

6. Modell illesztés, alakzatok

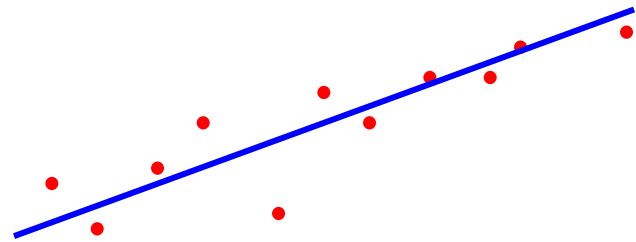
Kató Zoltán

**Képfeldolgozás és Számítógépes Grafika tanszék
SZTE**

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

ROBOSZTUS EGYENES ILLESZTÉS

Egyenes illesztés



- Adott a síkban egy P_1, P_2, \dots, P_N ponthalmaz
- Illesszünk rájuk egy egyenest úgy, hogy
 - A pontok egyenestől vett négyzetes távolságának összege minimális (*geometriai hiba*)
 - Az egyenest az egyenletével reprezentáljuk:

$$c + n_x x + n_y y = 0, \quad n_x^2 + n_y^2 = 1$$

- (n_x, n_y) az egyenes normálvektora (egységvektor)
- Ha egy P pont nincs az egyenesen, akkor $|r|$ a pont távolsága az egyenestől:

$$c + n_x x + n_y y = r$$

- ➔ Túlhatározott egyenletrendszer megoldása (legkisebb négyzetes probléma):

- Minimalizáljuk:

$$\|r\| = \sum_{i=1}^N r_i^2$$

$$\underbrace{\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & y_N \end{bmatrix}}_A \underbrace{\begin{bmatrix} c \\ n_x \\ n_y \end{bmatrix}}_x = \underbrace{\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}}_r \quad \text{és} \quad n_x^2 + n_y^2 = 1$$

Megoldás

- **N** nagy \rightarrow **A** sok sorból áll
 - **A QR** felbontásával (Q ortogonális, R felső trianguláris) egy kisebb egyenletrendszert kaphatunk
 - Mivel a normát az ortogonális transzformáció nem változtatja (hiszen csak elforgatás: $||Q^T r|| = ||r||$), az alábbi egyenletrendszert kapjuk:

$$\mathbf{A} = \mathbf{QR} \Rightarrow \mathbf{Q}^T \mathbf{Ax} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c \\ n_x \\ n_y \end{bmatrix} = \mathbf{Q}^T \mathbf{r} \quad \text{és} \quad n_x^2 + n_y^2 = 1$$

Megoldás

- Mivel a nemlineáris mellékfeltétel csak 2 ismeretlent tartalmaz, az alábbi feladatot kell megoldani:

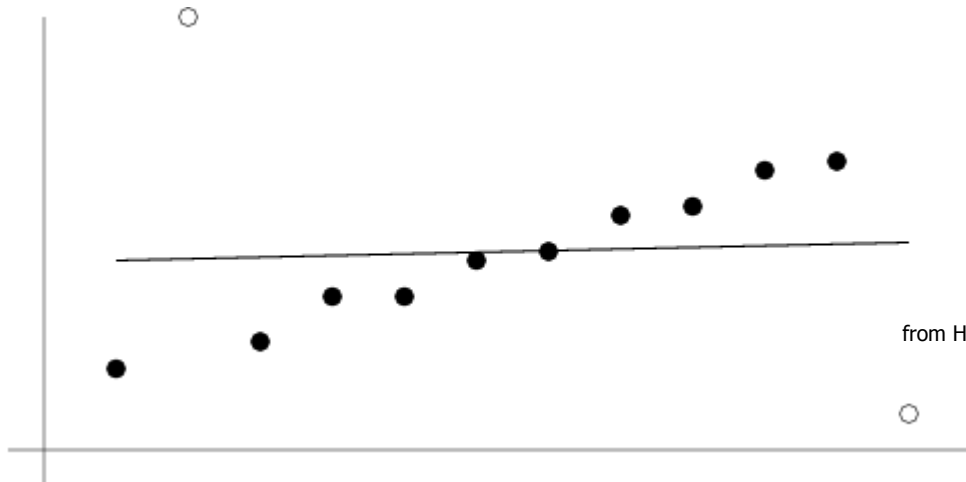
$$\begin{bmatrix} R_{22} & R_{23} \\ 0 & R_{33} \end{bmatrix} \begin{bmatrix} n_x \\ n_y \end{bmatrix} \approx 0 \quad \text{és} \quad \mathbf{n}_x^2 + \mathbf{n}_y^2 = 1$$

- Ez egy klasszikus LSE probléma: $\|Bx\| = \min$, és $\|x\| = 1$.
 - A minimum értéke **B legkisebb szinguláris értéke**
 - A megoldás pedig az ennek *megfelelő szinguláris vektor*
 - ➔ (n_x, n_y) B SVD felbontásából kapható meg
 - ➔ **c** pedig ezek után az eredeti egyenletekbe való behelyettesítéssel adódik

Kilógó adatok problémája (Outliers)

- Az LSE módszer feltételezi, hogy az adatpontjaink egy Gauss hiba erejéig jól illeszkednek a modellre
 - Vagyis a pont koordináták nagyjából egy egyenesre esnek egy korlátos szórással.
- Mit tehetünk, ha az adatpontokra nem teljesül ez a feltétel?
- *Kilógó adatok (outliers)* – az adatpontok nem ugyanahhoz a modellhez (egyeneshez) tartoznak: a becslésünk hibás lesz

Line fitting using regression is biased by outliers

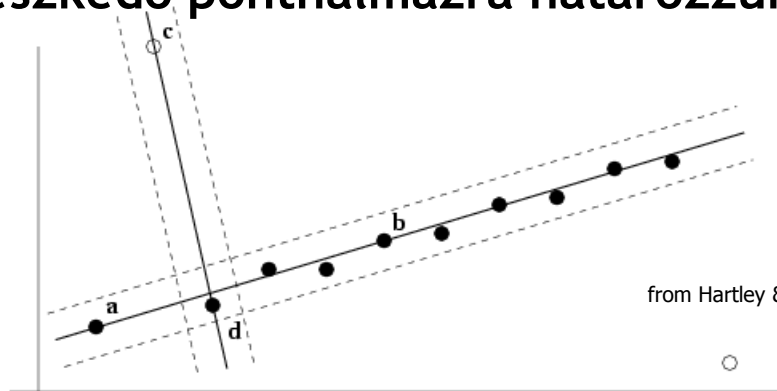


from Hartley & Zisserman

Robosztus becslés: RANSAC

- A becslést tekintjük egy két-lépéses eljárásnak:
 - Osztályozzuk az adatpontjainkat kilógókra és illeszkedőkre
 - Illesszük a modellt az illeszkedőkre.
- 1. Vegyünk egy véletlen minimális mintát az adatpontokból, amire a modell illeszthető (a *minta*)
- 2. Az illesztett modelltől t távolságra levő ponthalmaz a *konszenzus halmaz*, melynek mérete a modell *támogatottsága*.
- 3. Ismételjük az eljárást N mintára → a legnagyobb támogatottságú modell lesz a robusztus illesztés
 - Ettől t távolságon kívül lévő pontok a kilógók
 - A végső modellt az illeszkedő ponthalmazra határozzuk meg

Two samples
and their supports
for line-fitting



from Hartley & Zisserman

RANSAC: t és N meghatározása

- t : általában empirikusan határozzuk meg
- N : határozzuk meg úgy, hogy p valószínűséggel legalább egy minta tiszta illeszkedő (S pontot tartalmazó) adatpontokból álljon:

$$N = \frac{\log(1 - p)}{\log(1 - (1 - \epsilon)^S)}$$

ahol ϵ annak a valószínűsége, hogy egy pont kilógó

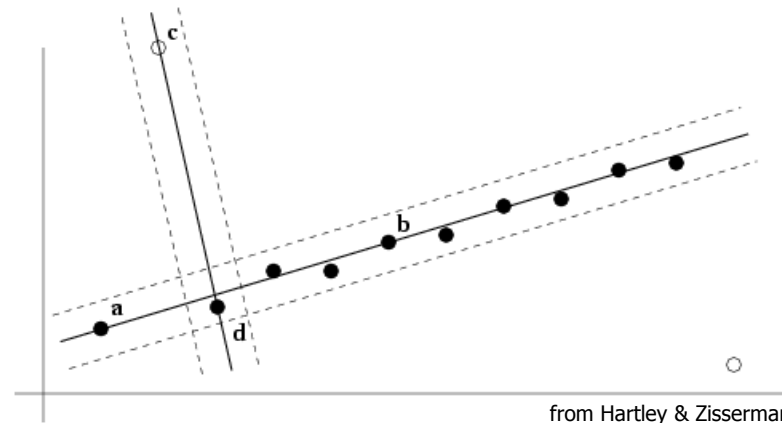
- Tipikusan $p = 0.99$

RANSAC: példa ($p = 0.99$)

Sample size	Proportion of outliers ε						
	5%	10%	20%	25%	30%	40%	50%
S							
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293
7	4	8	20	33	54	163	588
8	5	9	26	44	78	272	1177

Példa: robusztus egyenes illesztés

- $n = 12$ pont
- Minimális minta mérete $S = 2$
- 2 kilógó pont $\rightarrow \varepsilon = 1/6 = 20\%$
- Tehát $N = 5$ 99% valószínűséggel már ad legalább egy mintát, ami nem tartalmaz kilógó pontokat.
 - Ezzel szemben az összes lehetséges pontpár végigpróbálgatása $N = 66$ esetet jelentene.

