

V. SIFT és RANSAC

1. Kulcspont illesztés két kép között affin transzformációval három pontpár bevonásával. Az egy pontpár esetén keletkező két egyenlet alakja.

SIFT (Scale-Invariant Feature Transform)

Affin transzformáció: lineáris transzformáció és eltolás ($x \rightarrow Ax + b$)

3 db kulcspont illesztése két kép (sztereó, optical flow) között, ezek alapján keressük az affin transzformációt. Minden pontpár között:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

Ismeretlenek: m_i és t_x, t_y (forgatás, skálázás, nyújtás és eltolás)

3 pontpár esetén megoldható lineáris egyenletrendszerre vezet:

$$\begin{bmatrix} x & y & 0 & 0 & 1 & 0 \\ 0 & 0 & x & y & 0 & 1 \\ & & \dots & & & \\ & & \dots & & & \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} u \\ v \\ \vdots \end{bmatrix}$$

Inhomogén lineáris egyenletrendszer, $Ax = b$ alakban, ennek LS megoldása:

$$x = [A^T A]^{-1} A^T b$$

2. A kulcspont (feature) detektorral és deskriptorral szembeni elvárások. A Harris detektor matematikája, régió, él és sarok detektálás sajátérték technikával. A kulcspont (feature) illesztéssel szembeni elvárások.

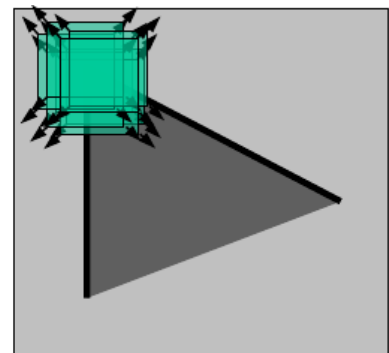
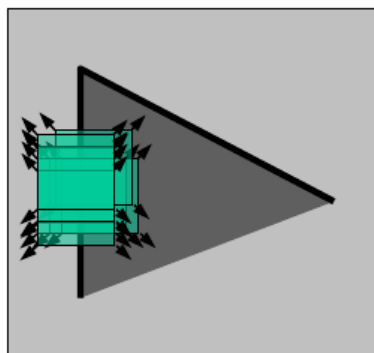
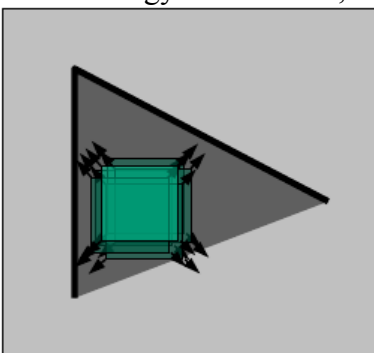
A kulcspont (feature) illesztéssel szembeni elvárások:

A kulcspont (feature) detektorral és deskriptorral szembeni elvárások:

- Könnyű megtalálni
- Invariáns: képzajra, megvilágításra, skálázásra, eltolásra, forgatásra
- Könnyen összehasonlítható egy objektum tulajdonságokat tartalmazó adatbázis elemeivel
- deskriptor legyen megkülönböztető

A Harris detektor:

Élkeresés egy kis ablakkal, intenzitásváltozás eltolás hatására:



sík: nincs intenzitásváltozás

él: az él mentén nincs

sarok: minden irányban van

Intenzitásváltozás az $[u,v]$ elmozdulás függvényében:

