

8. Pontmegfeleltetések

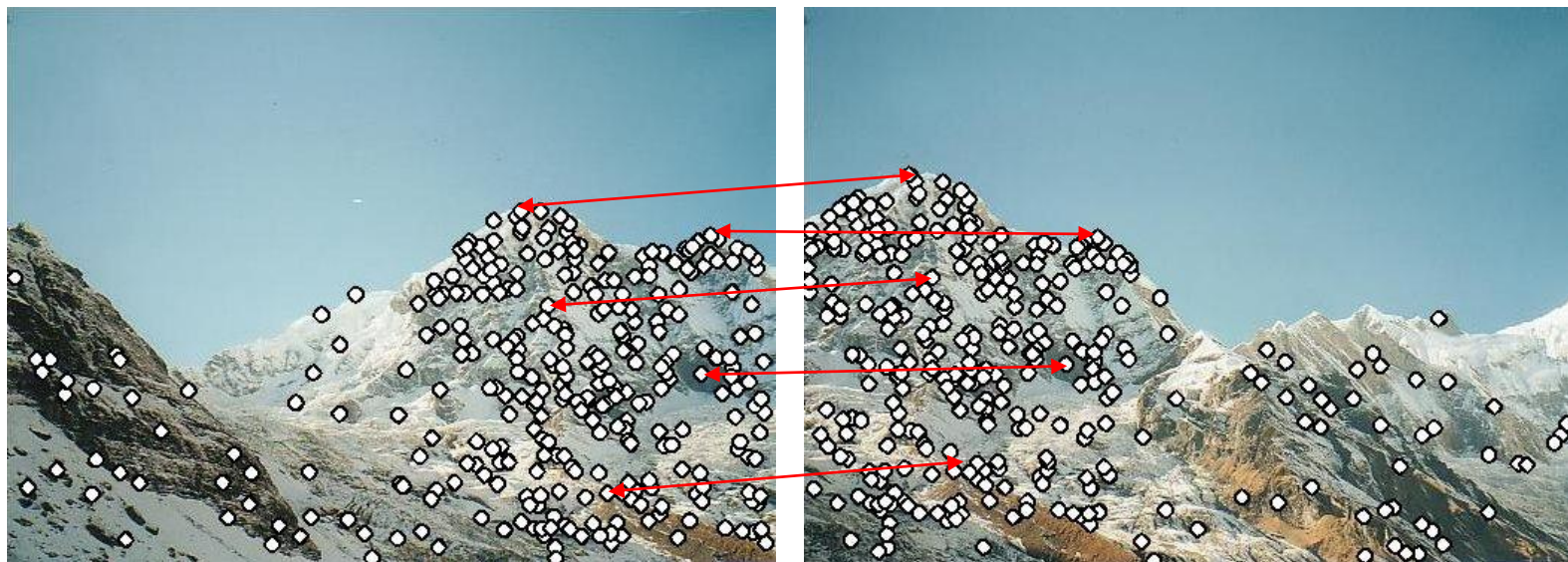
Kató Zoltán

Képfeldolgozás és Számítógépes Grafika tanszék

SZTE

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

Példa: panoráma kép készítés



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

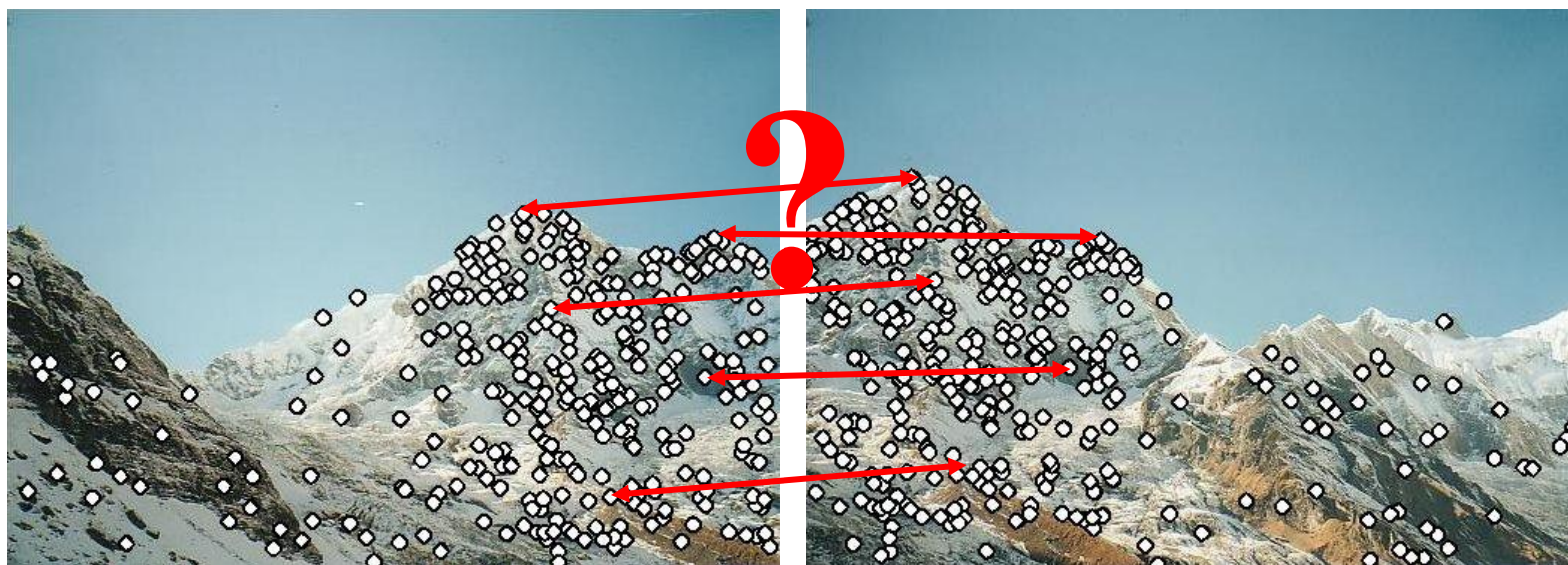


Megválaszolandó kérdések

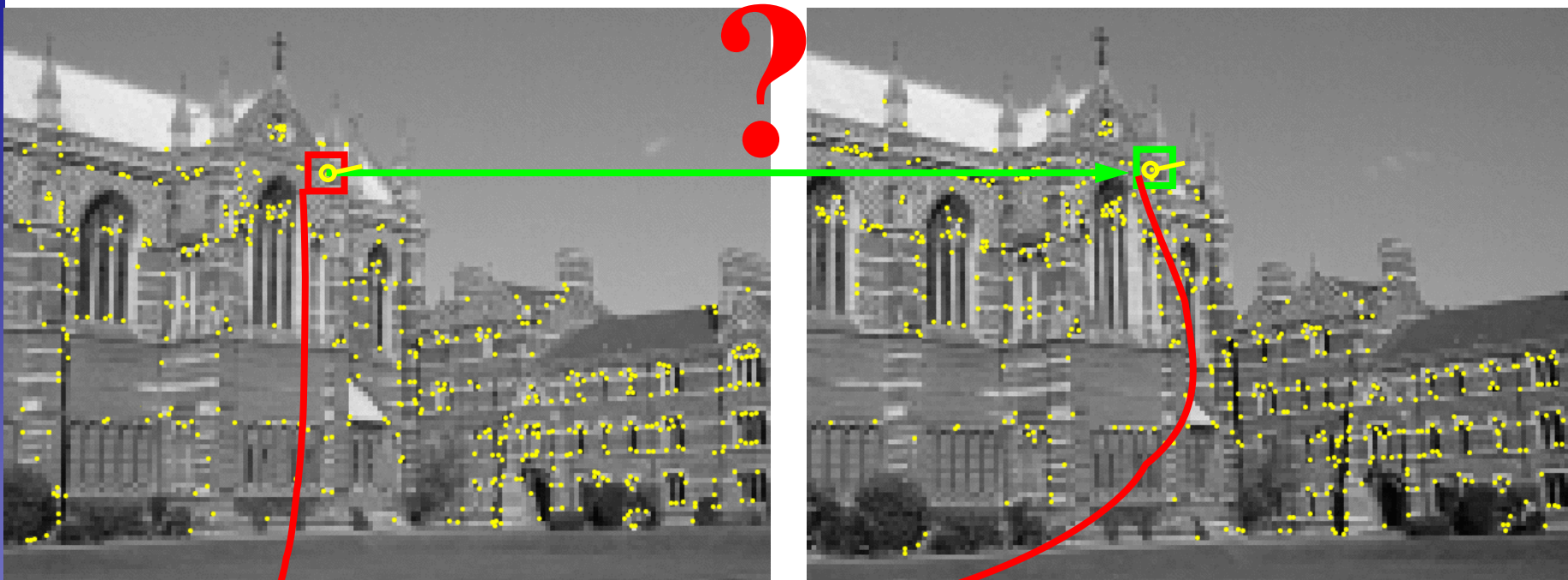
- Melyek azok a pontok, amelyeket megbízható módon detektálhatunk a képeken?
 - Sarokpontok
- Hogyan tudjuk leírni/jellemezni a kinyert pontokat?
 - Invariáns jellemzők
- Hogyan feleltessünk meg két képről kinyert pontokat?
 - Jellemzők összehasonlítása, robusztusság