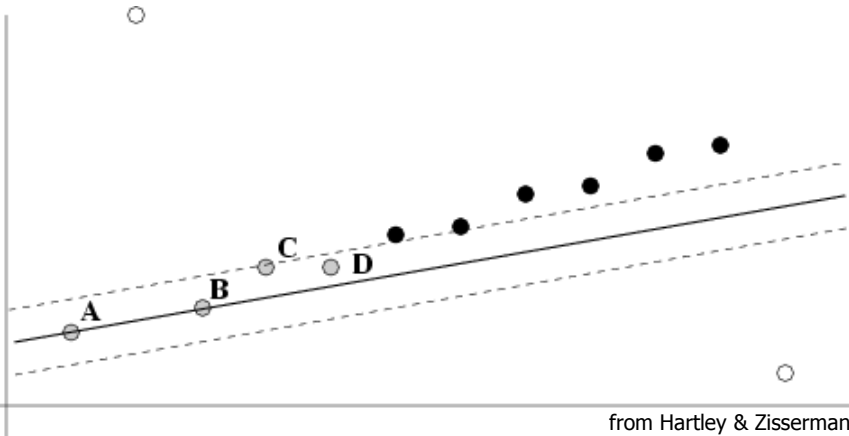


N adaptív meghatározása

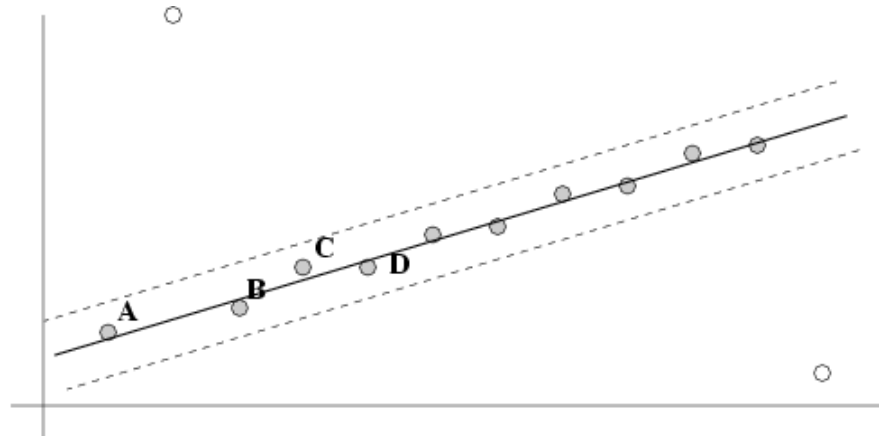
- Ha ε ismeretlen, akkor az alábbi módon meghatározhatjuk:
 1. Legyen $N=1$ és $\varepsilon=0.5$ (legrosszabb eset)
 2. Minden egyen mintára számoljuk meg az illeszkedő adatpontokat
 3. Frissítsük a kilógó adatok arányát, ha kisebb mint az előző:
 $\varepsilon = 1 - (\text{number of inliers}) / (\text{total number of points})$
 4. Állítsuk be N értékét a formula alapján
 5. Ha az eddigi minták száma meghaladja N értékét \rightarrow stop

RANSAC és LSE

- RANSAC kiszűri a kilógó adatokat, és a végén egy olyan becslést ad, amely a legnagyobb támogatottságú minimális adathalmazra épül
- Ezt javíthatjuk, ha az így megtalált összes illeszkedő pontra egy LSE illesztést végzünk.
- Így azonban megváltozhat az illeszkedő pontok halmaza is
 - Érdemes tehát alternálva kiszűrni a kilógó adatokat majd LSE illesztéssel meghatározni a többi pontra illeszkedő modellt.



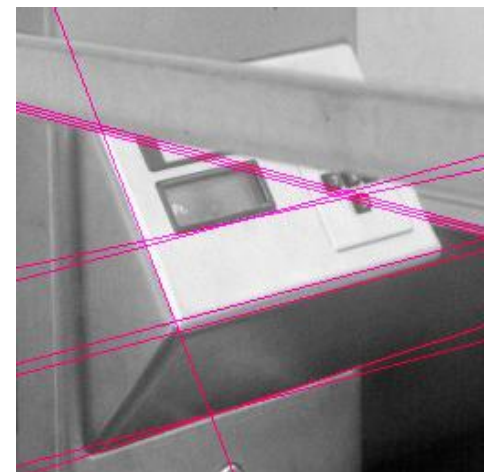
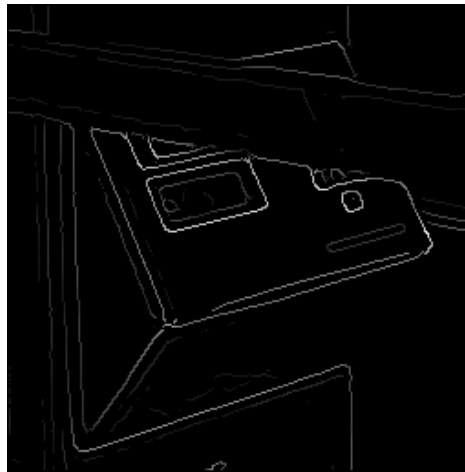
from Hartley & Zisserman



ALAKZATOK ÉS MINTÁK DETEKTÁLÁSA

Hough transzformáció

- Éldetektálás során csak élpontok halmazát kapjuk.
- Hogyan kereshetünk magasabb rendű struktúrákat, alakzatokat az élpontok halmazában?



- A Hough-transzformáció során a képen általában az $f(x, y; a_1, a_2, \dots, a_n) = 0$ a_1, a_2, \dots, a_n paraméterekkel explicit alakban megadható görbéket keressük
- A Hough transzformáció alkalmazása célravezető, ha
 - ismert alakú (és méretű) objektumokat keresünk a képen.
 - Akkor is, ha azok részben takartak vagy zajosak.

Egyenesek detektálása

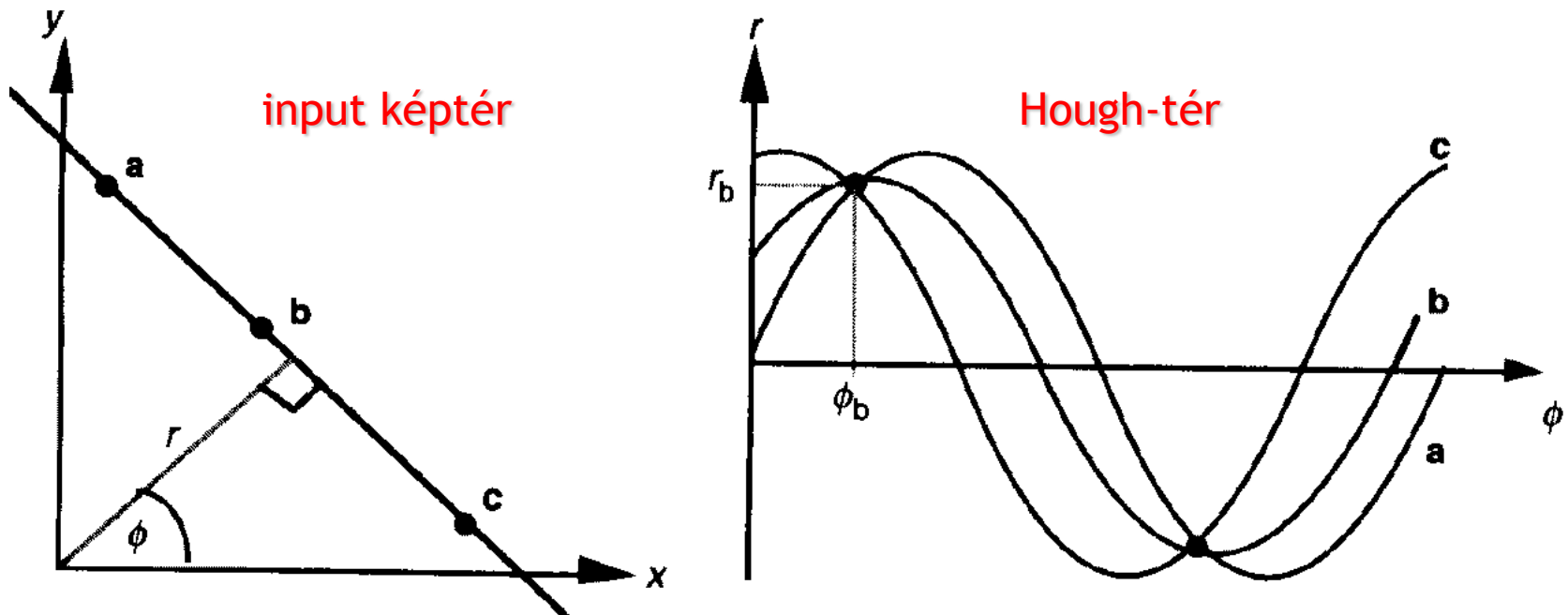
- Ekkor az input tér egy (x_i, y_i) pontjának az $0 \leq \varphi \leq 2\pi$

