

Analisi Feature descriptors con SURF

Corso di Analisi Immagini e Video a.a. 2019/2020

Emilio Casella matr.204898

Concetti preliminari:

- keypoints: punti particolari dell'immagine o punti di interesse detti features locali.
Un insieme di keypoints può essere visto come una “patch” saliente dell'immagine, caratterizzata da un alto contenuto informativo locale della stessa. Il campionamento può essere casuale o a griglia regolare;
- descriptors: vettore numerico che identifica un determinato punto (keypoint) dell'oggetto univocamente, indipendentemente dalle condizioni di acquisizione dell'immagine;

SURF: Speeded Up Robust Features

- Nasce nel 2006 con lo scopo di migliorare SIFT;
- Può essere diviso in 3 fasi:
 1. Costruzione dell'immagine integrale;
 2. Ricerca keypoints nell'immagine tramite la matrice Hessiana;
 3. Creazione del descrittore.

1. Costruzione dell'immagine integrale

Data in ingresso un'immagine I e un punto di coordinate (x, y) , l'integrale dell'immagine $I_{\Sigma}(x, y)$ è calcolata tramite la somma dei valori dell'intensità dei pixel compresi tra il punto dato e l'origine. Formalmente è definito dalla formula:

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

Utilizzando questa tecnica, il calcolo si riduce alla sola esecuzione di semplici somme e alcuni accessi in memoria. Il grande vantaggio di tale metodo è che tale calcolo risulta, inoltre, essere indipendente dalle dimensioni dell'area considerata.

2. Ricerca keypoints nell'immagine tramite la matrice Hessiana

Il rivelatore dei keypoint è basato sulla *matrice Hessiana*, la quale viene approssimata mediante l'utilizzo delle immagini integrali, permettono ottime performance computazionali ed elevata accuratezza. Il detector così modificato prende il nome di **Fast-Hessian**.

Il SURF si serve di un filtro Gaussiano del secondo ordine normalizzato, che permette un'analisi spaziale e su diversi fattori di scala. Infatti, dato un punto $x = (x, y)$ in un'immagine I , la matrice Hessiana $\mathcal{H}(x, \sigma)$ alla scala σ è definita come:

$$\mathcal{H}(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

dove $L_{xx}(x, \sigma)$ è la convoluzione della Gaussiana derivata al secondo ordine $\frac{\partial^2}{\partial x^2}g(\sigma)$ con l'immagine I nel punto x , e similmente per $L_{xy}(x, \sigma)$ e $L_{yy}(x, \sigma)$.

2.Ricerca keypoints nell'immagine tramite la matrice Hessiana

PRO:

- Le funzioni Gaussianhe risultano essere ottime per le analisi di tipo spaziale (scale space);

CONTRO:

- Nella pratica vanno discretizzate;
- Perdita di reperibilità a $\pi/4$ per detector;
- Keypoints devono essere individuati a scale diverse.

2. Ricerca keypoints nell'immagine tramite la matrice Hessiana

- Spingendo l'approssimazione della matrice Hessiana all'estremo, si può utilizzare un box filter, calcolato ad un costo computazionale basso (immagini integrali) e performance paragonabili a quelle di una Gaussiana discretizzata e approssimata;

Per esempio, i box filter 9×9 in Figura 2.6 sono le approssimazioni di una Gaussiana con $\sigma = 1.2$ e rappresentano la scala più bassa per calcolare le risposte di convoluzione.

