

# Feature Detection: SURF and Matching

Moroz-Dubenco  
Cristiana

# Feature Detection and Matching

Parte importantă în aplicațiile de computer vision (image retrieval, object detection)

Caracteristică = informație relevantă pentru a soluționa o sarcină într-o aplicație

structuri specifice ale imaginii (puncte, colțuri, obiecte)

rezultatul unei operații aplicate imaginii

# Features

## Keypoint Features

- se află în locații specifice ale imaginii (vârful unui munte, colțurile unei clădiri)
- sunt descrise prin apariția de patch-uri de pixeli care înconjoară punctul

## Edges

- pot fi potrivite în funcție de orientare și aspect
- sunt buni indicatori ai limitelor obiectelor și evenimentelor de ocluzie dintr-o secvență de imagini

# Feature Detection and Matching

Detectare: identificarea punctului de interes

Descriere: descrierea aspectului local al fiecărui punct caracteristic într-o maniera independentă de schimbare luminii, rotației, translației și scalării  $\Rightarrow$  un vector descriptiv pentru fiecare caracteristică

Potrivire: compararea descriptorilor pentru identificarea caracteristicilor similare  $\Rightarrow$  un set de perechi  $(X_i, Y_i) \leftrightarrow (X_i^{\prime}, Y_i^{\prime})$ , unde  $(X_i, Y_i)$  este o caracteristică într-o imagine, iar  $(X_i^{\prime}, Y_i^{\prime})$  este caracteristica potrivită dintr-o altă imagine

# Feature Detection and Matching

Găsirea unei mulțimi de puncte de interes

Definirea unei regiuni în jurul fiecărui punct de interes

Extragerea și normalizarea conținutului regiunilor

Calcularea descriptorilor locali pe baza regiunilor normalizate

Potrivirea descriptorilor locali

# SURF

Algoritm rapid și robust pentru reprezentarea invariabilă a similarităților locale și compararea imaginilor

Inspirat din SIFT, mai rapid și mai robust în privința transformărilor imaginilor

# SURF

Detectarea punctelor de interes



Descrierea regiunilor locale



Potrivirea descriptorilor

Detectarea punctului  
de interes

Punctul în care direcția  
marginilor unui obiect se  
schimbă abrupt sau punctul  
de intersecție între două sau  
mai multe margini.

- Poziție bine definită în spațiul imaginii
- Stabil în raport cu perturbațiile locale sau globale ale imaginii
- Poate fi detectat într-un mod eficient



## Detectarea punctelor de interes

Filtrarea imaginii

Filtre în formă de pătrat (square-shaped) ca o aproximare a blur-ului gaussian

Imaginea integrală pentru rapiditatea filtrării:

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j)$$

Valoarea imaginii integrale într-un punct dat reprezintă suma tuturor pixelilor din imaginea de intrare într-o regiune dreptunghiulară determinată de origine și punct.

Suma intensităților în orice regiune dreptunghiulară se calculează rapid, evaluând doar vârfurile dreptunghiului.

## Detectarea punctelor de interes

Matricea Hessiană

Folosește determinantul matricei Hessiene pentru a măsura schimbările locale în jurul unui punct și alege punctele unde determinantul are valoare maximală.

Pentru un pixel dat, Hessian-ul lui este:

$$H(f(x, y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Detectarea  
punctelor de interes  
Matricea Hessiană

Pentru adaptabilitate la scalare,  
 imaginea este filtrată cu un kernel Gaussian,  
 rezultând, pentru un punct  $X = (x, y)$ ,  
 matricea Hessiană  $H(x, \sigma)$  la scală  $\sigma$ :

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{yx}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix}$$

$L_{xx}(x, \sigma) =$   
convoluția derivatei gaussiene de ordin  
secund cu imaginea  $I$  în punctul  $x$

