

Caratteristiche dell'immagine

Francesco Isgr`o

17 ottobre 2023

Quali sono le caratteristiche dell'immagine?

una proprietà globale di un'immagine

livello medio di grigio

la zona

una parte dell'immagine con alcune proprietà speciali

un cerchio

una

linea una regione con

texture un bordo

Definizione

Le caratteristiche dell'immagine sono parti locali, significative e rilevabili dell'immagine

Significativo: la caratteristica viene associata ad elementi interessanti della scena attraverso il processo di formazione dell'immagine.

Rilevabile: deve esistere un algoritmo di localizzazione, altrimenti la funzionalità non è di alcuna utilità.

Caratteristica significativa

forti variazioni di intensità, create da regioni dei contorni
degli oggetti con livelli di grigio uniformi, immagini di superfici planari

Caratteristica significativa

forti variazioni di intensità, create da regioni dei contorni
degli oggetti con livelli di grigio uniformi, immagini di superfici planari

Tieni presente che...

la caratteristica può corrispondere o meno a parti della scena.

Rilevabile

diverse caratteristiche sono associate a diversi algoritmi di rilevamento. L'output
è un descrittore

di caratteristica: posizione più altre proprietà essenziali

Descrittore delle caratteristiche della linea: punto centrale, lunghezza, orientamento

A cosa servono le funzionalità dell'immagine?

L'estrazione delle funzionalità è un passaggio intermedio

Ad esempio, l'estrazione della linea può essere per:

muovere un robot in un corridoio

decidere se un'immagine contiene un determinato oggetto

eseguire la calibrazione della fotocamera

Non ha molto senso perseguire l'estrazione perfetta delle caratteristiche di per sé.

L'adeguatezza di un algoritmo di rilevamento delle caratteristiche dovrebbe essere valutata in definitiva nel contesto di un sistema completo.

I criteri generali di prestazione possono e devono essere applicati per testare i moduli di estrazione delle funzionalità in modo indipendente.

Rilevamento dei bordi

Definizione dei bordi

I punti dei bordi sono pixel in corrispondenza o attorno ai quali i valori dell'immagine subiscono una brusca variazione.

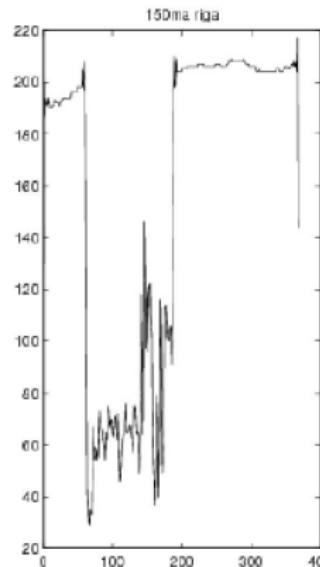
Il termine **bordo** è usato anche per riferirsi a catene collegate di punti di bordo: frammenti di contorno.

Rilevamento dei bordi

Definizione dei bordi

I punti dei bordi sono pixel in corrispondenza o attorno ai quali i valori dell'immagine subiscono una brusca variazione.

Il termine bordo è usato anche per riferirsi a catene collegate di punti di bordo: frammenti di contorno.



I bordi sono cambiamenti locali significativi nell'immagine. Sono punti dell'immagine in cui si riscontra un'elevata variazione dei valori di intensità.

Solitamente si verificano al confine tra due diverse regioni dell'immagine.

Non corrispondono necessariamente ai contorni reali, ma potrebbero essere causati da ombre, riflessi, ...

Come studiamo queste variazioni?

Gradiente dell'immagine

Una misura di questa variazione è il gradiente. Data una funzione $f(x, y)$, il suo gradiente è definito come

$$\vec{y}(f(x,y)) = \begin{matrix} J_x \\ J_y \end{matrix} = \begin{matrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{matrix}$$