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

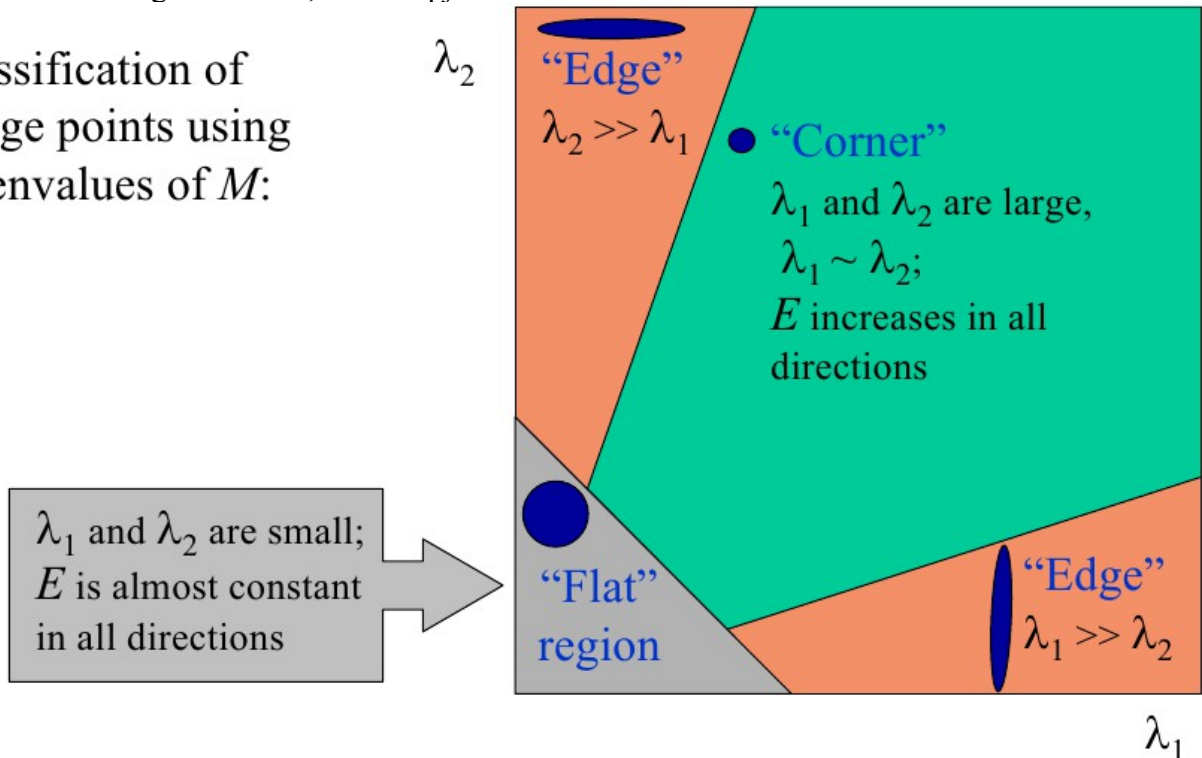
ahol $I(x,y)$ az intenzitás, $w(x,y)$ az ablak függvénye (Gauss vagy egyszerű ablak)

Kis eltolásokra bilineáris approximációt kapunk:

$$E(u,v) \approx [u,v] M \begin{bmatrix} u \\ v \end{bmatrix} \text{ ahol } M = \sum_{x,y} w(x,y) \begin{bmatrix} I_{xx}^2 & I_x I_y \\ I_x I_y & I_{yy}^2 \end{bmatrix}$$

M sajátértéket meghatározzuk, ezek alapján:

Classification of image points using eigenvalues of M :