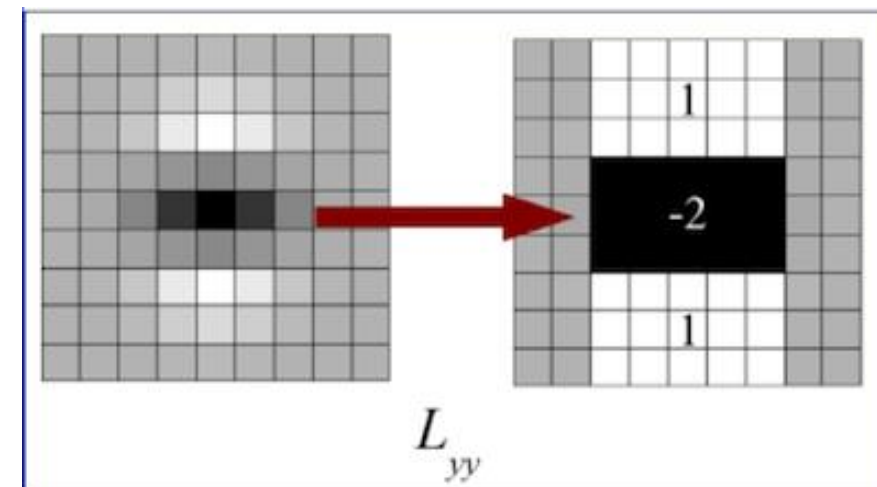
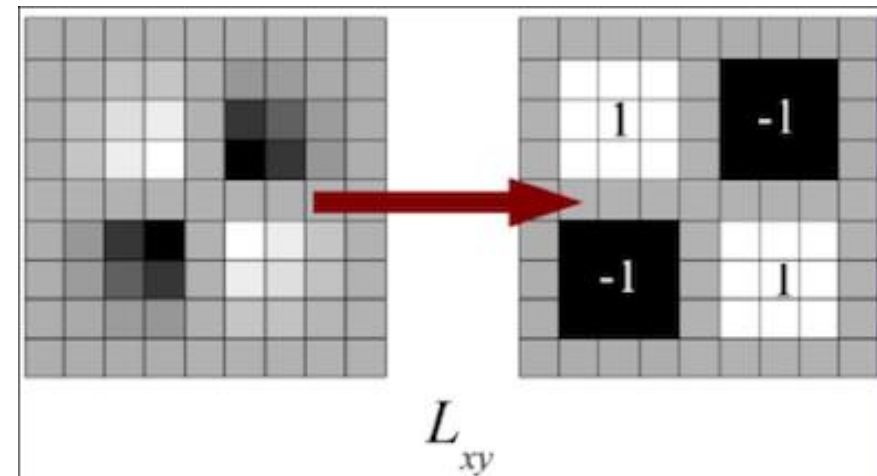
## Detectarea punctelor de interes

### Determinantul Hessian

- Pentru calculul determinantului, se aplică întâi convoluția cu kernel-ul gaussian, apoi se calculează derivatele de ordin doi.
- Convoluțiile și derivatele se aproximează folosind box filters.

$L_{xy}$  = derivata gaussiană parțială în  $xy$

$L_{yy}$  = derivata gaussiană parțială în  $y$



Detectarea punctelor  
de interes  
Determinantul Hessian

Filtrele dreptunghiulare de 9x9 folosite  
reprezintă aproximările  
pentru derivata gaussiană de ordin doi cu  $\sigma = 1.2$ .

$$\det(\mathcal{H}_{\text{approx}}) = D_{xx}D_{yy} - (wD_{xy})^2.$$

$D_{xx}$ ,  $D_{yy}$  și  $D_{xy}$  = aproximările derivatelor  
parțiale

## Detectarea punctelor de interes

Reprezentarea spațiului la scară

Punctele de interes se pot găsi la diferite scări, întrucât căutarea corespondențelor se face, de obicei, prin compararea de imagini unde punctele sunt văzute la scări diferite.

În alți algoritmi, se realizează o piramidă de imagini care sunt nivelate (smoothed) cu un filtru gaussian, apoi sunt sub-eșantionate pentru a atinge un nivel mai înalt al piramidei.