grandezza del gradiente: $\sqrt{J_x^2 + J_y^2}$

Direzione del gradiente $J_y : \arctan \frac{J_y}{J_x}$



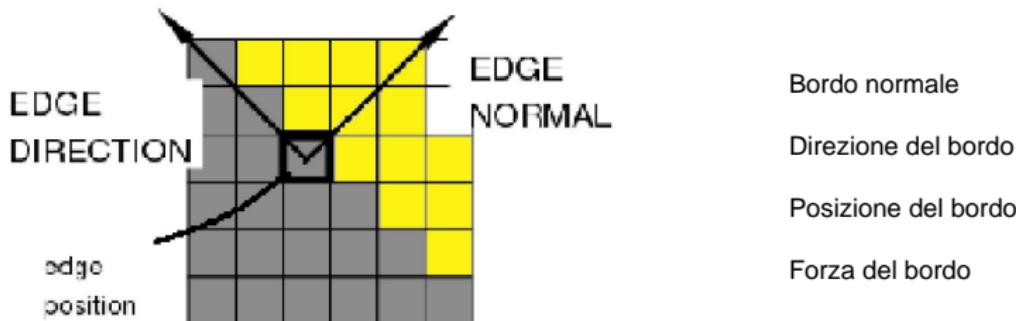
lx



ly



Descrittore essenziale del bordo



Rilevamento dei bordi incrociati di Roberts

funzione `RobertsEdgeDetect(I, Ÿ)`

Applicare lo smussamento del rumore (es. filtro gaussiano)
ottenendo l'immagine I_s

Filter I_s con le maschere

$$\begin{array}{cc} 10 & 01 \\ 0-1 & \ddot{\gamma}10 \end{array}$$

ottenendo le immagini I_x e I_y

Stimare l'intensità del gradiente $G(i, j)$

$$\sqrt{I_x(i, j)^2 + I_y(i, j)^2}$$

Prendi come bordi tutti i pixel tali che $G(i, j) > \theta$

=



Immagine originale

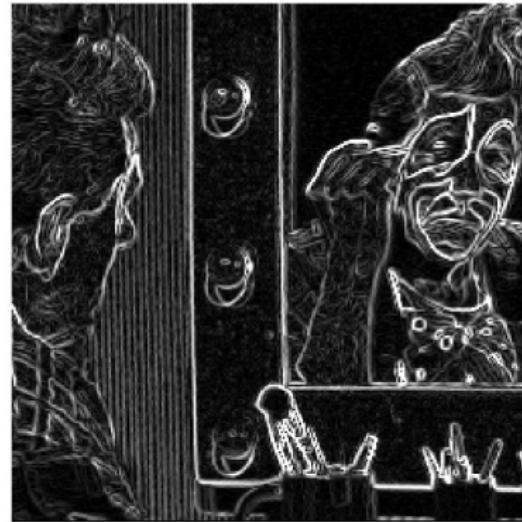


Immagine $G(i, j)$.



Immagine originale



$G(i, j)$ immagine $\bar{y} = 80$

Operatore molto veloce poiché il kernel è piccolo
L'immagine di output presenta punti luminosi spuri
L'operatore è suscettibile al rumore
Solo i bordi più forti sono stati rilevati con una certa affidabilità I valori
di output
da possono superare il valore massimo consentito dei pixel: una soluzione è
impostare i pixel in overflow al massimo valore.

Rilevamento dei bordi Sobel

funzione SobelEdgeDetect(I , \hat{y})

Applicare lo smussamento del rumore (es. filtro gaussiano)

ottenendo l'immagine I_s

Filter I_s con le maschere

$$\begin{array}{ccc} \hat{y}_1 & 0 & 1 & 1 & 2 & 1 \\ \hat{y}_2 & 0 & 2 & 0 & 0 & 0 \\ \hat{y}_y & \hat{y}_1 & 0 & 1 & \hat{y}_y & \hat{y}_1 & \hat{y}_2 & \hat{y}_1 & \hat{y}_y \end{array}$$