Lokális leírók

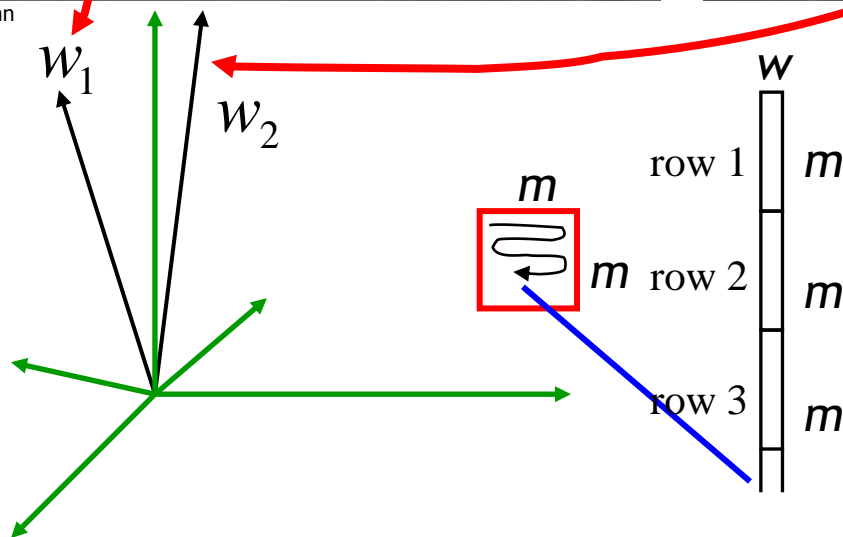
- A detektált pontokat hogyan tudnánk leírni úgy, hogy az
 - invariáns és
 - egyedi legyen
- A kinyert pontok önmagukban nem jellemezhetőek jól
 - egyetlen intenzitás-érték nem elég stabil és egyedi
 - A pontok környezete már elég egyedi lehet, de az invariancia biztosítása nem triviális
- Tekintsük a pontot tartalmazó ablak tartalmát



Képfoltok mint vektorok



from Hartley & Zisserman



Mindkét $m \times m$ ablak tartalma ábrázolható egy m^2 dimenziós vektorként.

Az ablakok tartalmát normalizálva a vektorok egységnyiek lesznek, miáltal könnyen összehasonlíthatóak a két vektor különbségeként.