Detectarea punctelor  
de interes

Reprezentarea spațiului la  
scară

SURF nu aplică iterativ același filtru rezultatului stratului filtrat anterior, datorită folosirii de box filters.

Filtrele de orice dimensiune pot fi aplicate direct pe imaginea inițială, chiar și în paralel.

Reprezentarea spațiului la scară se face prin scalarea dimensiunii filtrului, în loc să se diminueze dimensiunea imaginii.

## Detectarea punctelor de interes

Reprezentarea spațiului la scară

Rezultatul filtrării cu filtrul =  $9 \times 9$  se considera scara inițială ( $s=1.2$ , corespunzător derivatelor gaussiene cu  $\sigma = 1.2$ ).

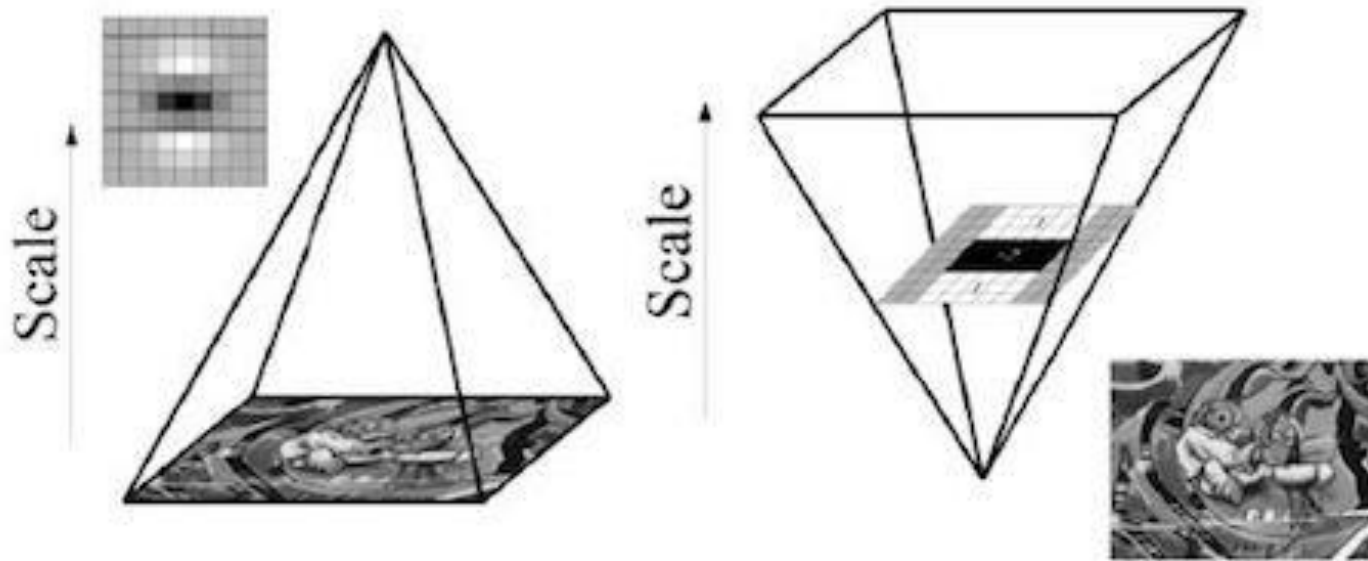
Următoarele straturi se obțin prin filtrarea imaginii cu măști gradul mai mari, ținând cont de natura discretă a imaginilor integrale și de structura specifică a filtrului ( $9 \times 9 \rightarrow 15 \times 15 \rightarrow 21 \times 21 \rightarrow 27 \times 27$  etc.).

Pentru a localiza punctele de interes, se aplică o suprimare non-maximă într-o regiune de  $3 \times 3 \times 3$ .