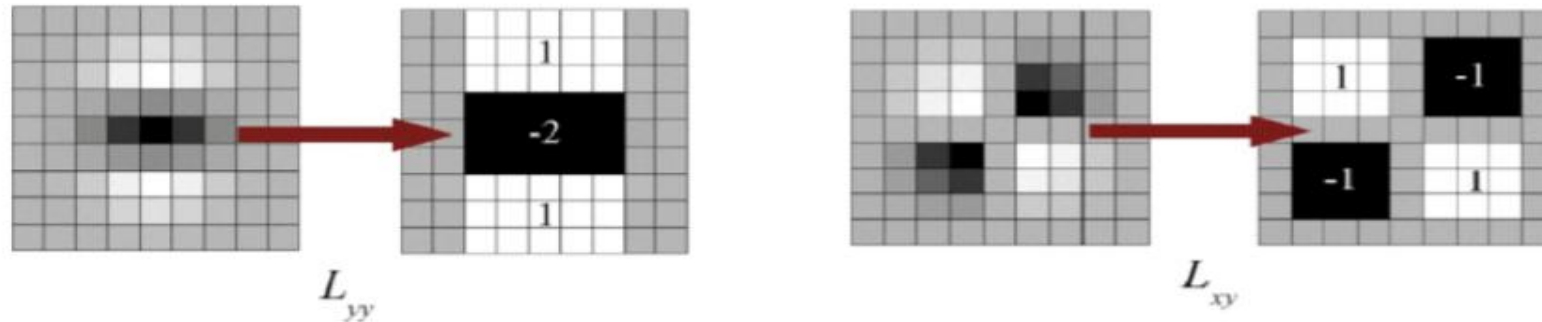


Figura 2.6: Rappresentazione dell'approssimazione apportata dal filtro box 9×9 [8]

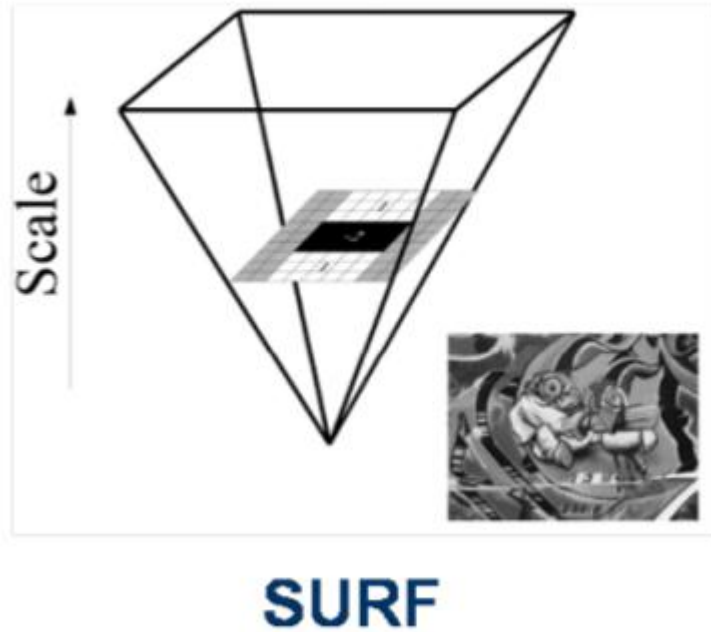
Indicando le approssimazione con D_{xx} , D_{xy} e D_{yy} e si calcola il determinante dell'Hessiana con:

$$\det(\mathcal{H}_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xy})^2$$

2.Ricerca keypoints nell'immagine tramite la matrice Hessiana

- Lo Scale-Space è generalmente implementato come una piramide di immagini, ottenuta applicando funzione di tipo Gaussiano alle diverse immagini;
- Grazie all'impiego dei Box-filter e all'integrale dell'immagine, non è più necessario applicare l'analisi Scale-Space; effettuando, quindi, con l'up-scaling del filtro piuttosto che il ridimensionamento iterativo dell'immagine filtrata;
- I vari layer della piramide sono ottenuti filtrando l'immagine con filtri man mano crescenti in dimensione.

2. Ricerca keypoints nell'immagine tramite la matrice Hessiana



- Per ogni scala i massimi del determinante dell'Hessiana approssimata vengono confrontati con i 26 punti (9x9 meno il punto centrale) connessi al punto candidato.;
- I massimi vengono poi interpolati con i punti vicini per trovare la vera posizione del massimo;
- Lo spazio delle scale viene diviso in ottave, ognuna delle quali rappresenta una serie di mappe di risposta, ottenute convolvendo la medesima immagine di input con un filtro di dimensione crescente, con un passo di almeno 2^n pixel; in modo da mantenere sempre un pixel centrale.