Az ablakok tartalmának normalizálása

- Még azonos kamerák esetén is lehet eltérés a rögzített intenzitás értékekben (pl. eltérő gain/érzékenység).
- Ennek kiegyenlítésére célszerű a képeket normalizálni
 - Ezáltal az ablakokból képzett vektorok egységnyiek lesznek, így különböző képek között is összehasonlíthatóak

$$\bar{I} = \frac{1}{|W_m(x,y)|} \sum_{(u,v) \in W_m(x,y)} I(u,v)$$

Átlag pixelérték

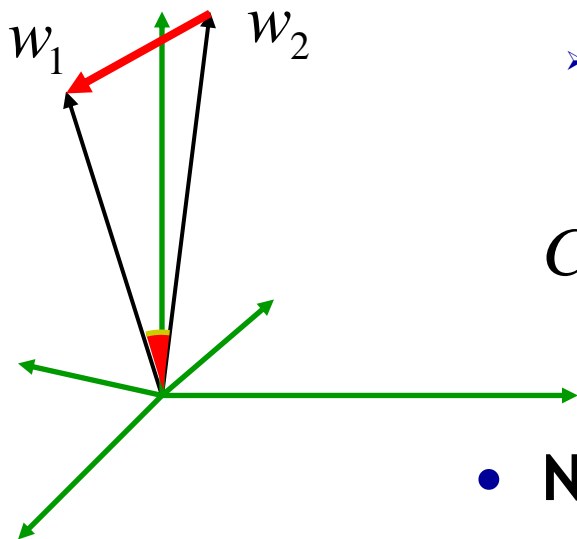
$$\|I\|_{W_m(x,y)} = \sqrt{\sum_{(u,v) \in W_m(x,y)} [I(u,v)]^2}$$

Ablak magnitúdó

$$\hat{I}(x,y) = \frac{I(x,y) - \bar{I}}{\|I - \bar{I}\|_{W_m(x,y)}}$$

Normalizált pixelérték

Hasonlósági mértékek



- **SSD: (Normalized) Sum of Squared Differences**
 - Minél kisebb annál jobban illeszkedik a két vektor

$$C_{\text{SSD}} = \sum_{(u,v) \in W_m} [\hat{I}_1(u,v) - \hat{I}_2(u,v)]^2 = \|w_1 - w_2\|^2$$

- **NC: Normalized Correlation**
 - Minél nagyobb, annál jobban illeszkedik a két vektor

$$C_{\text{NC}} = \sum_{(u,v) \in W_m} \hat{I}_1(u,v) \hat{I}_2(u,v) = w_1 \cdot w_2 = \cos \theta$$

Képfolt alapú leíró invarianciája

- Fotometriai invariancia a normalizálással elég jól teljesül
- Azonos méretű ablakok hasonlíthatók össze → skálafüggő
- Azonos állású (vagyis rasztorsorokra illeszkedő) ablakok hasonlíthatók össze → orientáció függő
- Összességében tehát
 - nincs geometriai invariancia,
 - van részleges fotometriai invariancia

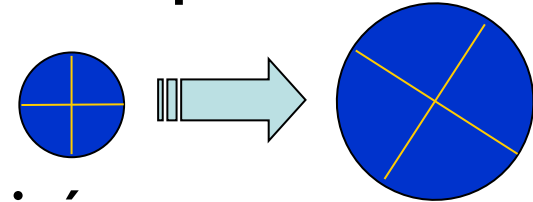
SIFT: SCALE INVARIANT FEATURE TRANSFORM

SIFT - Scale Invariant Feature Transform

- Skála- és irányfüggetlen fotometriailag invariáns pontleírókat állít elő az alábbi főbb lépésekben:

1. skála meghatározása (ez már megtörténik a pontok detektálása során)

- DoG szélsőhelyek térben és skálában



2. lokális orientáció a domináns gradiens irány

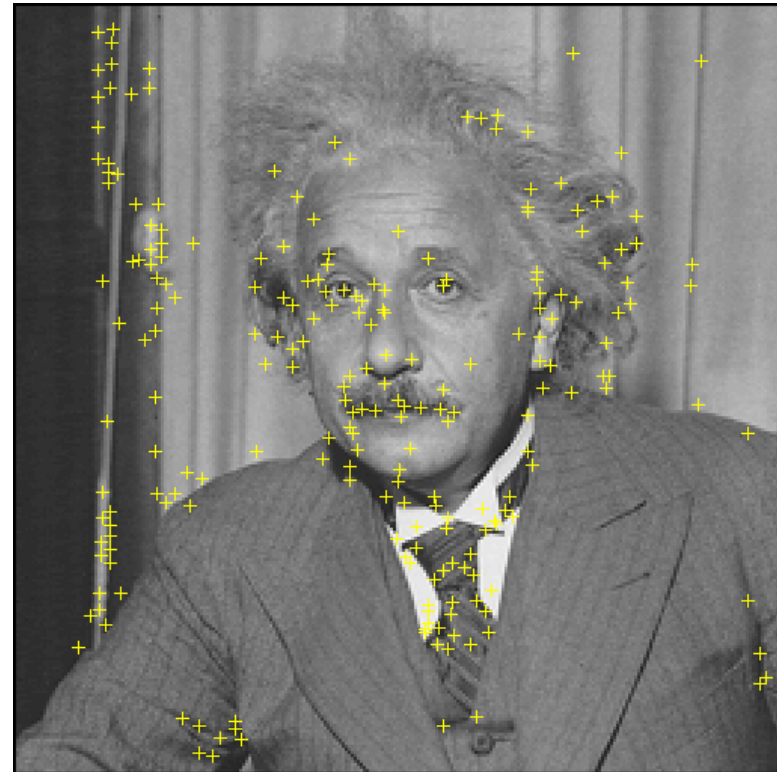
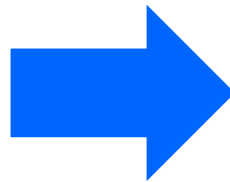
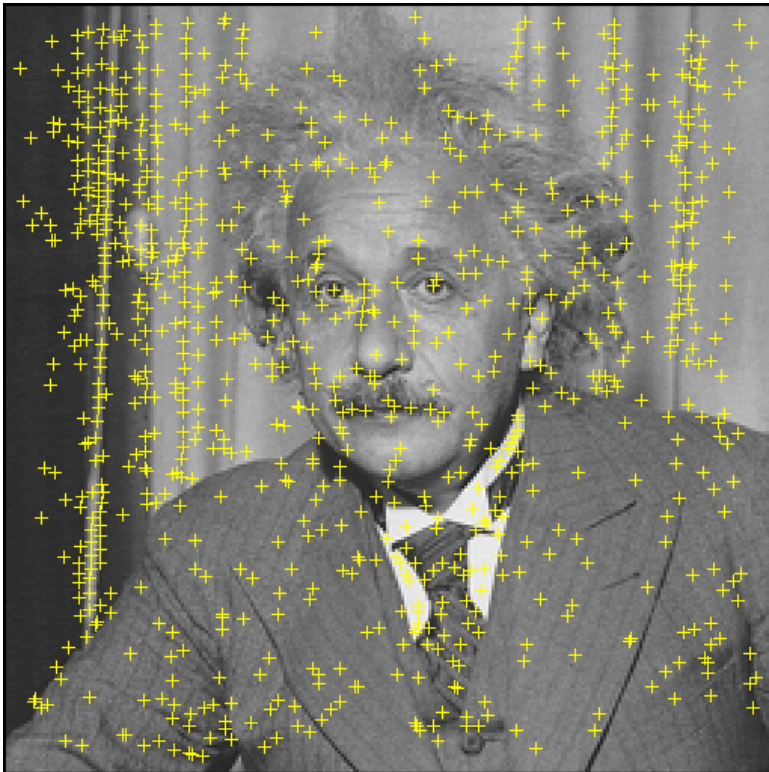
3. Az így kapott skála és orientáció minden egyes kinyert pontban egyértelműen meghatároz egy lokális koordinátarendszert