$$r = x_i \cdot \cos\varphi + y_i \cdot \sin\varphi$$

$$0 \leq r \leq r_{\max}$$

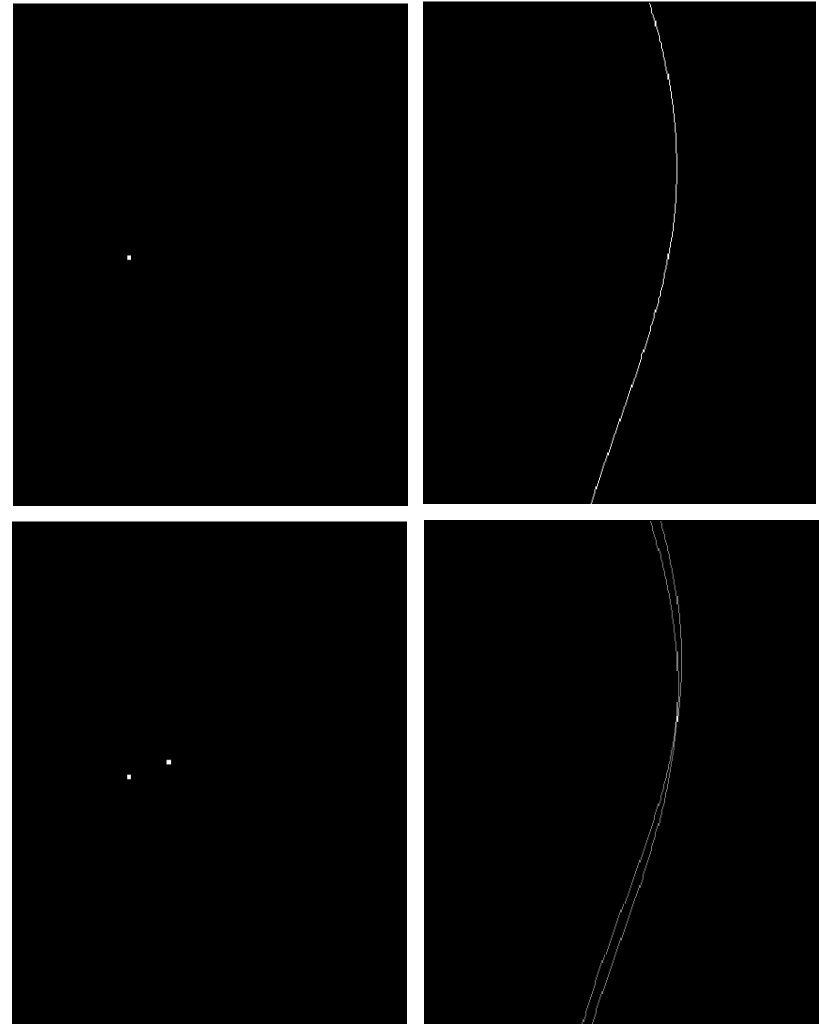
szinuszoid görbe felel meg a Hough-térben.

- Az egy egyenesbe eső pontokhoz tartozó szinuszoid görbék egy pontban metszik egymást.



Hogyan találjuk meg az egyeneseket?

- Egy (él)pont a képtérben megfelel egy szinusz görbének a Hough térben
 - Két pontnak két görbe felel meg
- Két (vagy több) ilyen görbe metszéspontja által reprezentált egyenesre ekkor kettő (vagy több) szavazat esett.
 - Az így kapott egyenes valamennyi rá szavazó ponton átmegy a képtérben.
- A Hough tér küszöbölésével megkapjuk a képtér egyeneseit



Hough transzformáció algoritmus

Create ϕ and r for all possible lines

Create an array A indexed by ϕ and r

for each point (x, y)

 for each angle ϕ

$$r = x \cdot \cos(\phi) + y \cdot \sin(\phi)$$

$$A[\phi, r] = A[\phi, r] + 1$$

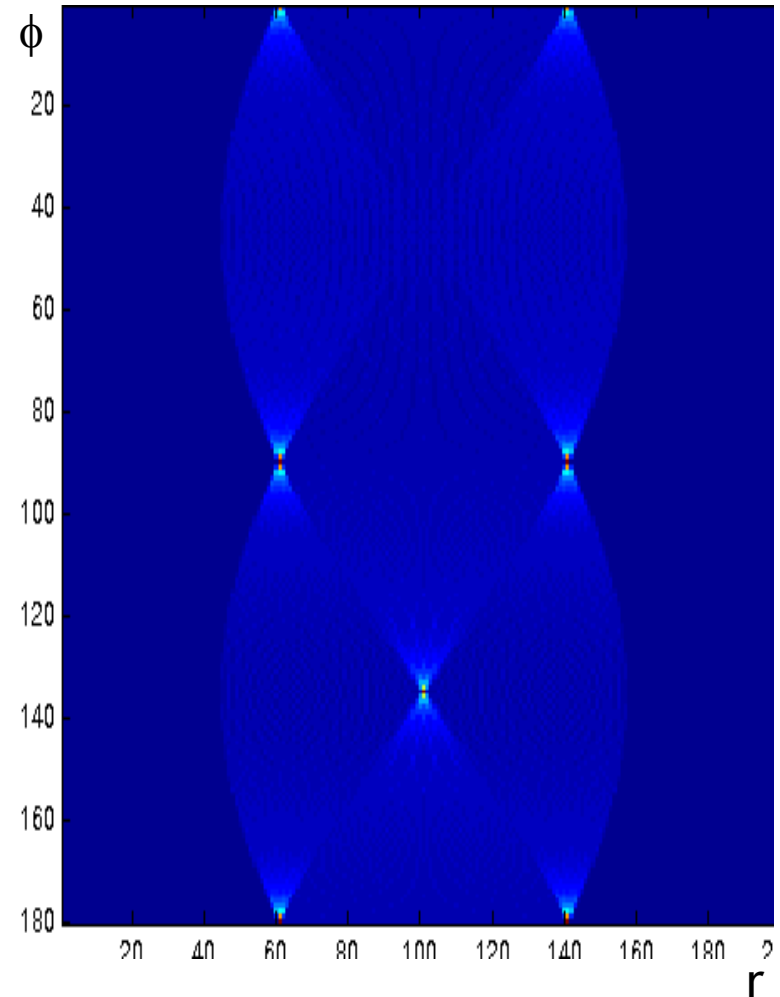
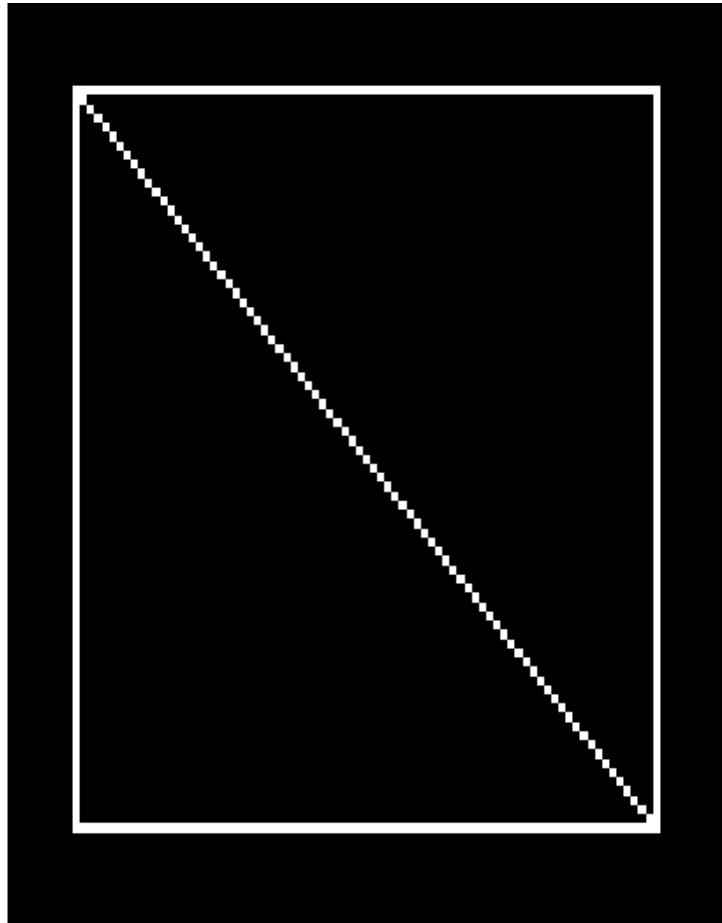
 end