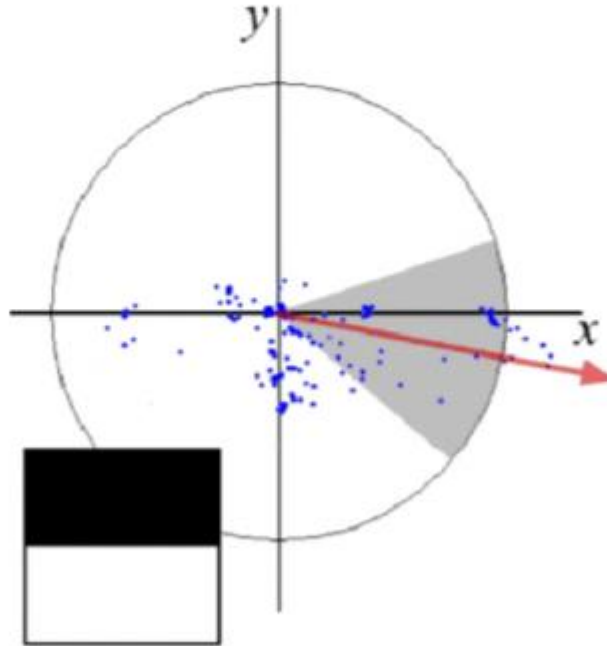
3.Creazione del descrittore

Passi:

1. **Fissare un orientamento principale valutando un area circolare nell'intorno del punto di interesse, in modo da raggiungere l'invarianza rispetto alla rotazione dell'immagine;**
 - a) calcolare la risposta di Haar wavelet nelle direzioni x e y in un intorno pari a 6 volte s attorno al punto di interesse (dove s rappresenta la scala alla quale il punto è identificato);
 - b) Si usano le immagini integrali per calcolare la risposta, poichè le wavelet di Haar sono dei Box filter;
 - c) Le risposte calcolate sono quindi pesate con una funzione gaussiana (con $\sigma = 2s$) centrata nel punto di interesse e successivamente rappresentate nello spazio.
 - d) La risposta orizzontale viene rappresentata sull'asse delle ascisse, mentre quella verticale sull'asse delle ordinate. La direzione dominante è stimata calcolando la somma di tutte le risposte con una finestra scorrevole, avente apertura $\pi/3$ radianti;assunta come orientamento locale;

3. Creazione del descrittore

c) Il vettore con modulo maggiore, tra tutti quelli calcolati con le diverse finestre, è l'orientamento principale del punto di interesse.



3.Creazione del descrittore

Passi:

2. Determinazione del descriptors;

- a) Si costruisce una griglia quadrata centrata sul keypoint e orientata lungo l'orientamento principale assegnato al passo precedente;
- b) La regione viene divisa in sotto regioni regolari di dimensioni 4x4 in modo da preservare un maggior numero di informazioni spaziali;
- c) Per ognuna di queste regioni viene calcolata la risposta di Haar Wavelet con una spaziatura regolare in direzione orizzontale (dx) e verticale (dy), in modo da incrementare la robustezza alle deformazioni geometriche;
- d) Le risposte dx e dy sono pesate con una funzione gaussiana, centrata nel keypoint;

3.Creazione del descrittore

- e) Le risposte dx e dy vengono sommate su tutta la sub-regione, formando il primo set di valori del vettore del descriptor;
- f) Per mantenere le informazioni in merito alla polarità delle variazioni di intensità, sono estratti anche i valori assoluti delle somme ($|dx|$) e ($|dy|$), fornendo un descrittore di 4 dimensioni:
- g) Concatenando tale descriptors, per tutte le sub-regioni di dimensione 4×4 considerate, si ottiene un vettore di lunghezza 64. L'invarianza rispetto al contrasto viene raggiunta normalizzando il vettore.

$$v = (\sum dx; \sum dy; \sum |dx|; \sum |dy|).$$



Remarks

- Esistono varie versioni che variano nella dimensione di creazione dei descriptors, ad esempio da 36 a 128 bit;
- Un esempio di utilizzo del modello descritto viene fornito nel notebook in allegato; con riferimento all'algoritmo SURF implementato nella libreria OpenCv;

Credits

- https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_surf_intro/py_surf_intro.html
- https://amslaurea.unibo.it/6855/1/Andrea_Annovi_tesi.pdf