- Minden további számítás ebben a koordinátarendszerben történik, így a kapott leírók skála- és irányfüggetlenek lesznek

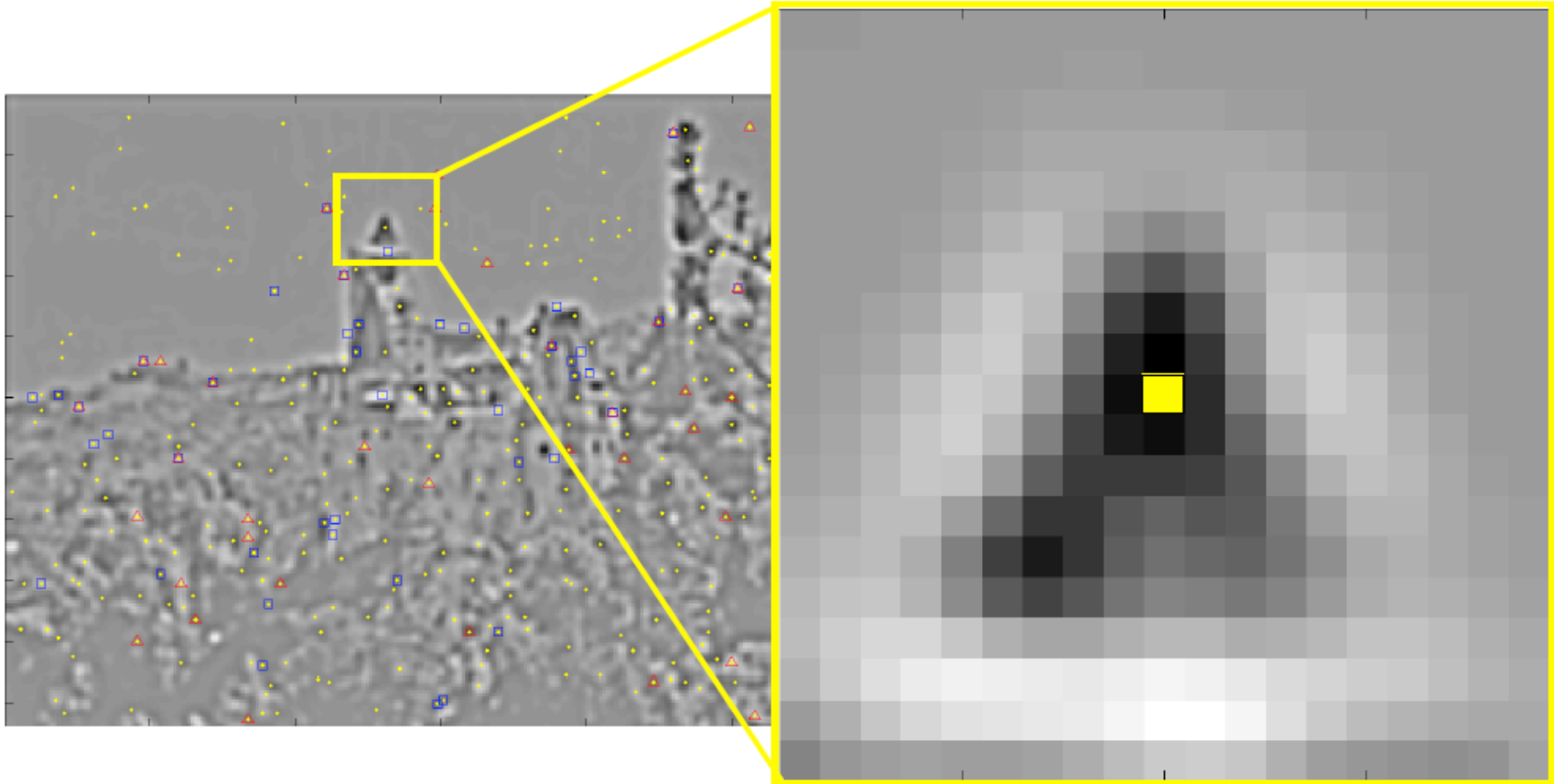
4. Számítsunk **gradiens irány-hisztogramokat** több kisebb ablakban, amiből leíró vektort képezünk.

