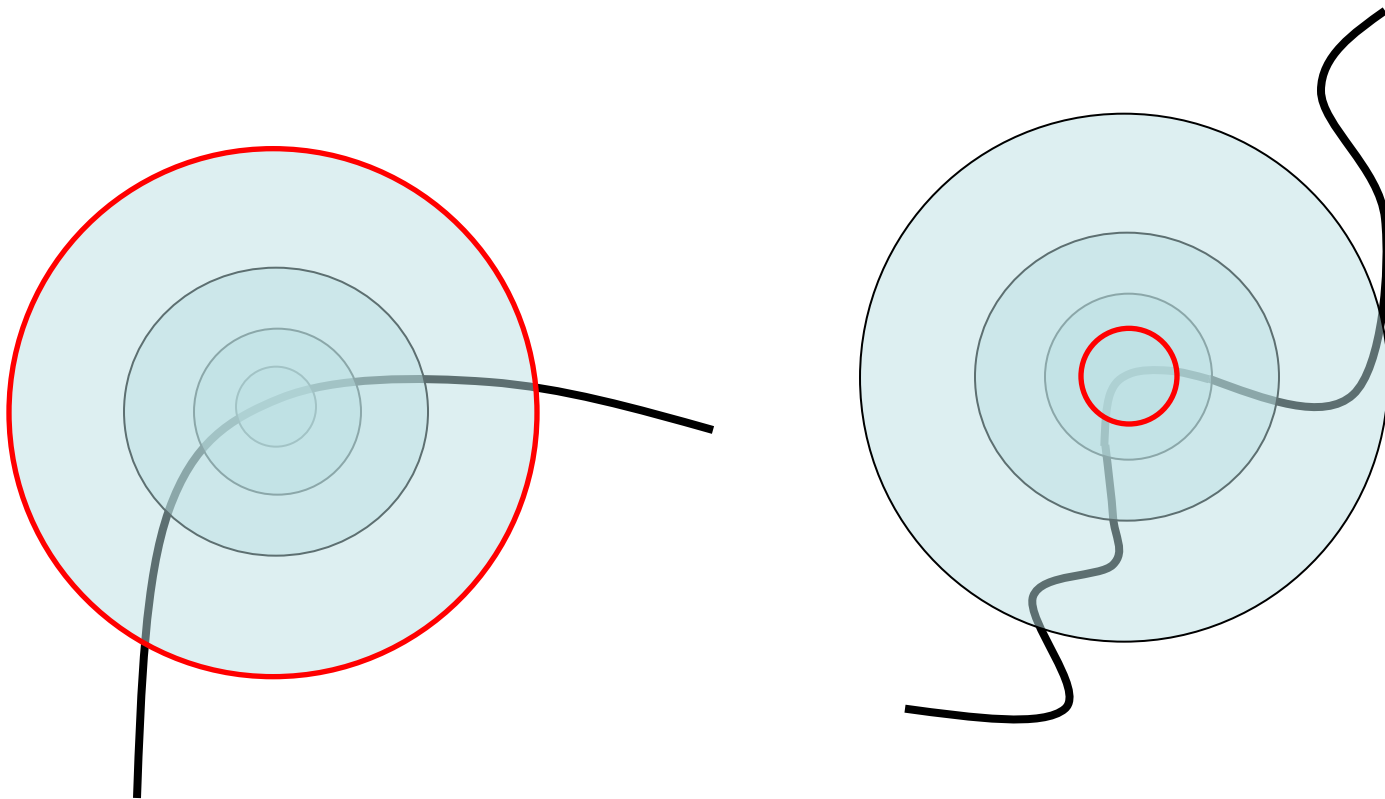
Skála-invaráns detektálás

- Tekintsünk különböző méretű régiókat (pl. köröket) egy pont körül
 - Ezzel ekvivalens, ha megfelelő képpiramis változó felbontású szintjein azonos mérettel keresünk
- Az egymásnak megfelelő méretek hasonlóan néznek ki mindkét képen



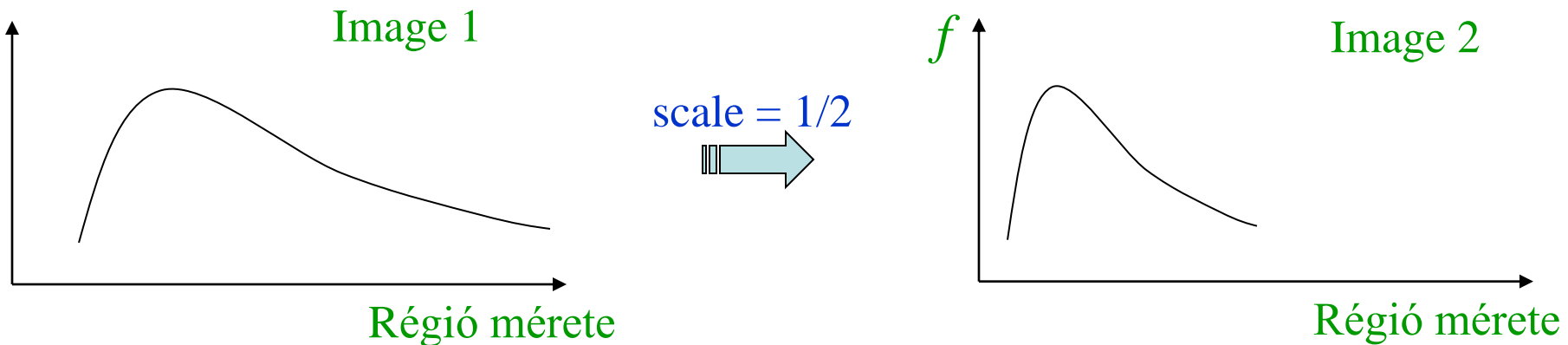
Skála-invaráns detektálás

- Hogyan válasszuk ki egymástól függetlenül a “megfelelő” köröket?