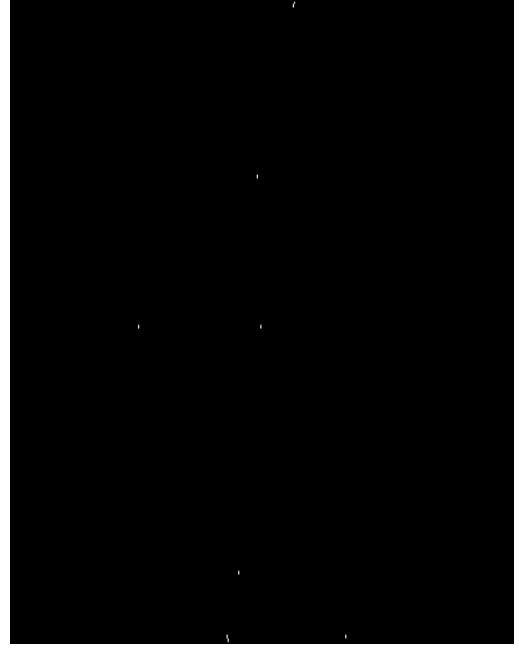
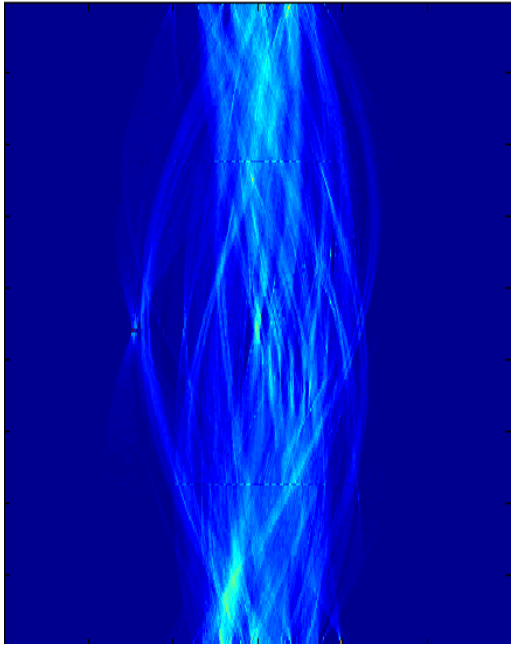
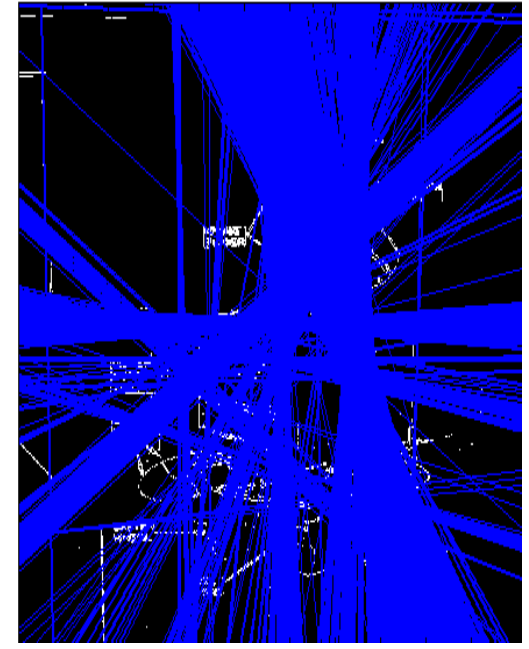
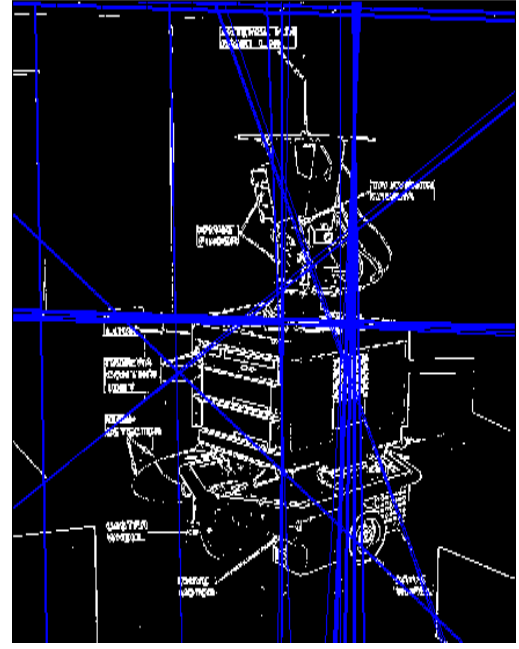
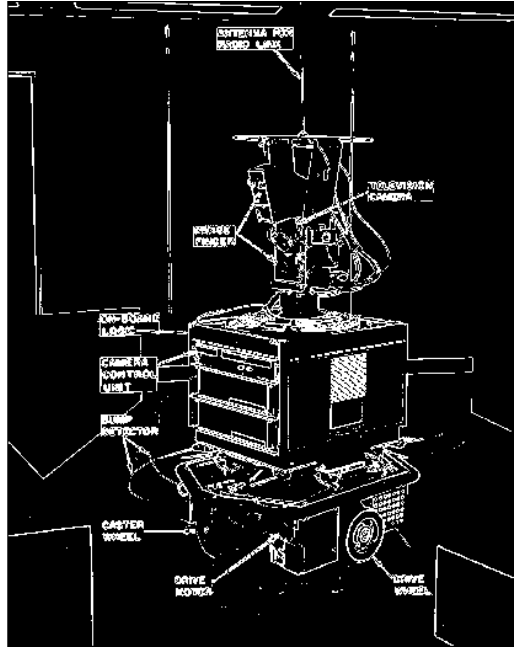
end

where $A > \text{Threshold}$ return the corresponding line parameters

Példa egyenesek detektálására

- A képen 5 egyenes található (4 oldal + 1 átló)



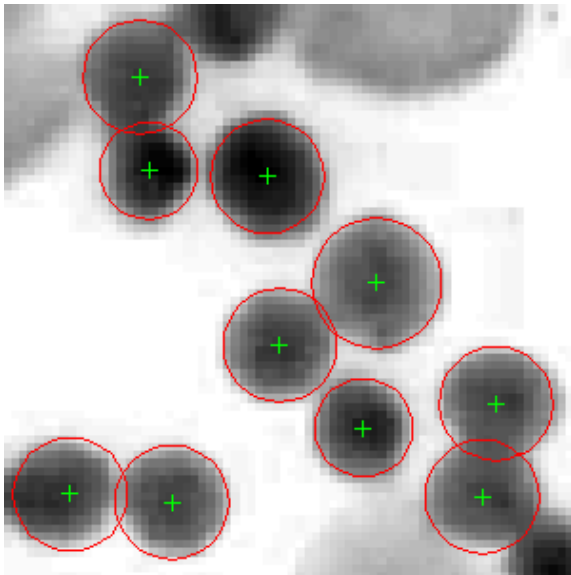


Körvonal detektálása

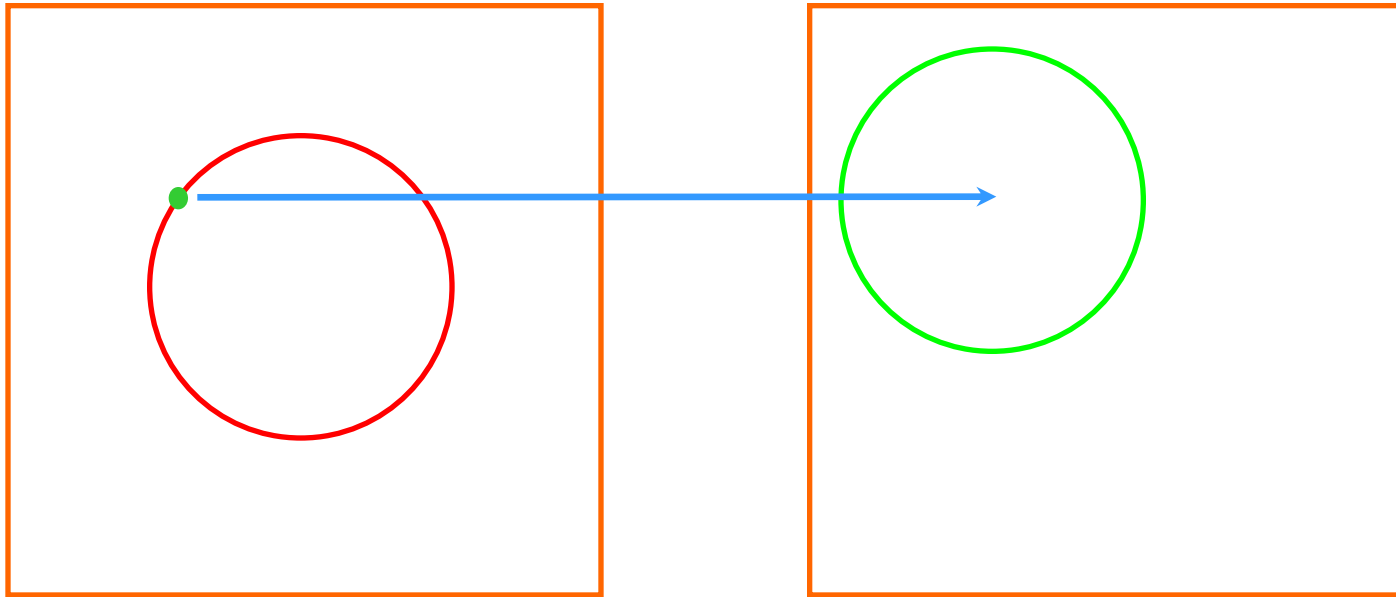
- Általános körök esetén az (a,b,r) Hough-tér 3-dimenziós lesz:

$$ax^2+by^2=r^2 \rightarrow f(x,y,a,b,r)=ax^2+by^2-r^2=0$$

- Amennyiben pl. adott (konstans) r -sugarú kört keresünk, akkor a paraméter-tér 2-dimenziósra csökken – a gyakorlatban ez használatos