Jellemző pontok kinyerése

- DoG skálatérben válasszunk ki minden pontot, ami a $3 \times 3 \times 3$ környezetben szélsőhely (max, vagy min)
- Töröljük az instabil pontokat
 - Alacsony kontraszt
 - Nem elég magas sarkossági jellemző
- Megkapjuk a képi pozíciót (x,y) és a hozzá tartozó skálát (σ)

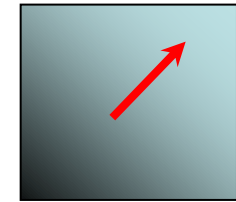
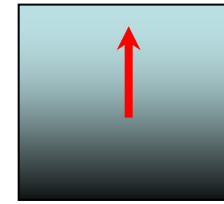


Kinyert jellemző pont

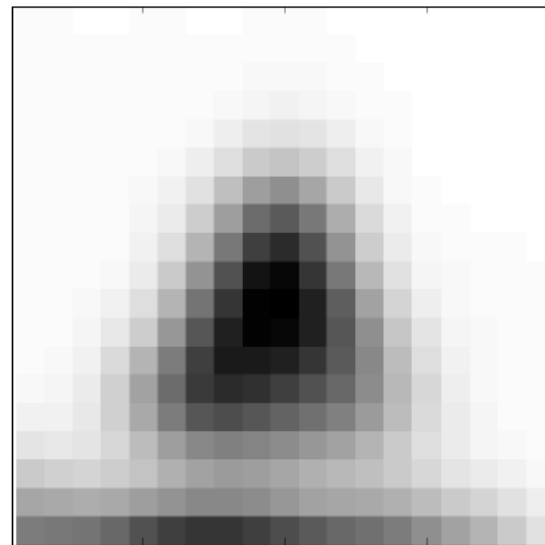


- **Keypoint location = extrema location**
- **Keypoint scale is scale of the DOG image**

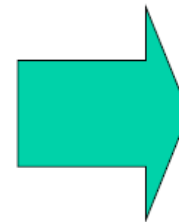
Orientáció



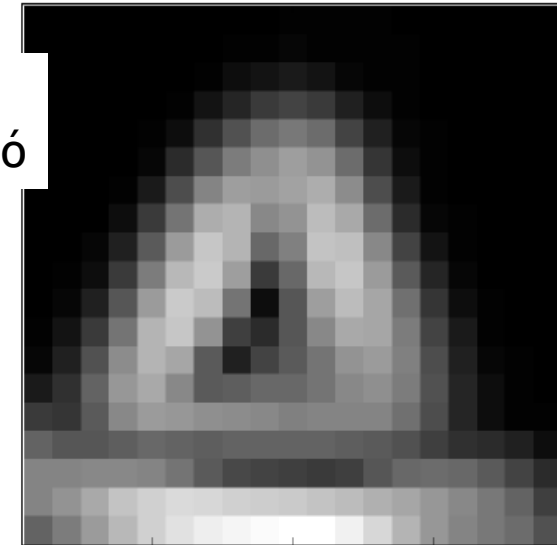
- A kinyert pontokhoz rendeljük egy konzisztens irányt
 - Minden ponthoz a hozzá tartozó skálán számoljunk gradienst



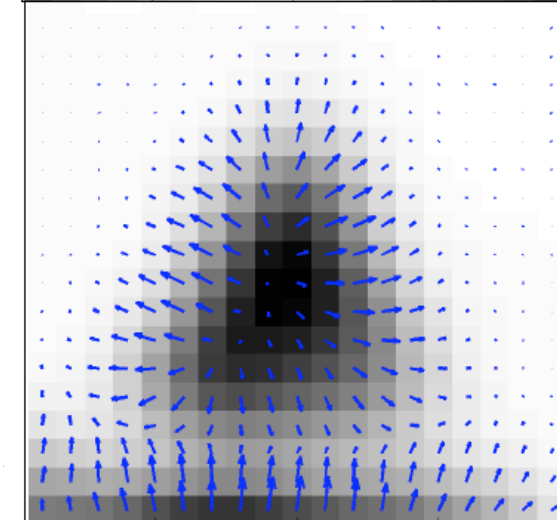
A képpiramis megfelelő skálájáról vett (Gauss simított) kép



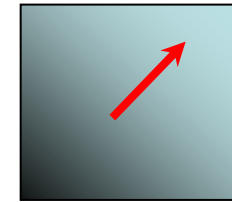
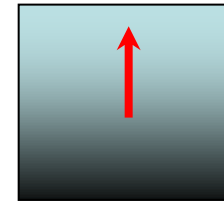
gradiens
magnitúdo