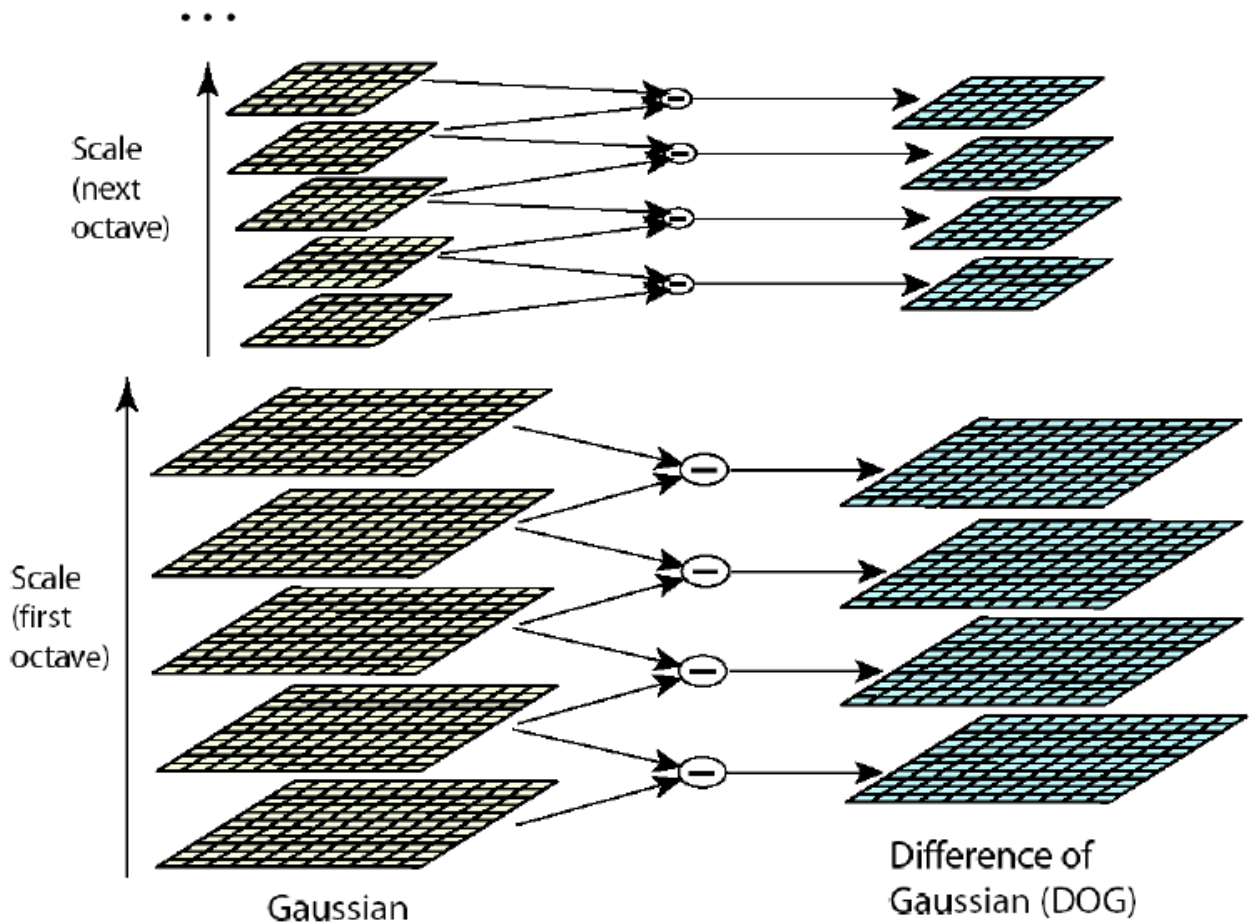
A sarok minősége (corner response): $R = \frac{\det(M)}{\text{trace}(M)}$, $\det(M) = \lambda_1 \lambda_2$, $\text{trace}(M) = \lambda_1 + \lambda_2$

Az algoritmus: olyan pontok keresése, ahol R nagyobb egy adott küszöbnél, R lokális maximum pontjainak keresése

R invariáns a forgatásra, a Harris detektor invariáns az intenzitás skálázására és eltolására, de a kép skálázásra nem.

3. SIFT deskriptor vektor felépítése. Detektálás többszörös skálában. A Difference of Gaussians (DOG) piramis. A kanonikus orientáció meghatározása. A 128 dimenziós SIFT vektor alak. A lokális kulcsponatok (local features) előnyei.