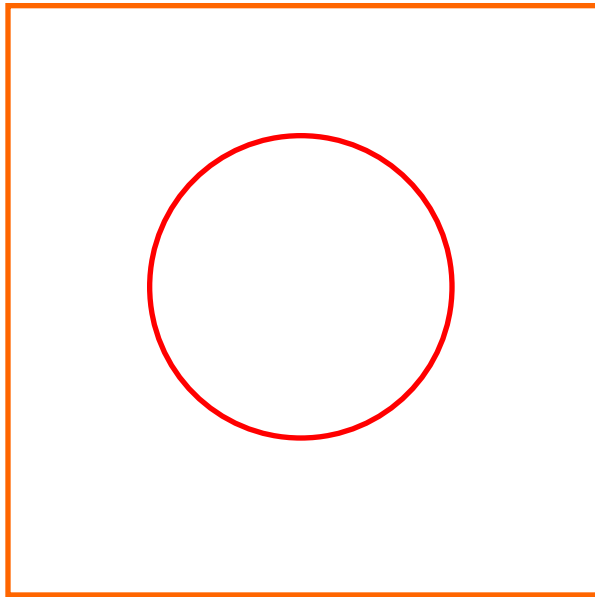
Körvonal detektálása



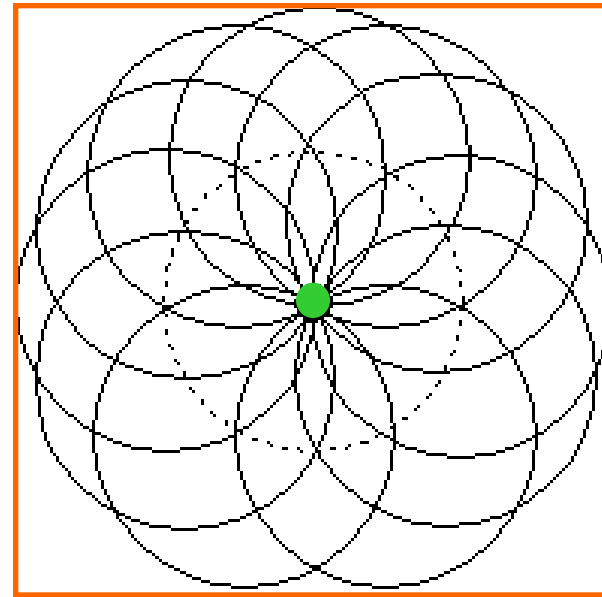
él-kép, amin ismert
sugarú kört keresünk

2D Hough-tér, ahol minden
élpontnak egy, a potenciális
középpontokat tartalmazó kör
felel meg

Körvonal detektálása

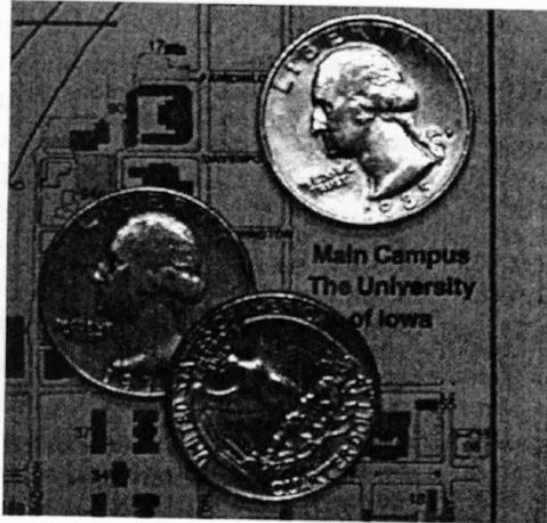


él-kép

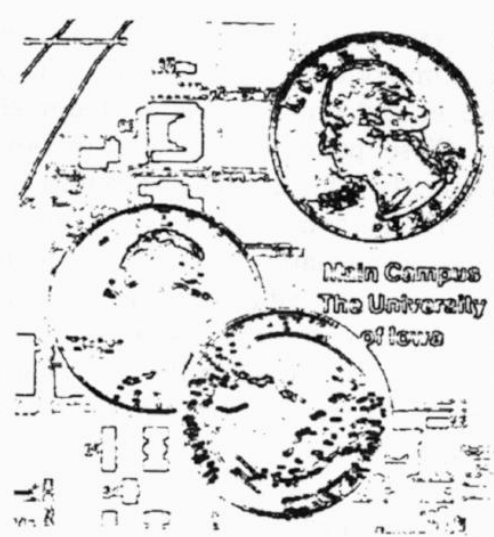


maximumhely(ek) a Hough-térben → a detektált kör(ök) középpontja(i)

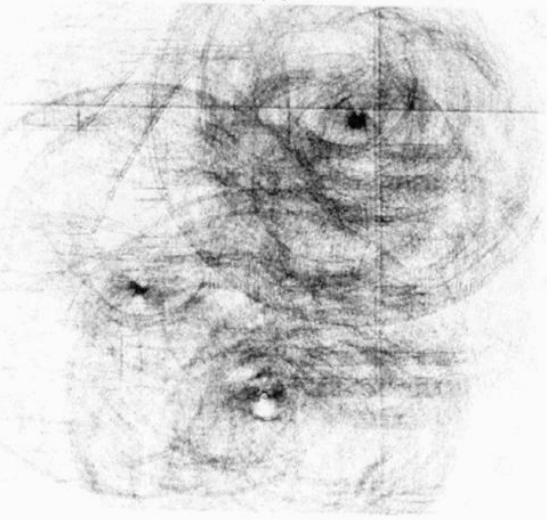
Példa körvonal detektálására



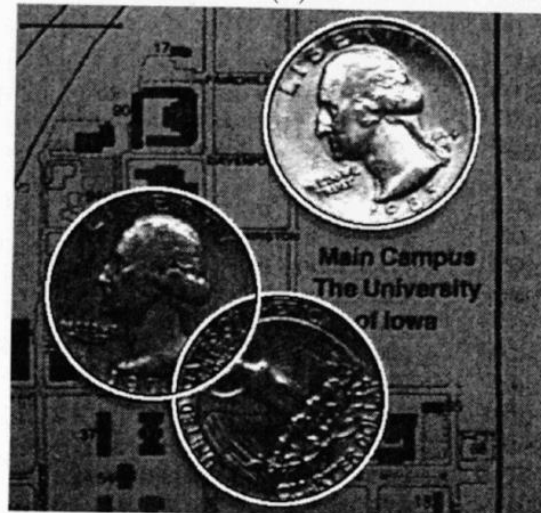
(a)



(b)



(c)



(d)

- (a) eredeti kép
- (b) él-kép
- (c) 2D paraméter tér (adott sugarú köröket keresünk)
- (d) detektált körök az eredeti képen

Minták keresése képeken

- Az illesztési módszerekkel ismert tárgyakat, minták előfordulásait keressük a képen.



Normalizált kereszt-korreláció (2D)

- A legegyszerűbb a normalizált kereszt-korreláció alapján illeszteni
 - Az illesztési minta $M \times N$ méretű
 - A kép minden pontjára kiszámoljuk a mintával vett kereszt-korreláció értékét az alábbi képlet alapján
 - x és y a minta illetve a kép vizsgált pixeleit jelölik

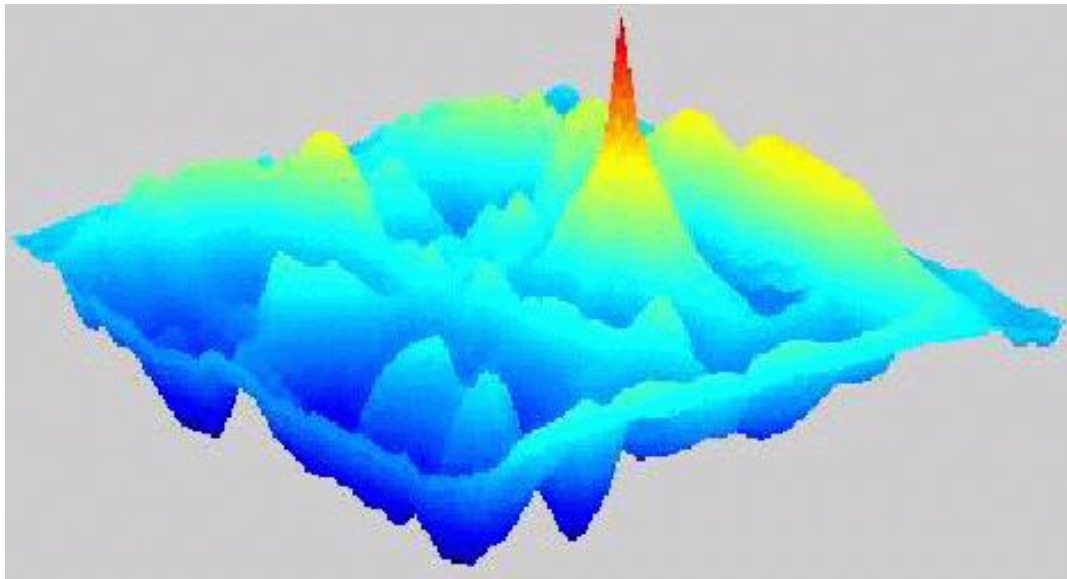
$$r_{xy}(d, e) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x_{ij} - \mu_x) \cdot (y_{i-d, j-e} - \mu_y)}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x_{ij} - \mu_x)^2} \cdot \sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (y_{ij} - \mu_y)^2}}$$