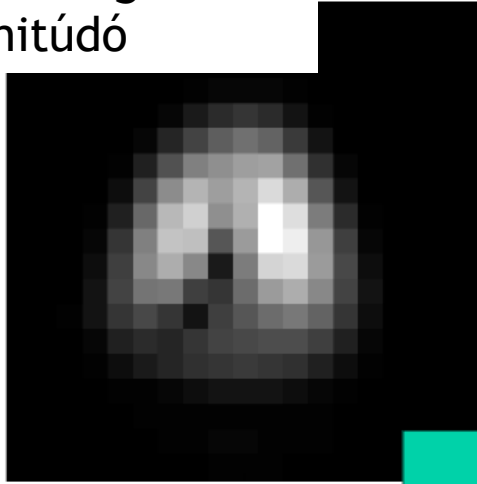
Gradiens
irány



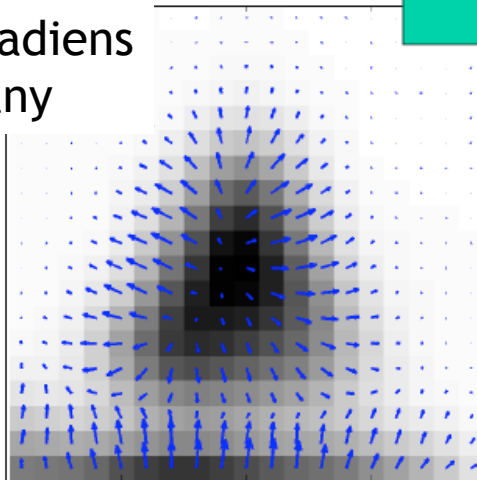
Orientáció



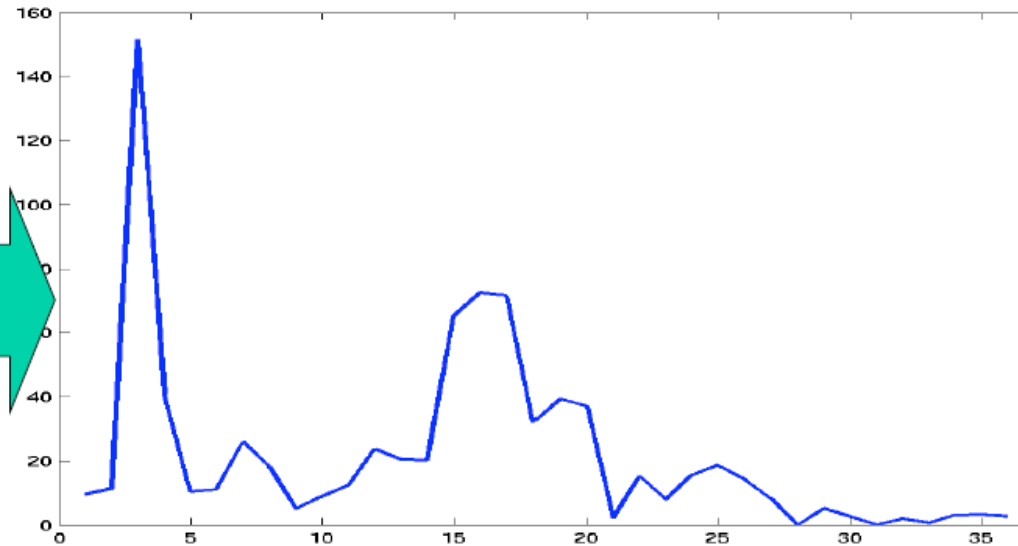
Súlyozott gradiens
magnitúdó



Gradiens
irány

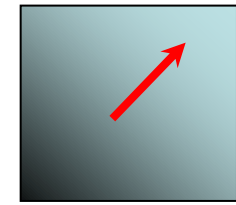
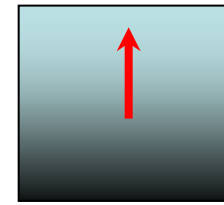


Súlyozott irány-hisztogram: Minden csoport azon gradiens-magnitúdók súlyozott összegét tartalmazza, amelyek a csoportba eső irányúak

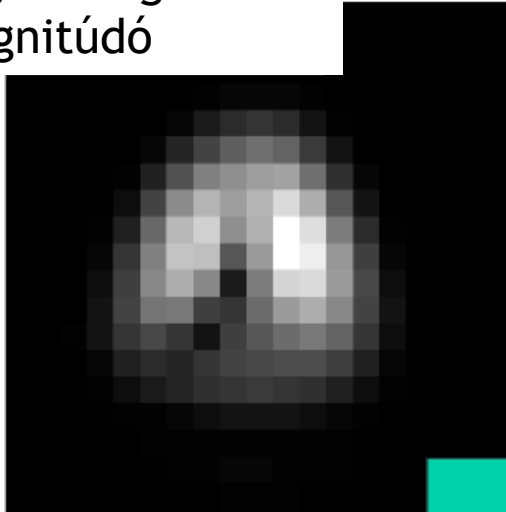


Irány csoportok: 10 fokonként tekintünk a gradiens vektorokat. Ez összesen 36 csoportot jelent.

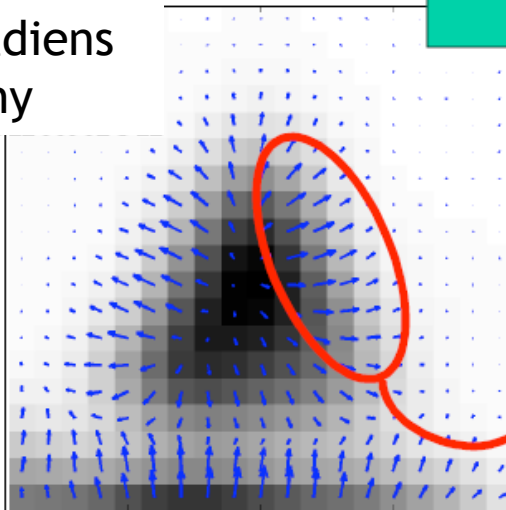
Orientáció



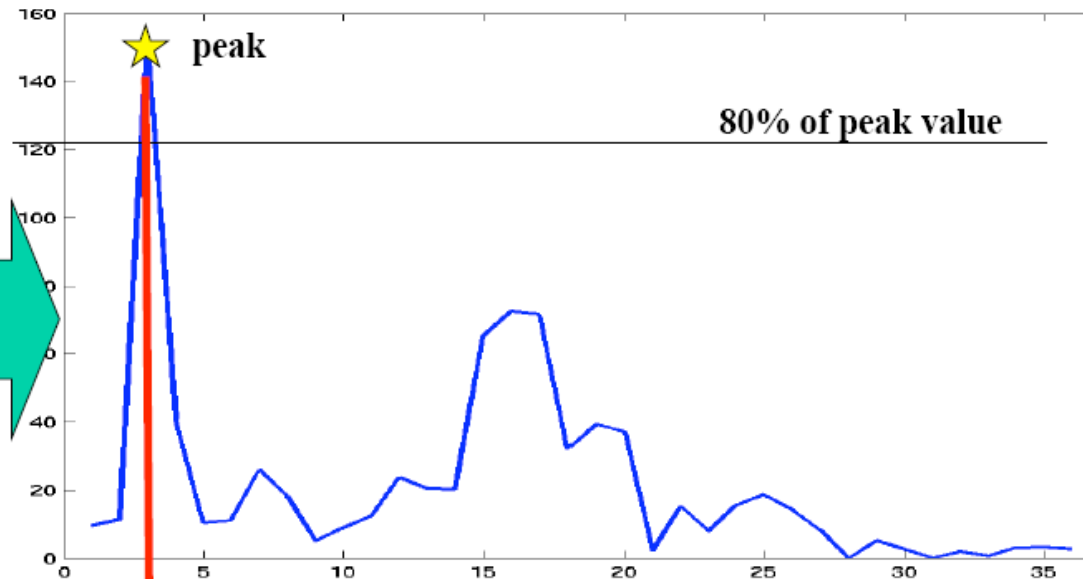
Súlyozott gradiens
magnitúdó



Gradiens
irány



Írány kiválasztás: Minden olyan irányt kiválasztunk, amelyek a csúcsérték 80%-n felül van és minden ilyen irányhoz létrehozunk egy pontot (tipikusan a pontok 15%-a több irányú)