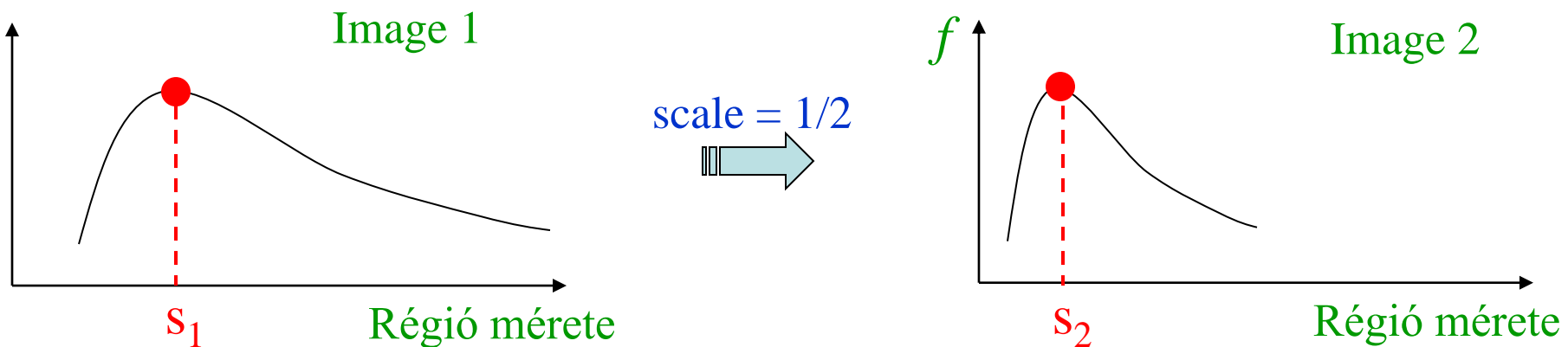
Automatikus skála meghatározás

- Konstruáljunk egy skála-invariáns függvényt
 - Ugyanazt az értéket kell adnia az egymásnak megfelelő régiókon akkor is, ha azok különböző felbontásúak
 - Például átlagintenzitás: állandó marad akkor is, ha a felbontás változik
- Egy adott képpont esetén ezt tekinthetjük az ablakméret függvényének



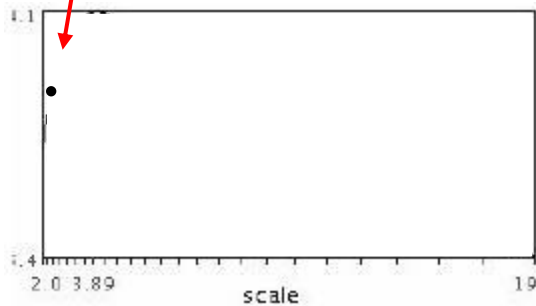
Automatikus skála meghatározás

- Tekintsük a függvény lokális maximumát
- Vegyük észre, hogy ez független lesz a mérettől!
- **FONTOS:** ez a skála-invariáns méret a két képen egymástól függetlenül megtalálható!

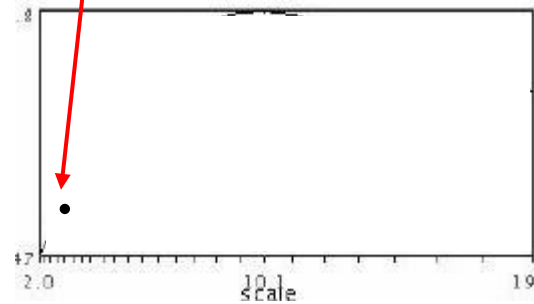


Automatikus skála meghatározás

- Függvény érték növekvő skála esetén (scale signature)



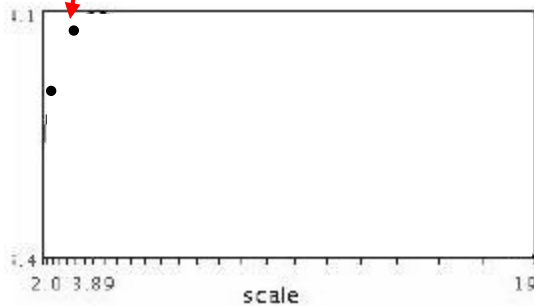
$$f(I_{i_1 \dots i_m}(x, \sigma))$$



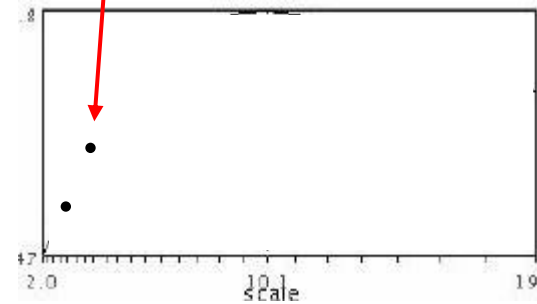
$$f(I_{i_1 \dots i_m}(x', \sigma))$$

Automatikus skála meghatározás

- Függvény érték növekvő skála esetén (scale signature)