Illesztési algoritmus

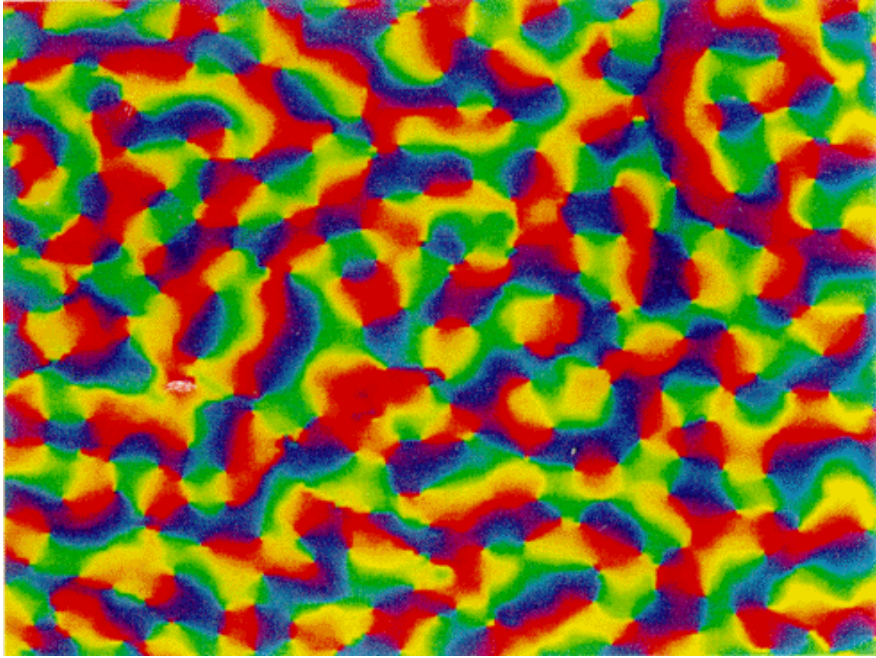
1. Számítsuk ki a mintának megfelelő illeszkedési kritériumot minden helyre (esetleg több méretre és irányra is) a képen.
2. Egy küszöbérték feletti lokális maximumhelyek megadják a minta előfordulási helyeit a képen.



az illesztési minta
és a kép

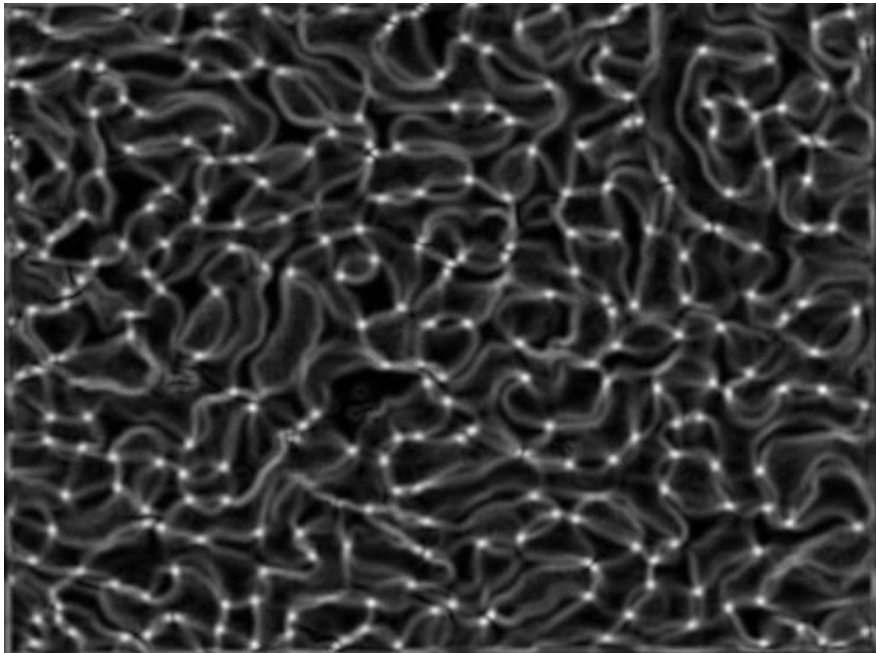


a kereszt-korreláció



← input kép

 ← illesztési minta



← kereszt-korreláció kép,
(fehér: magas korreláció,
fekete: alacsony
korreláció)

Alternatív illesztési kritériumok

$$C_1(u, v) = \frac{1}{1 + \max_{(i,j) \in V} |f(i+u, j+v) - h(i, j)|}$$

$$C_2(u, v) = \frac{1}{1 + \sum_{(i,j) \in V} |f(i+u, j+v) - h(i, j)|}$$

$$C_3(u, v) = \frac{1}{1 + \sum_{(i,j) \in V} (f(i+u, j+v) - h(i, j))^2}$$

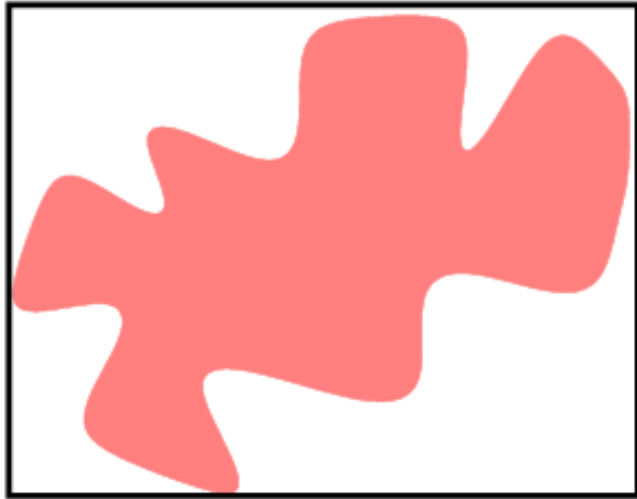
ahol f a feldolgozandó kép, h a keresendő minta és V a képpontok halmaza,

Hierarchikus illesztés

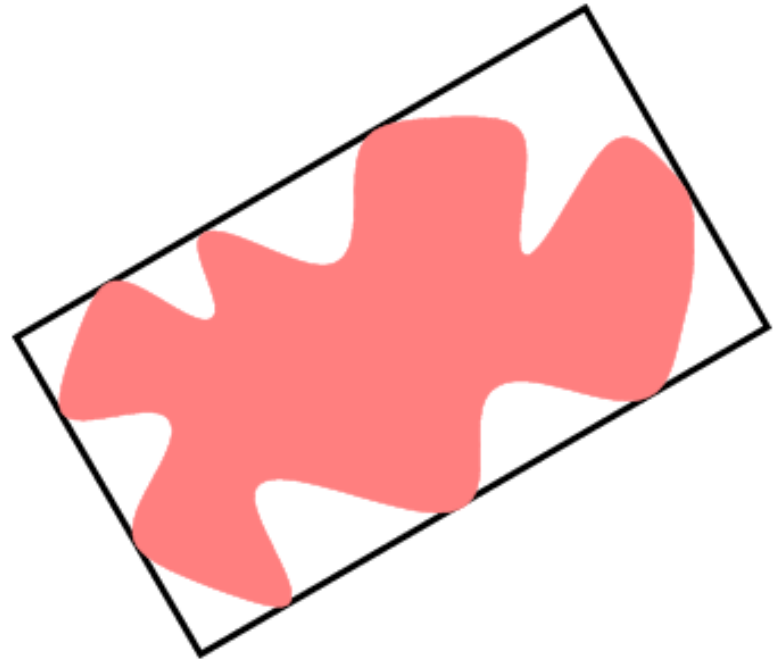
- A képpiramisok itt is jól használhatóak:
 - A piramis struktúrával csökkenthető a műveleti komplexitás.
- Először egy durvább mintát illesztünk egy durvább képen (kevesebb művelet).
- Utána már csak azokat az illesztéseket vizsgáljuk finomabb felbontásban, amelyek kritériuma meghaladt egy bizonyos küszöböt.