ottenendo le immagini I_x e I_y

Stimare l'intensità del gradiente $G(i, j)$

$$\sqrt{I_x(i, j)^2 + I_y(i, j)^2}$$

Prendi come bordi tutti i pixel tali che $G(i, j) > \hat{y}$

=



Immagine originale

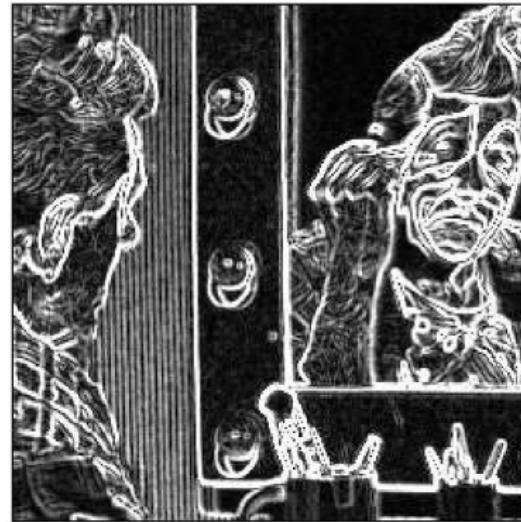
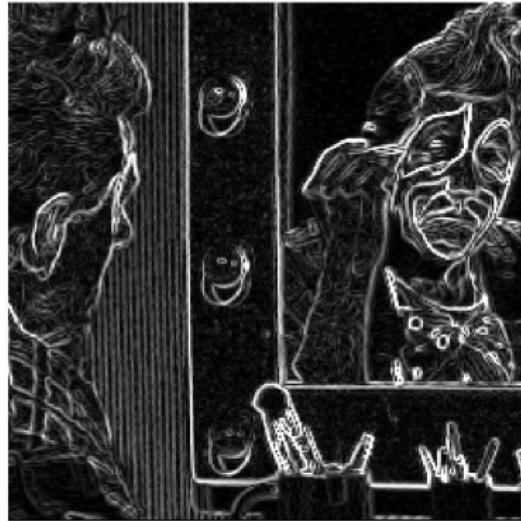
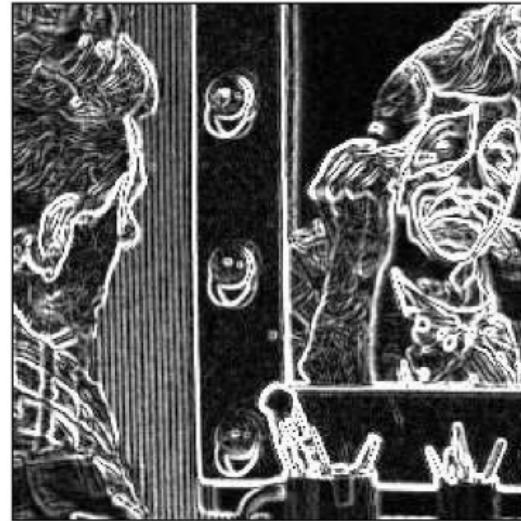


Immagine $G(i, j)$.



$G(i, j)$ Roberts



$G(i, j)$ Sobel

Per quanto riguarda l'operatore Roberts Cross

Il rumore spurio è meno intenso

Le linee sono più spesse

Eseguire il diradamento: rendere la posizione più precisa

Rilevamento dei bordi

Un tipico algoritmo per il rilevamento dei bordi contiene i seguenti passaggi:

Attenuazione del rumore: in assenza di informazioni utilizzare un filtro gaussiano

Miglioramento dei bordi: progetta un filtro che risponda ai bordi, in modo che i bordi siano massimi locali

Localizzazione dei bordi: decidere quali massimi locali rappresentano la soglia
del diradamento
dei bordi

Rilevatore di bordi di Canny

È il rilevatore di bordi più utilizzato.

È ottimale in un preciso senso matematico. Formula criteri di prestazione quantitativi. Formalizza le proprietà desiderabili per il rilevatore

Criteri per il rilevamento ottimale dei bordi

- Buon rilevamento** Il rilevatore ottimale deve ridurre al minimo la probabilità di falsi positivi e quella di mancare i bordi reali
- Buona localizzazione** I bordi rilevati devono essere il più vicino possibile ai bordi reali
- Vincolo di risposta singola** Il rilevatore deve restituire un punto per ogni punto del bordo vero

Rilevamento dei bordi e filtri lineari

Un ampio filtro lineare

migliora il rilevamento

peggiora la localizzazione

Compromesso tra rilevamento e localizzazione

Possiamo raggiungere un compromesso ottimale tra i criteri di localizzazione e di rilevamento regolando la scala del filtro, ma non possiamo migliorare entrambi i criteri contemporaneamente.