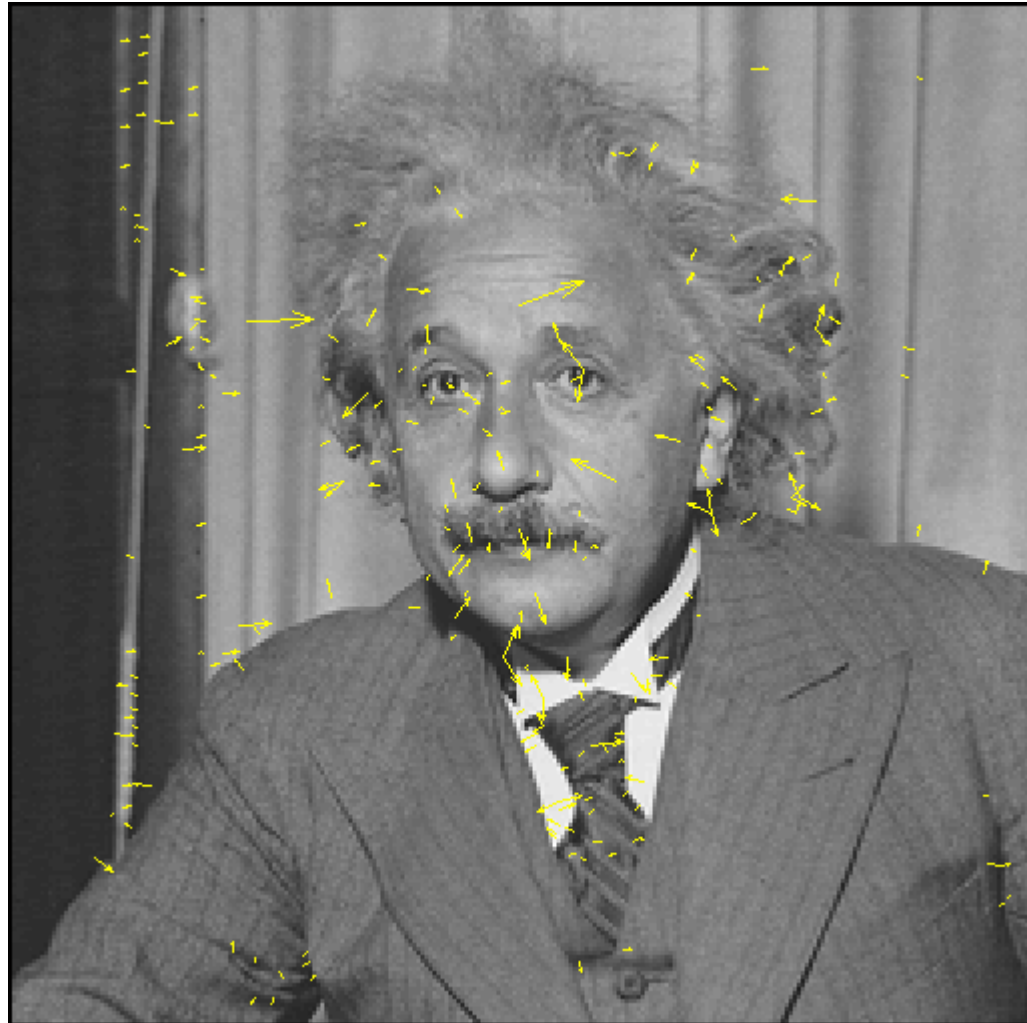
20-30 degrees

Jelen példában a pont iránya 25 fok lesz, mivel csúcs a 20-30 fokos csoportban van, és nincs másik domináns irány.

SIFT leíró

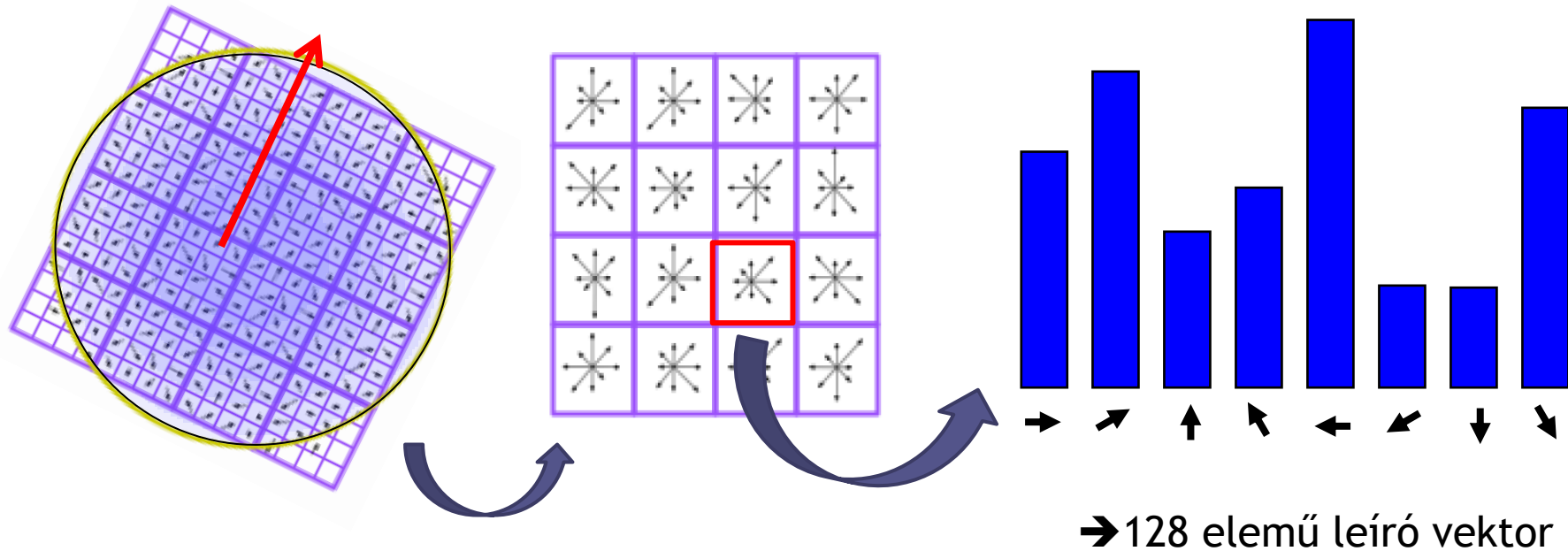
- Eddig minden jellemző ponthoz tartozik:
 - pozíció
 - skála
 - magnitúdó
 - irány