# Detectarea punctelor de interes

Reprezentarea spațiului la scară



În loc să se reducă iterativ dimensiunea imaginii (stânga), se scalează filtrele la un cost constant, folosind imaginile integrale (dreapta).

## Descrierea regiunilor locale

Descriptorul de caracteristici

Algoritm care primește o imagine și produce descriptori (vectori) de caracteristici

Înglobează informația într-o serie de numere

Se comportă ca o „amprentă” numerică ce poate fi folosită pentru a diferenția caracteristicile între ele

## Descrierea regiunilor locale

Descriptorul de caracteristici

### Descriptor local

- Reprezintă compact vecinătatea locală a unui punct
- Compară forma și aspectul doar în vecinătatea locală
- Potrivit pentru matching

### Descriptor global

- Descrie întreaga imagine
- Nu este robust (orice schimbare în imagine afectează descriptorul rezultat)

## Descrierea regiunilor locale

Descriptorul de caracteristici

Produce o descriere unică și robustă a unei caracteristici (de exemplu, prin descrierea intensității distribuției pixelilor în regiunea punctului de interes).

Majoritatea descriptorilor sunt calculați local, producând o descriere pentru fiecare punct de interes detectat anterior.

Dimensionalitatea descriptorului influențează direct complexitatea computațională și acuratețea potrivirii punctelor.

Un descriptor scurt poate fi mai robust în privința variațiilor de aspect, dar să nu ofere suficiente criterii de discriminare, ducând la multe rezultate fals-pozitive.

Descrierea regiunilor  
locale

Descriptorul de caracteristici

Fixarea unei orientări  
reproductibile pe baza informației  
dintr-o regiune circulară în jurul  
punctului de interes

Construirea unei regiuni pătrate  
aliniată orientării selectate și  
extragerea descriptorului din  
regiune

# Descrierea regiunilor locale

## Atribuirea orientării

Pentru a obține invarianță la rotație, este necesară detectarea unei orientări reproductibile a punctelor de interes, astfel:

Se calculează răspunsurile Haar-wavelet în direcțiile x și y într-o regiune de rază  $6s$  în jurul punctului de interes ( $s$  = scara la care a fost detectat punctul de interes).

Răspunsurile obținute sunt ponderate cu o funcție gaussiană centrată în punctul de interes, apoi sunt reprezentate ca puncte într-un spațiu bidimensional, cu răspunsul orizontal în abscisă și răspunsul vertical în ordonată.

Se estimează orientarea dominantă prin calcularea sumei tuturor răspunsurilor dintr-o fereastră de orientare glisantă cu dimensiunea  $\pi / 3$ . Se însumează răspunsurile orizontale și verticale din fereastră, producând un vector local de orientare.

Se alege cel mai lung astfel de vector, care definește orientarea punctului de interes.

# Descrierea regiunilor locale

## Extragerea descriptorului


Se construiește o regiune pătrată centrată în jurul punctului de interes și orientată de-a lungul orientării selectate anterior.



Se împarte regiunea de interes în sub-regiuni de 4x4.