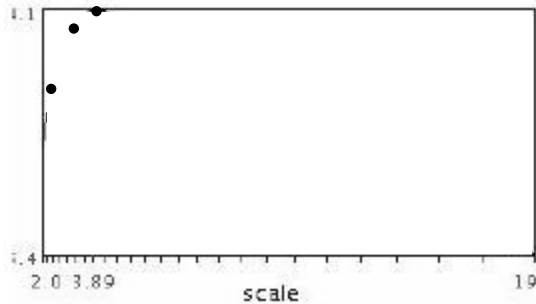
$$f(I_{i_1 \dots i_m}(x, \sigma))$$



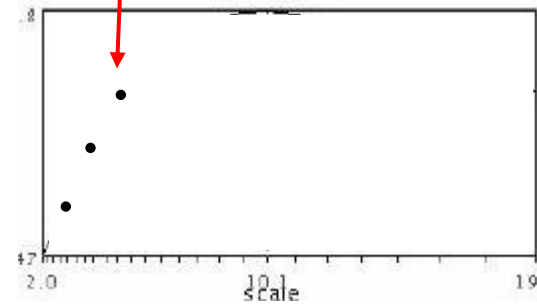
$$f(I_{i_1 \dots i_m}(x', \sigma))$$

Automatikus skála meghatározás

- Függvény érték növekvő skála esetén (scale signature)



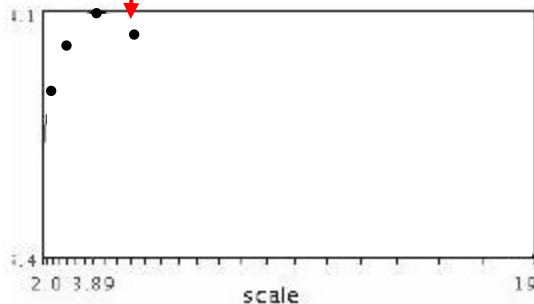
$$f(I_{i_1...i_m}(x, \sigma))$$



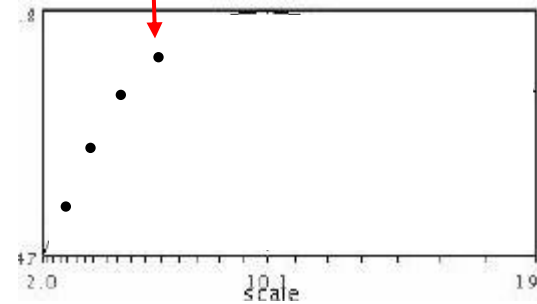
$$f(I_{i_1...i_m}(x', \sigma))$$

Automatikus skála meghatározás

- Függvény érték növekvő skála esetén (scale signature)



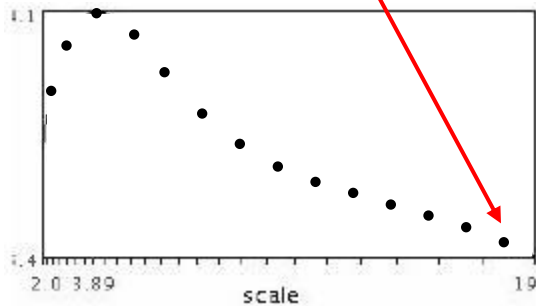
$$f(I_{i_1...i_m}(x, \sigma))$$



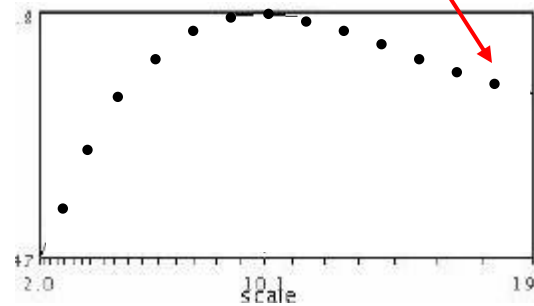
$$f(I_{i_1...i_m}(x', \sigma))$$

Automatikus skála meghatározás

- Függvény érték növekvő skála esetén (scale signature)



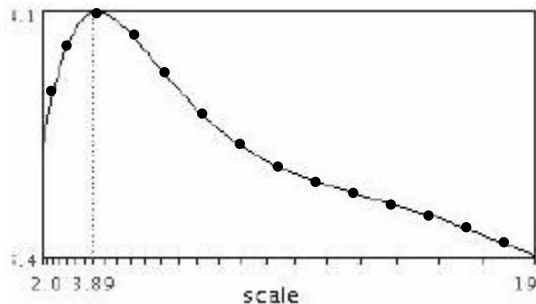
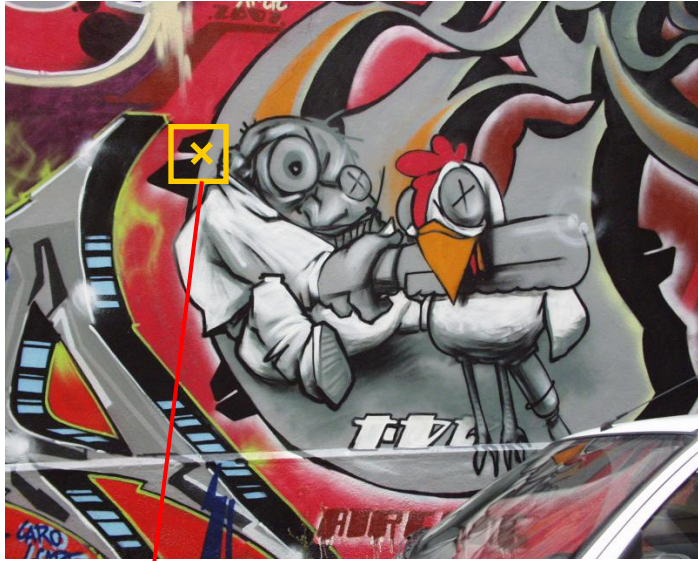
$$f(I_{i_1 \dots i_m}(x, \sigma))$$