Algoritmo di rilevamento di Canny

1. Miglioramento del fronte
2. Soppressione dei non massimi
3. Soglia di isteresi

Miglioramento dei bordi

1. Applica lo smoothing gaussiano all'immagine I, ottenendo J = IG
2. Per ogni pixel (i, j)
 - Stimare le componenti del gradiente Jx e Jy
 - Calcolare la resistenza del bordo

$$E_s(i, j) = J_2 \times (ij) + J_2(i, j) \text{ si}$$

Stimare l'orientamento della normale al bordo

$$E_o(i, j) = \arctan \frac{J_y}{J_x}$$

Soppressione dei non massimi

1. Per ciascuna (i, j) e per un dato insieme D di direzioni:

trovare la direzione d^* in D che meglio approssima $E_0(i, j)$;
se $E_S(i, j)$ è più piccolo dei suoi due vicini lungo la
direzione d^* allora imposta $E_S(i,$

j) a zero 2. un tipico insieme D include $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$

Soglia di isteresi

Dati due valori di soglia $\hat{y}_1 > \hat{y}_2$, scansionare Es e 1.

se $Es(i, j) > \hat{y}_1$ allora 2.

seguire i massimi locali connessi, in entrambe le direzioni

perpendicolare alla normale al bordo, purché $Es(h, k) > \hat{y}_2$;
contrassegnare i punti visitati e salvare un elenco delle posizioni
di tutti i punti visitati.



Immagine originale



$\bar{y} = 1, \bar{y}_1 = 255, \bar{y}_2 = 1$



Immagine originale



$\bar{y} = 1$, $\bar{y}_1 = 255$, $\bar{y}_2 = 220$



Immagine originale



$\bar{y} = 1$, $\bar{y}_1 = 128$, $\bar{y}_2 = 1$



Immagine originale



$\bar{y} = 2, \bar{y}_1 = 128, \bar{y}_2 = 1$

Precisione sub-pixel

Tutti i rilevatori di bordi descritti identificano il pixel del bordo

Il centro potrebbe trovarsi ovunque all'interno del pixel

In alcune applicazioni mezzo pixel può corrispondere ad un errore elevato

Potrebbe essere importante individuare il bordo con una precisione inferiore al pixel

Il modo più semplice è interpolare una parabola nell'immagine migliorata

Per ciascun pixel del bordo p nell'immagine migliorata

1. Considera il bordo normale \hat{y}_p
- 2.

Identifica i due pixel vicini, \hat{y}_{p+1} e \hat{y}_{p-1}

lungo \hat{y}_p

3. Interpolare una parabola utilizzando i tre valori in
l'immagine migliorata
4. Trova il vertice della parabola

Valutazione dei rilevatori di bordi

Rilevatori di bordi specifici possono essere

valutati:

teoricamente sperimentalmente, stimando gli indici di prestazione su
immagini

sintetiche probabilità di falsi

rilevamenti: probabilità

di mancamenti: la media e gli errori RMS delle posizioni dei bordi

Valutazione dei rilevatori di bordi

Rilevatori di bordi specifici possono essere

valutati:

teoricamente sperimentalmente, stimando gli indici di prestazione su
immagini

sintetiche probabilità di rilevamento falso: il numero di bordi
spuri probabilità di

mancamenti: la media e gli errori RMS delle posizioni dei bordi

Valutazione dei rilevatori di bordi

Rilevatori di bordi specifici possono essere valutati:
teoricamente sperimentalmente, stimando gli indici di prestazione su immagini sintetiche probabilità di rilevamento falso: il numero di bordi spuri probabilità di mancamenti: numero di bordi veri persi la media e gli errori RMS delle posizioni dei bordi

Rilevamento di linee e curve

Le immagini spesso contengono oggetti creati dall'uomo

I contorni di questi oggetti sono molto regolari

ellisse

del cerchio

della linea

Rilevamento della linea: corrispondenza del modello

Cerca i picchi della convoluzione tra l'immagine e una serie di maschere My sintonizzate per rilevare la presenza di linee di una particolare larghezza e con un particolare orientamento y.

Un punto p è su una linea di orientamento y

se La risposta di My è la più alta tra tutte le maschere La
risposta di My è superiore ad una determinata soglia

a)

-1	-1	-1
2	2	2
-1	-1	-1

b)

-1	2	-1
-1	2	-1
-1	2	-1

c)

-1	-1	2
-1	2	-1
2	-1	-1

d)