- Mivel írhatnánk le a pont körüli régiót?



SIFT leíró

- Tekintsünk egy 16x16 ablakot a gradiens-mezőn, amit tovább osztunk 4x4 blokkokra.
- Számítsunk a 4x4-es mintákon 8 irányban hisztogramot
- Alkalmazzunk Gauss súlyozást a középpont körül, aminek szórása= $0.5 \times \sigma$ (a jellemző pont skálája)
- $4 \times 4 \times 8 = 128$ dimenziós vektor leíró



SIFT leíró - fotometriai invariancia

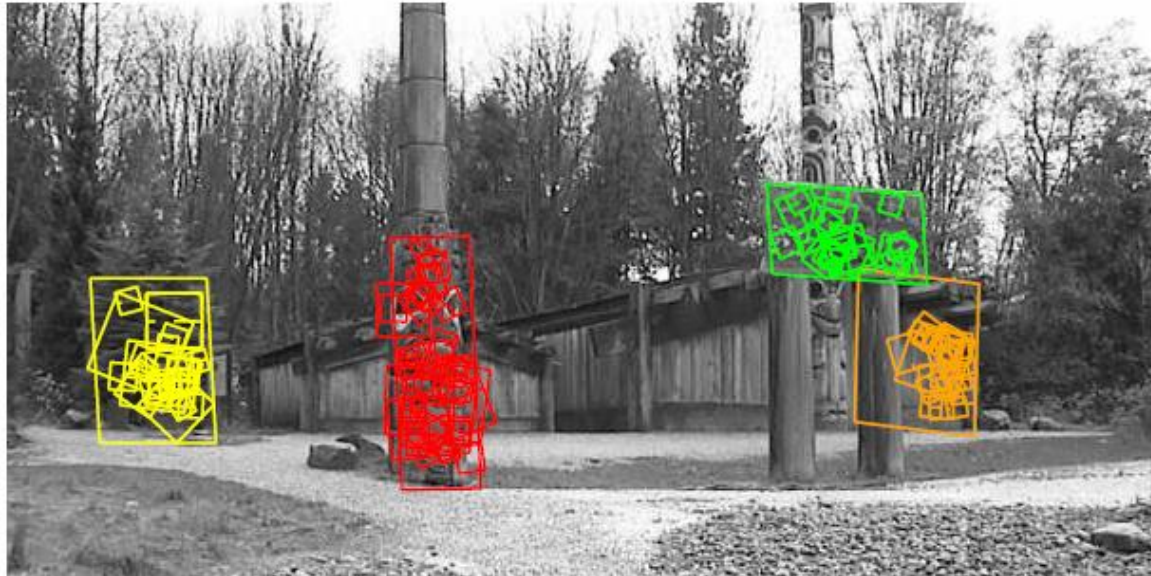
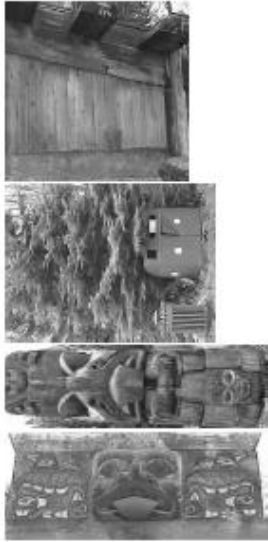
- **Lineáris (globális) megvilágítás változás:**
 - Az intenzitásértékek nagysága (gain) önmagában nem befolyásolja a gradienst
 - A vektorok egységnyivé normalizálása feloldja a kontrasztbeli különbségeket
- **Nem lineáris (lokális) megvilágítás változás:**
 - Telítettség inkább a magnitúdót mint az irányt befolyásolja.
 - Küszöböljük a gradiens magnitúdót 0.2-vel ($[0, 1]$ -re normált értékeket feltételezve), majd normalizáljunk újra.
 - empirikusan meghatározott érték 3D objektum többféle megvilágítása alapján
- **Az legnépszerűbb, rendkívül robusztus pontleíró [Mikolajczyk & Schmid 2005]**

SIFT leírók megfeleltetése

- L2 távolság alapján a legközelebbi szomszédal párosítjuk
- Hogyan szűrjük ki a rossz megfeleltetéseket (pl. ha nincs megfelelő párja egy pontnak)?
 - Küszöböljük az L2 távolságot → rossz teljesítmény
 - Küszöböljük inkább az arányt → jó teljesítmény

legjobb megfeleltetés
2. legjobb megfeleltetés

Képrészletek felismerése



Regisztráció panoráma képhez

onozó program)



Kató Zoltán: Digitális

[Brown & Lowe 2003]

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/>
- David Lowe
- James Hays: CS 143 Computer Vision, Brown University
➤ <http://www.cs.brown.edu/courses/cs143/>
- További források az egyes diákon megjelölve