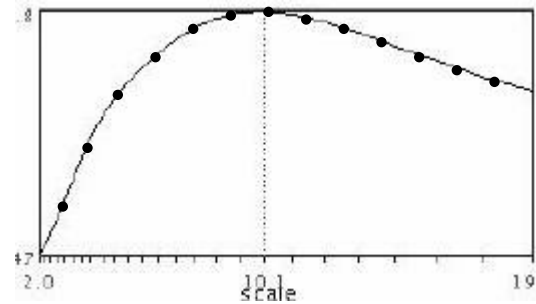
$$f(I_{i_1 \dots i_m}(x', \sigma))$$

Automatikus skála meghatározás

- Függvény érték növekvő skála esetén (scale signature)



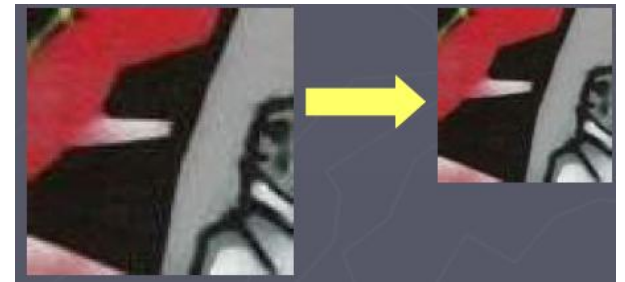
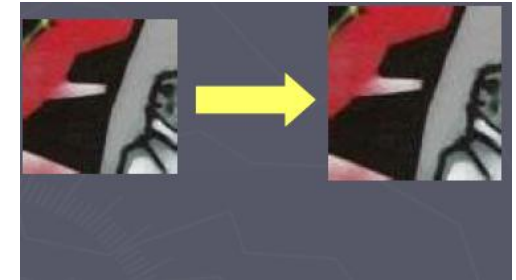
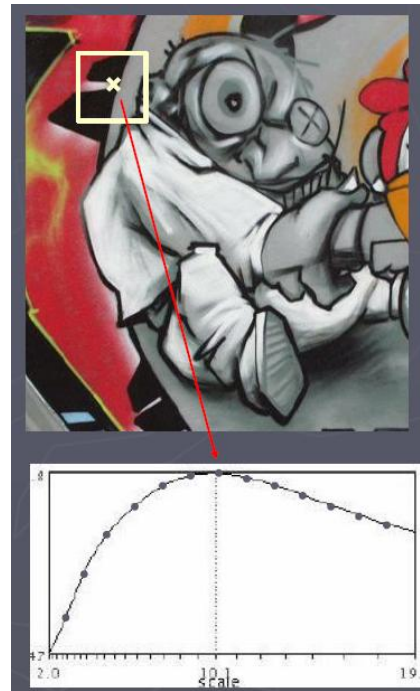
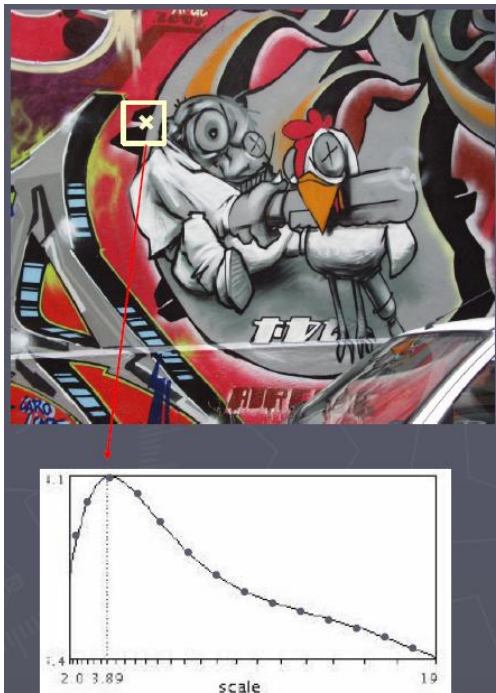
$$f(I_{i_1...i_m}(x, \sigma))$$



$$f(I_{i_1...i_m}(x', \sigma'))$$

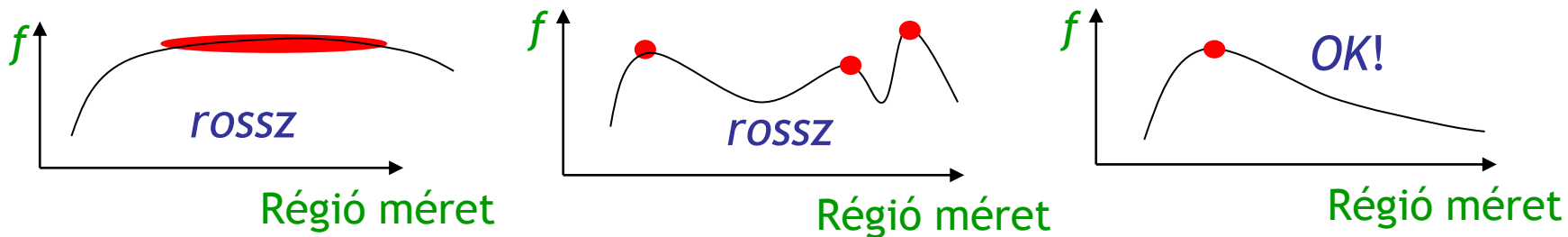
Automatikus skála meghatározás

- Használjuk a detektor által meghatározott skálát a megfeleltetéshez használt leírók kiszámításához a normalizált képtérben