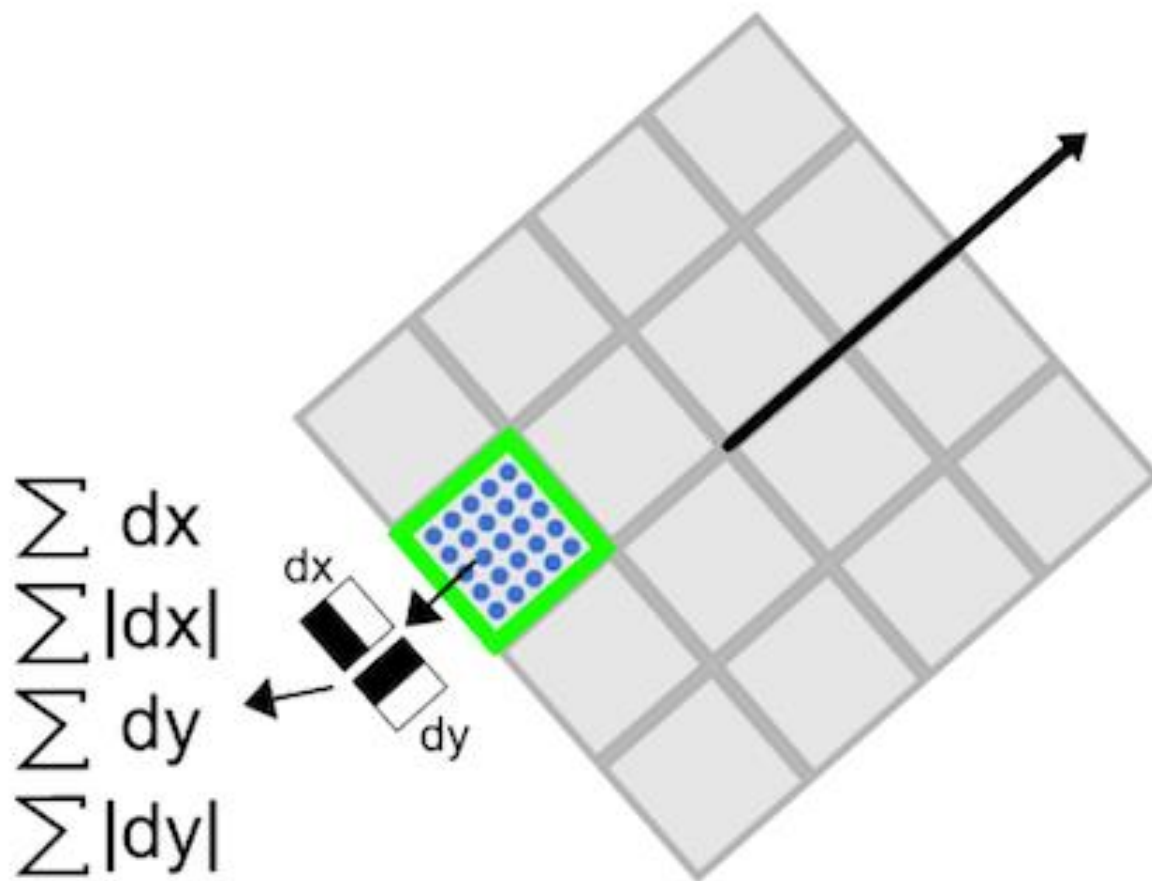
Se calculează răspunsurile Haar wavelet pentru puncte de probă distanțate la 5x5 pentru fiecare sub-regiune.



Se ponderează răspunsurile obținute cu un gaussian ( $\sigma = 3.3s$ ) centrat în punctul de interes, pentru a mări robustețea la deformații geometrice și erori de localizare.

# Descrierea regiunilor locale

## Extragerea descriptorului



$dx$  = răspunsul Haar wavelet în direcție orizontală

$dy$  = răspunsul Haar wavelet în direcție verticală



# Descrierea regiunilor locale

## Extragerea descriptorului

Se însumează răspunsurile  $dx$  și  $dy$  în fiecare sub-regiune, formând o primă mulțime de intrări pentru vectorul de caracteristici.

Se extrage suma valorilor absolute ale răspunsurilor,  $|dx|$  și  $|dy|$ , pentru a obține informații despre polaritatea modificărilor de intensitate.

Se obține un vector descriptiv patru-dimensional  $V = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$  pentru fiecare sub-regiune  $\Rightarrow$  un vector de lungime 64 pentru toate sub-regiunile de dimensiune  $4 \times 4$ .

# Potrivire

Găsirea corespondențelor între două imagini ale aceluiași obiect sau ale aceleiași scene

- După detectarea mulțimii punctelor de interes asociate cu descriptori din două sau mai multe imagini, se compară descriptorii și se stabilesc perechi potrivite între aceste imagini.
- Performanța metodelor de potrivire este dată de proprietățile punctelor de interes și de alegerea descriptorilor.