EGYSZERŰ ALAKZATJELLEMZŐK

Befoglaló téglalap



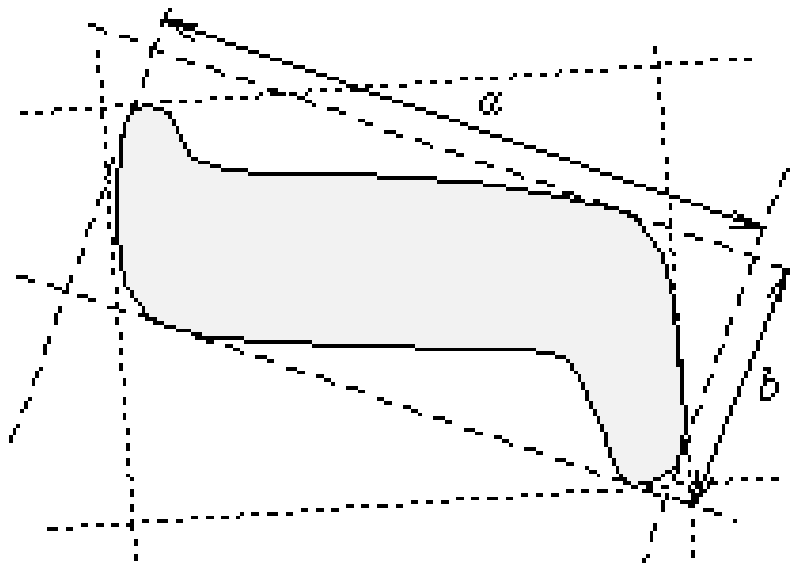
„álló” befoglaló téglalap



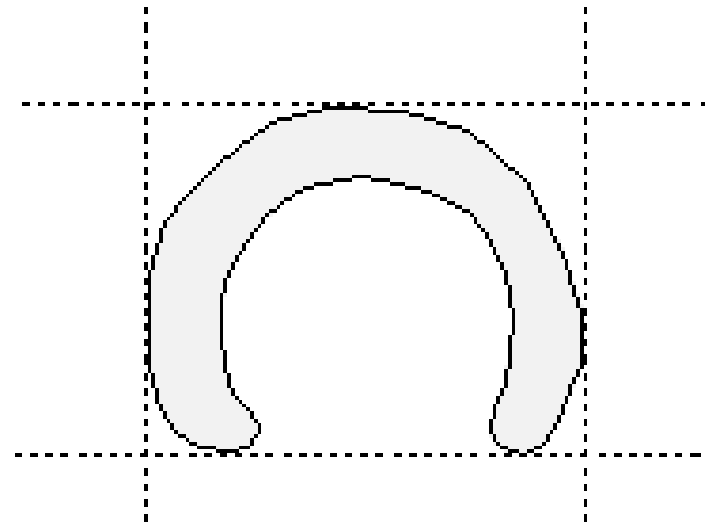
minimális befoglaló téglalap

Rektangularitás (téglalap-szerűség)

az alakzat területének és a minimális befoglaló téglalap területének a hányadosa

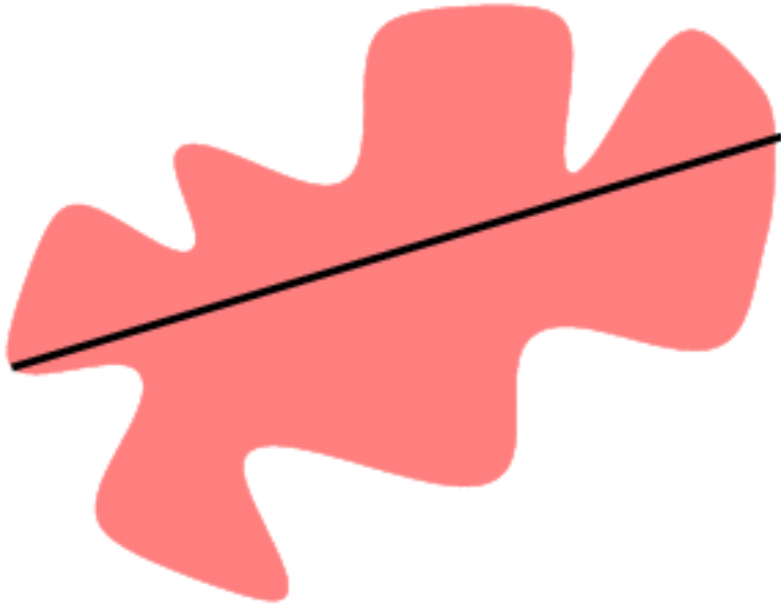


téglalap-szerű

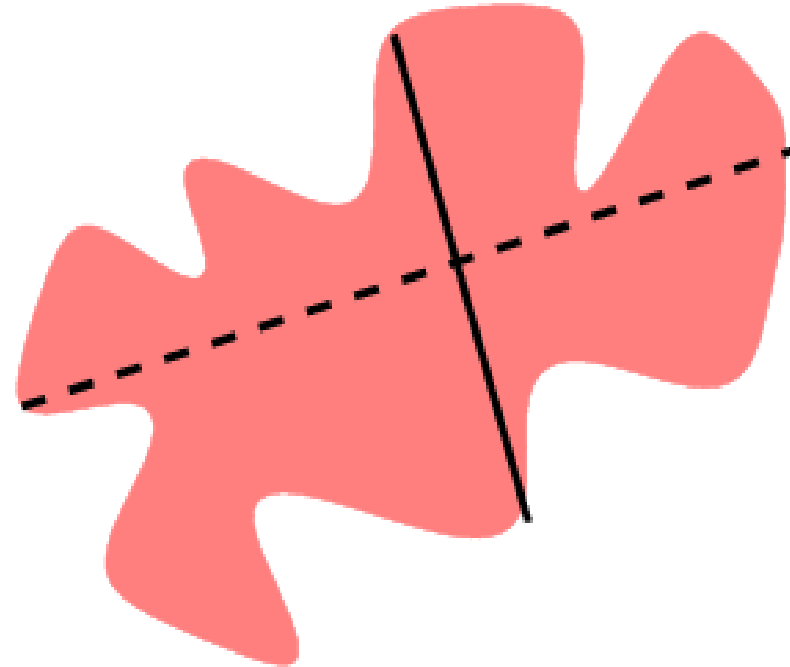


nem téglalap-szerű

Fő- és melléktengely



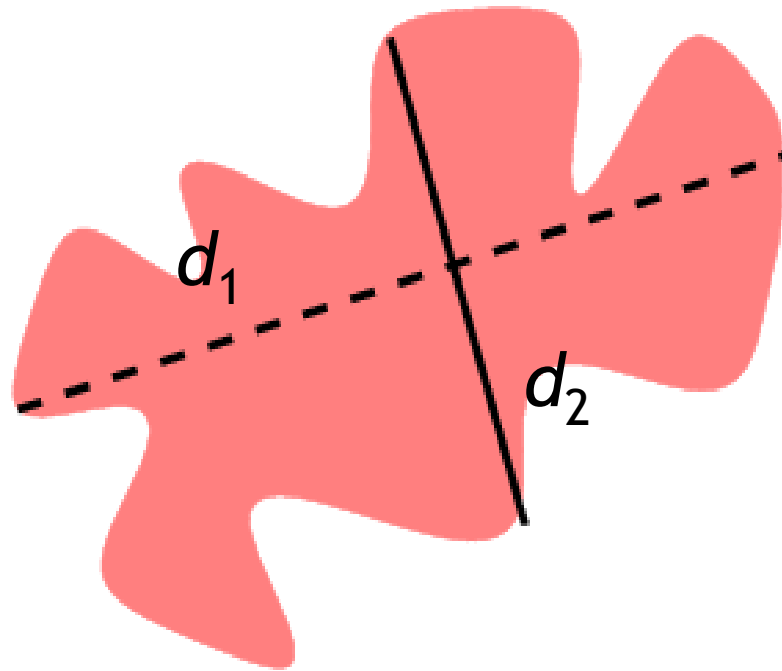
főtengely: az alakzaton belül haladó leghosszabb egyenes szakasz



melléktengely: az alakzaton belüli, a főtengelyre merőleges leghosszabb egyenes szakasz