Skála invariáns detektálás

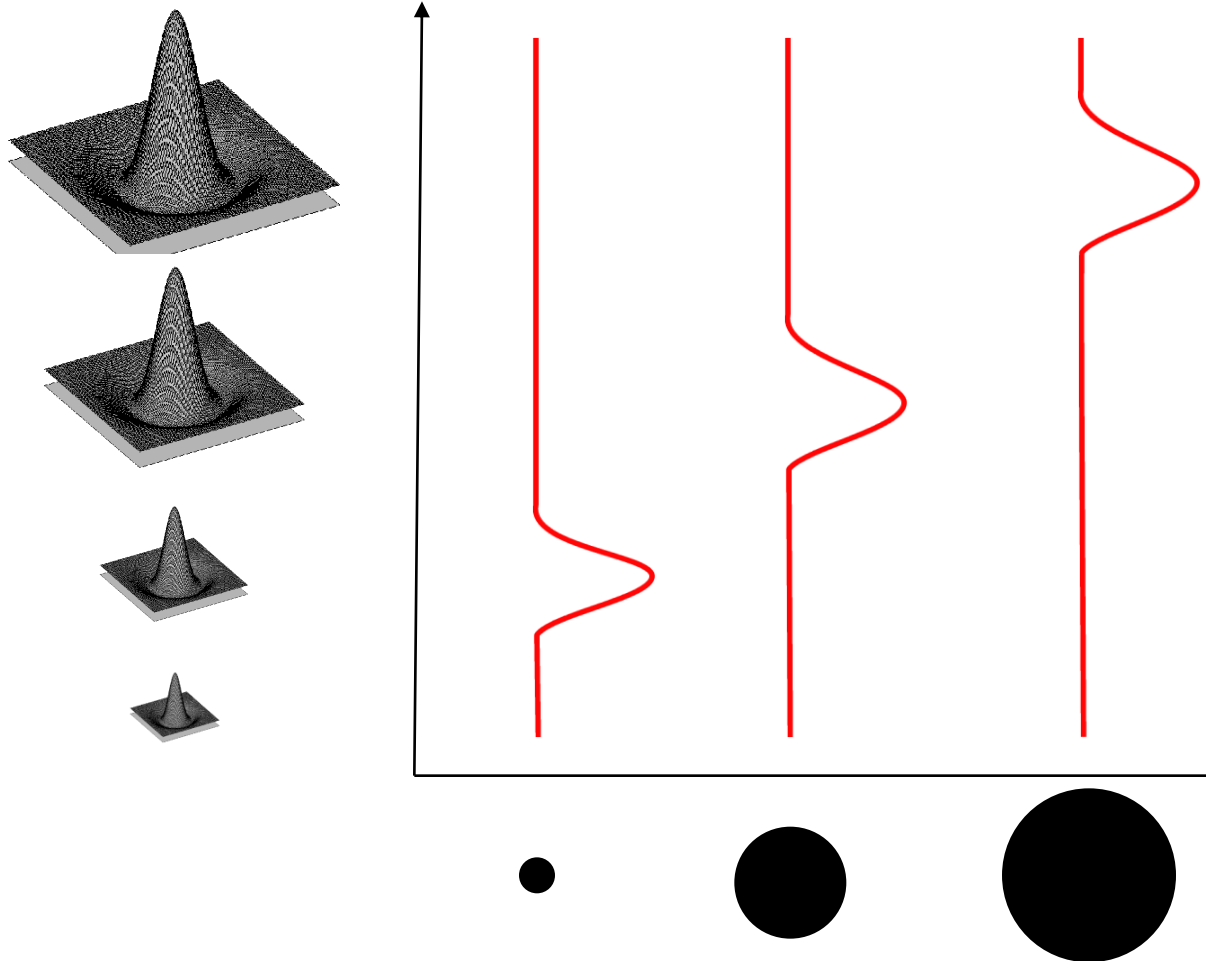
- A skála-detektáláshoz alkalmas függvénynek egyetlen jól meghatározott csúcsa van



- Ilyen például egy olyan függvény, ami a kontrasztra reagál (vagyis a lokális intenzitásváltozásokra)

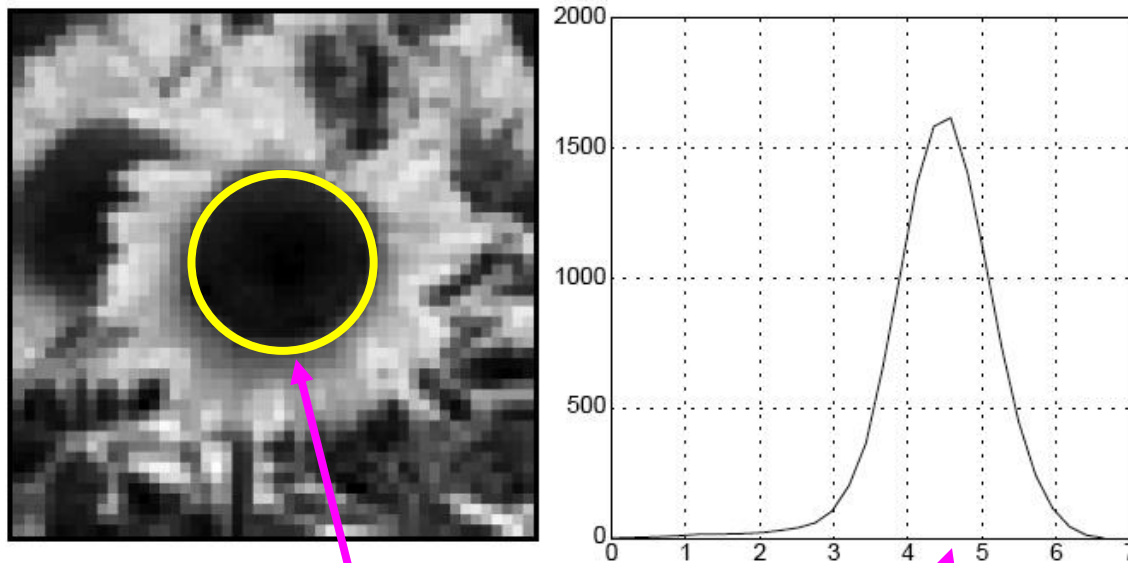
“Signature Function” megválasztása

- Laplacian-of-Gaussian = “blob” detector



Karakterisztikus skála

- *Karakterisztikus skálának* nevezzük azt a skálát, amelyen a Laplace válasz csúcsértéket produkál



Karakterisztikus skála

T. Lindeberg (1998). ["Feature detection with automatic scale selection."](#)
International Journal of Computer Vision 30 (2): pp 77--116.