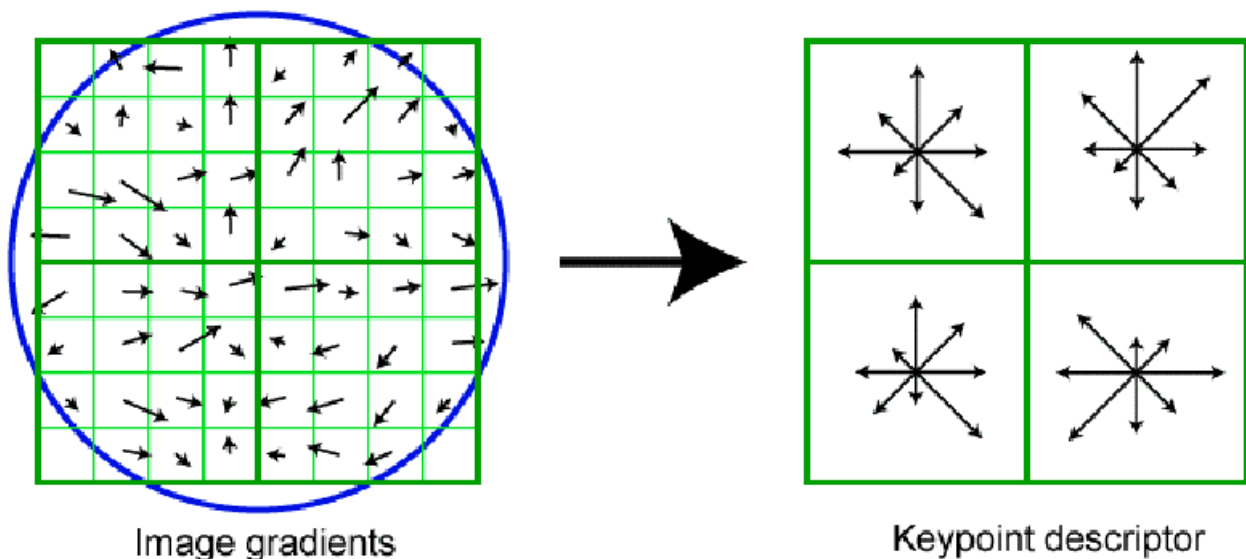
Minden méretet (skálát) meg kell vizsgálni, hogy meghatározható legyenek a skálázásra invariáns kulcsponatok. Erre egy módszer a Difference of Gaussians (DOG).



A lokális gradiensek irányáról egy választott skálázási értéknél hisztogram készítése $[0, 2\pi]$, a legmagasabb értékhez rendeljük a kanonikus orientációt.

A gradiens nagyságát és irányát egy 16×16 tömbön számítjuk ki a kulcspont körül.

- A gradienseket Gauss -féle ablak szerint súlyozzuk
- Az orientációk hisztogramjának tömböt hozunk létre.
- $8 \text{ orientáció} \times 4 \times 4 \text{ hisztogram tömb} = 128 \rightarrow 128 \text{ dimenziós SIFT vektor alak}$



A lokális kulcspontok (local features) előnyei:

- lokalitás: a tulajdonság lokálisak, robusztus
- megkülönböztethető: az egyéni tulajdonságok összehasonlíthatók egy adatbázis elemeivel
- minőség: sok tulajdonság illeszthető egy kis tárgyhoz is
- bővíthetőség: különböző tulajdonság-típusokkal bővíthető, melyek növelik a robusztusságot

4. A RANSAC algoritmus koncepciója, az algoritmus paramétereinek megválasztása. A RANSAC hurok lépései homográfia becslése esetén. Inliers/outliers meghatározás. A feature-bázisú képillesztés előnyei.

RANSAC algoritmus koncepciója:

- Véletlenszerűen S pont kiválasztása
- Az S pont segítségével megoldani a problémát
- A megoldás ellenőrzése a maradék pontok segítségével
- Többször ismételni az előzőeket
- Végül a legjobban illeszkedő megoldást választjuk, és azokat a pontokat, amelyek illenek hozzá

Paraméterek: Meg szeretnénk győződni róla, hogy az S minták közül P valószínűséggel legalább egyben nincsenek kívülálló (outlier) pontok.

Inlier: olyan pontok, amelyek eloszlása megmagyarázható egy modellel

Outlier: olyan pontok, amelyek nem illenek a modellbe

$$(1 - w^S)^N = 1 - P, \text{ ebből } N = \log \frac{(1 - P)}{\log(1 - w^S)}$$

RANSAC hurok homográfia becslése esetén:

1. Négy pont kiválasztása véletlenszerűen
2. H homográfia számítása
3. inlier pontok számítása, ahol $SSD(p_i', H p_i) < \epsilon$, majd folytatás az egyes ponttól
4. A legnagyobb inlier pontthalmaz megtartása
5. Minden inlier pontra újraszámítani az LS H becslést

Előnyök:

- kulcspontok keresésére létezik robusztus megoldás
- nem csak sarokpontokat keres
- Jó eloszlással rendelkezhet
- invariáns skálázásra, orientációra