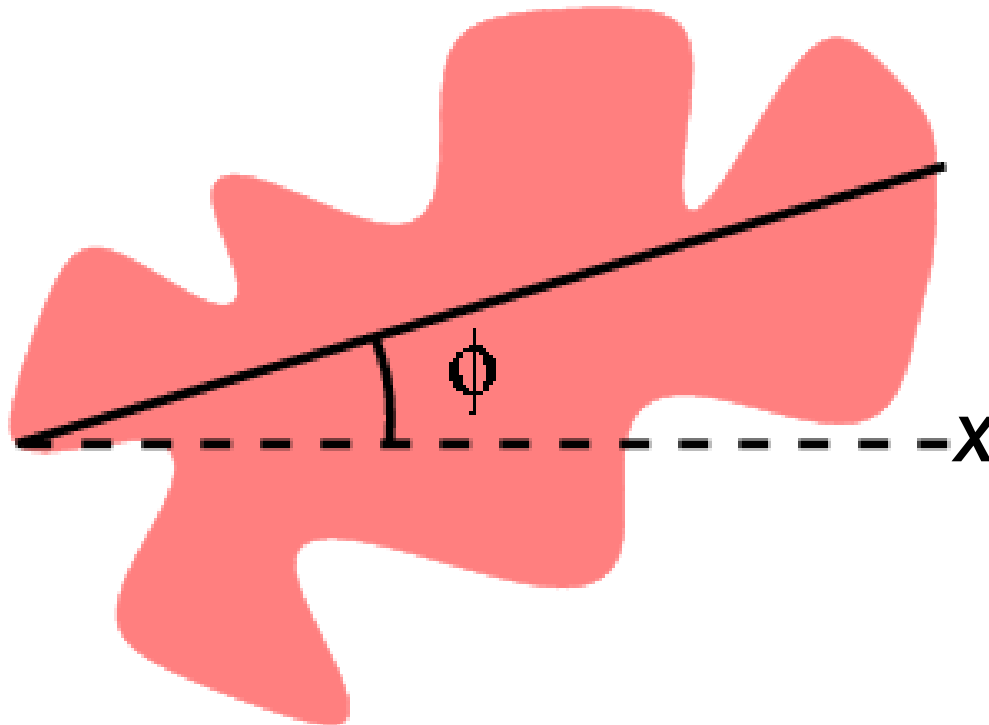
Excentricitás

a fő- és a melléktengely hosszaránya: d_1 / d_2



Főtengely szöge (az alakzat iránya)

a főtengely és az x-tengely által bezárt szög



Átmérő

$$\text{Diam}(S) = \max_{p, q \in \partial S} \{ d(p, q) \},$$

ahol:

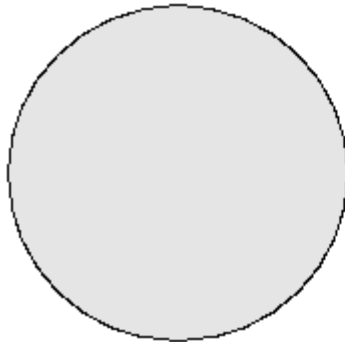
- S : alakzat,
- ∂S : alakzat határa,
- d : távolság.

Kompaktság

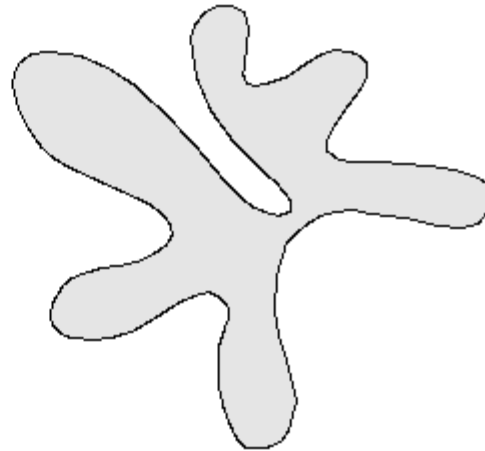
kompaktság = $(\text{kerület})^2 / \text{terület}$

pl. kör: $4\pi \approx 12.6$

négyzet: 16



erősen kompakt



gyengén kompakt

Cirkularitás (körszerűség)

$$\begin{aligned}\text{cirkularitás} &= 1 / \text{kompaktság} \\ &= \text{terület} / (\text{kerület})^2\end{aligned}$$

maximális a körre: $1 / (4\pi) \approx 0.08$.

Momentumok

Az I kép $(p+q)$ -adrendű (képi) momentumai
 $(p, q=0, 1, 2, \dots)$

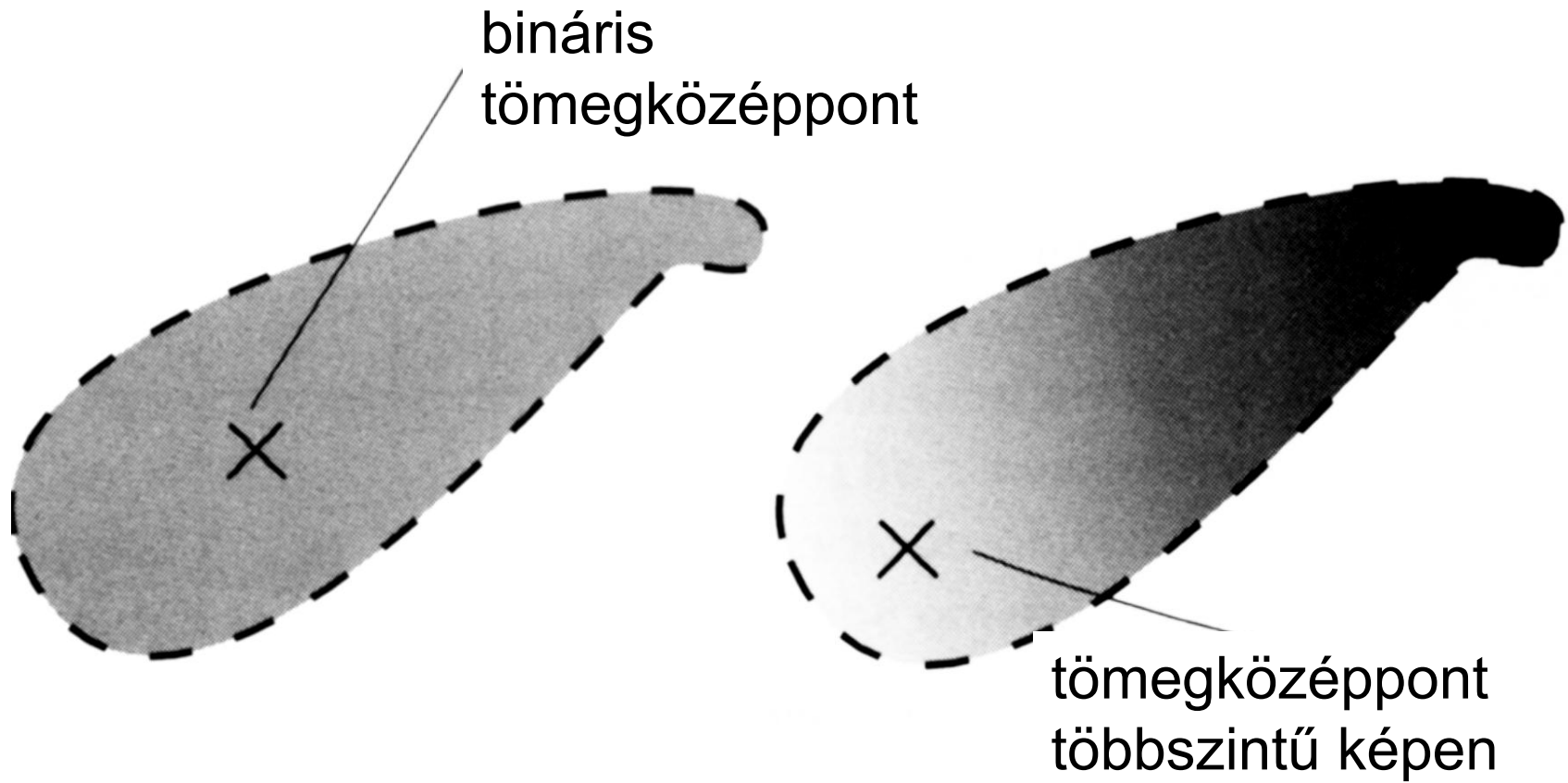
$$m_{p,q} = \sum_x \sum_y x^p \cdot y^q \cdot I(x, y)$$

Az S bináris alakzat $(p+q)$ -adrendű
 (geometriai) momentumai:

$$m_{p,q} = \sum_{(x,y) \in S} x^p \cdot y^q$$

$$(\bar{x}, \bar{y}) = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}} \right)$$

Súlypont



Centrális momentumok

Az I kép $(p+q)$ -adrendű centrális (képi) momentuma:

$$\mu_{p,q} = \sum_x \sum_y (x - \bar{x})^p \cdot (y - \bar{y})^q \cdot I(x, y)$$

Az S bináris alakzat $(p+q)$ -edrendű centrális (geometriai) momentuma:

$$\mu_{p,q} = \sum_{(x,y) \in S} (x - \bar{x})^p \cdot (y - \bar{y})^q$$

Normalizált centrális momentumok: $\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{p,q}^\gamma}$

$$\gamma = 1 + (p + q)/2$$

Excentricitás

$$\varepsilon = \frac{(\mu_{2,0} - \mu_{0,2})^2 + 4\mu_{1,1}^2}{(\mu_{2,0} + \mu_{0,2})^2}$$

0 és 1 közötti érték.

Főtengely szöge

$$\Phi = \frac{1}{2} \arctan \left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} \right)$$

Affin invariáns momentumok

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned}$$

$$\begin{aligned} \phi_6 = & (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned}$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ & + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]. \end{aligned}$$

Felhasznált anyagok

- **Palágyi Kálmán: Digitális Képfeldolgozás**
/pub/Digitalis_kepfeldolgozas
- **További források az egyes diákon megjelölve**