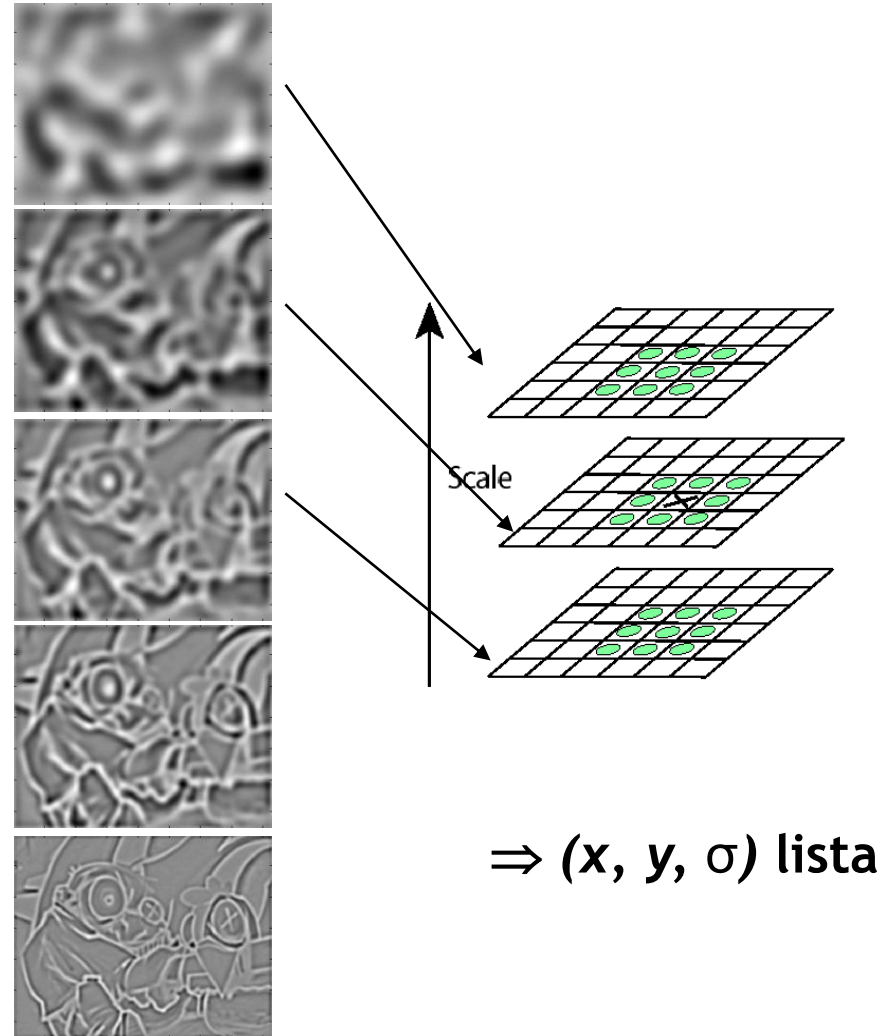
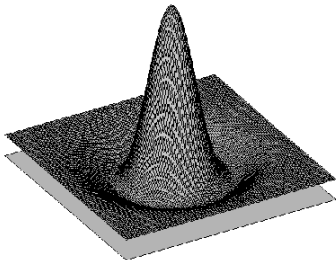
Laplacian-of-Gaussian (LoG) skála-tér

- Jellemző pontok:
A Laplacian-of-Gaussian (LoG) skála-tér lokális szélsőhelyei



$$L_{xx}(\sigma) + L_{yy}(\sigma) \rightarrow \sigma^3$$

σ^5
 σ^4
 σ^3
 σ^2
 σ



Skála-tér blob detektor példa

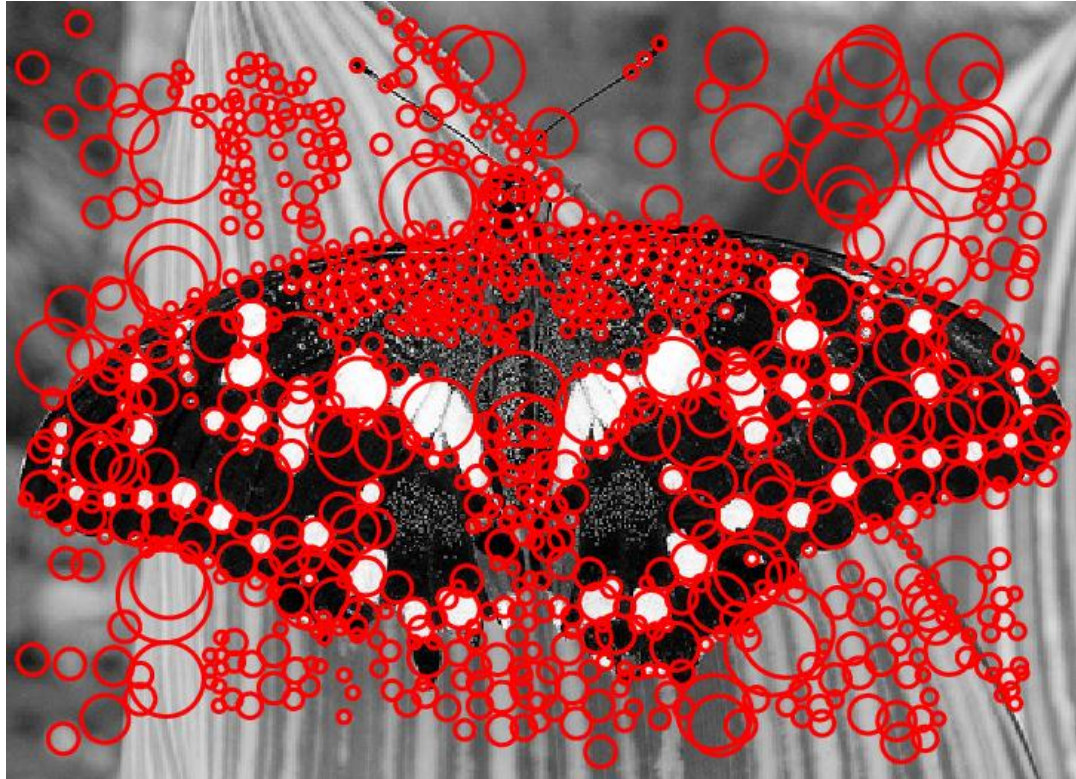


Skála-tér blob detektor példa



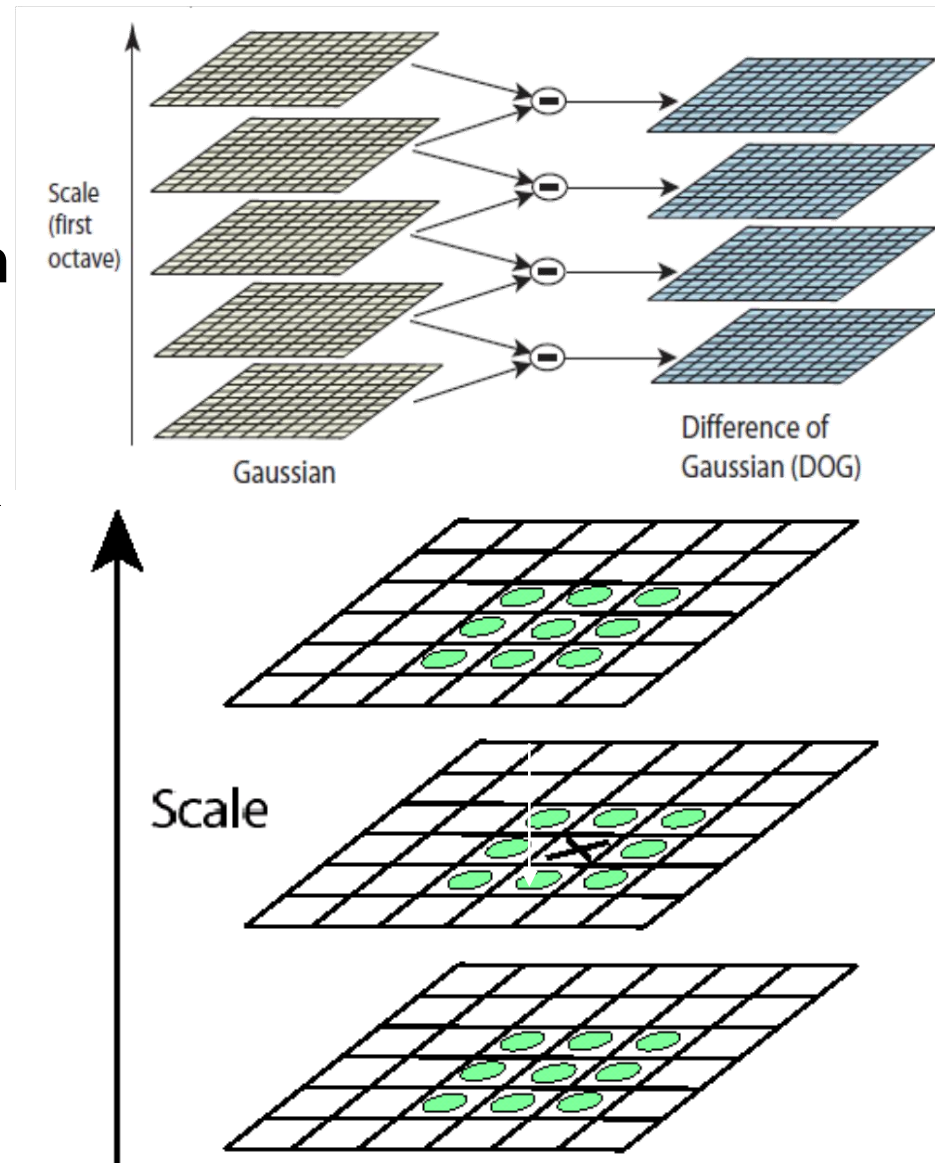
sigma = 11.9912

Skála-tér blob detektor példa

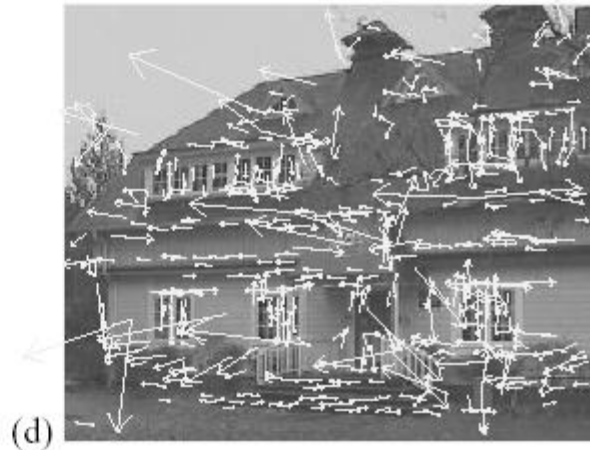
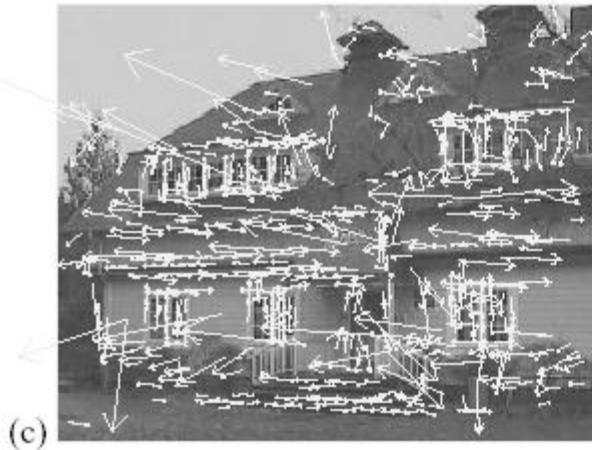


Jellemző pontok kinyerése DoG segítségével

- Detektáljuk a difference-of-Gaussian (DoG) szélsőhelyeit skála-térben
 - (max, min) egy $3 \times 3 \times 3$ szomszédságban
- Küszöböljük az értékeket
- Elimináljuk az él-válaszokat
- Az így kapott pontok listája: (x, y, σ)



Példa jellemző pontok kinyerésére



- (a) 233x189 kép
- (b) 832 DOG maximum
- (c) 729 maradt küszöbölés után
- (d) 536 maradt a sarkosság ellenőrzése után (élvázlatok törlése)

Skála invariáns sarok-detektálás

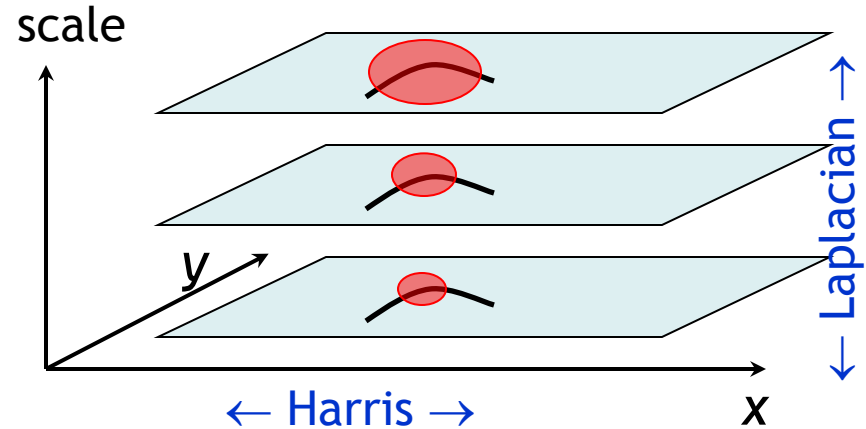
- **Input:** két kép ugyanazon látványról nagy felbontásbeli különbséggel
- **Cél:** detektáljuk *ugynazon* jellemző pontokat mindkét képen *függetlenül* a képmérettől
- **Megoldás:** keressük a *maximumát* egy megfelelően konstruált függvénynek a *skála-* és *képtérben*

Skála invariáns sarok-detektálás

- **Harris-Laplacian**¹

Keressük meg a lokális maximumait:

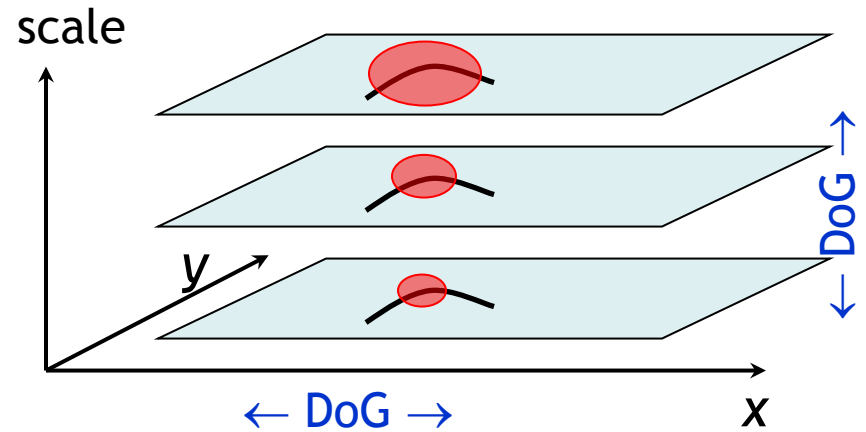
- Harris sarokdetektor a képen
- Laplacian a skálatérben



- **SIFT (Lowe)**²

Keressük meg a lokális maximumait:

- Difference of Gaussians (DoG) a skála- és képtérben



¹ K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

² D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". IJCV 2004

Felhasznált anyagok

- **Palágyi Kálmán: Digitális Képfeldolgozás**
/pub/Digitalis_kepfeldolgozas
- **Trevor Darrell: C280, Computer Vision**
<http://www.cs.cmu.edu/afs/cs/academic/class/15385-s06/lectures/ppts/>
- **Srinivasa Narasimhan: Computer Vision**
<http://www.cs.cmu.edu/afs/cs/academic/class/15385-s12/www/>
- **James Hays: CS 143 Computer Vision, Brown University**
 - <http://www.cs.brown.edu/courses/cs143/>
- **További források az egyes diákon megjelölve**