2	-1	-1
-1	2	-1
-1	-1	2

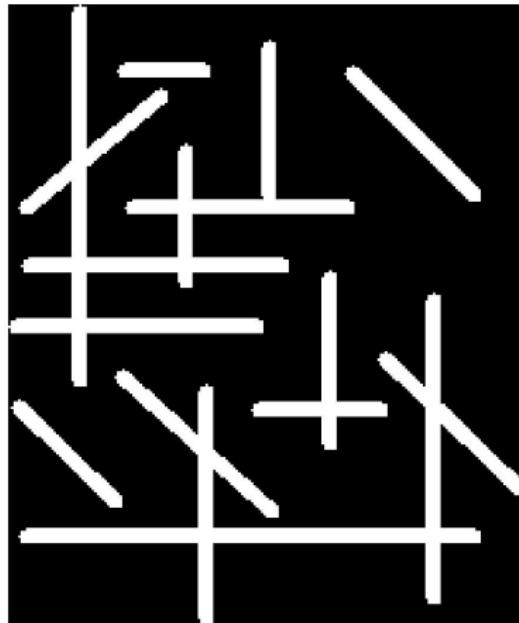


Immagine in ingresso

Filtra la risposta

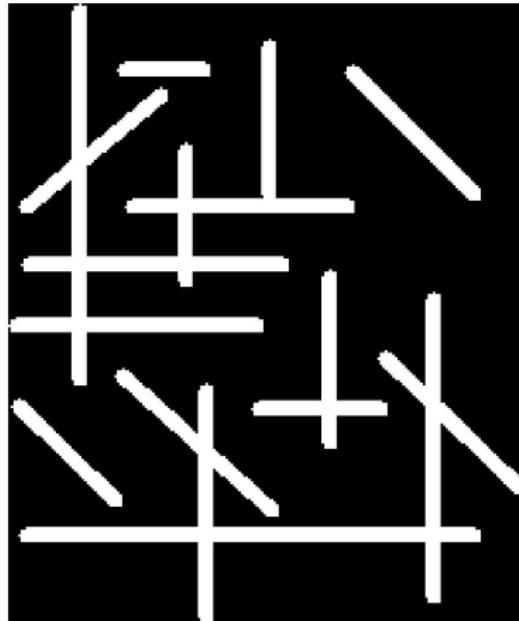
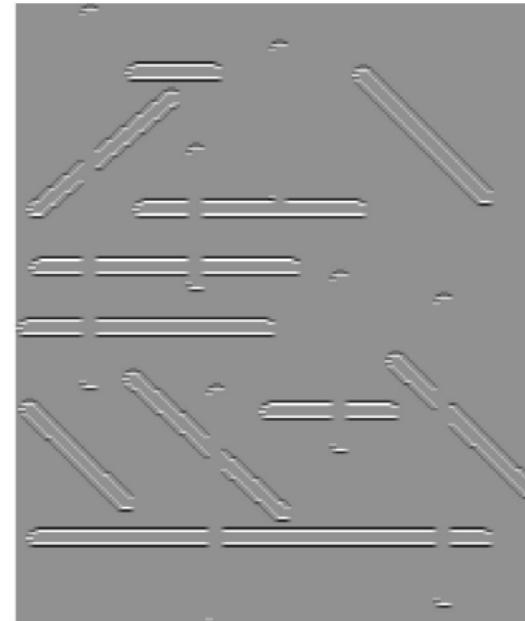


Immagine in ingresso



Filtra la risposta

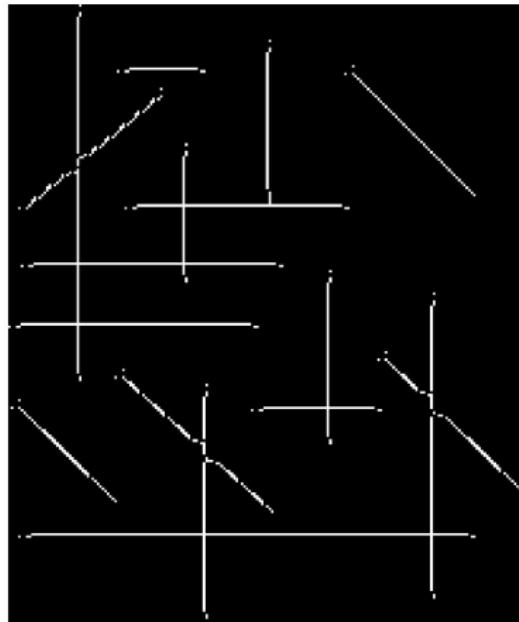


Immagine di input scheletrata

Filtra la risposta

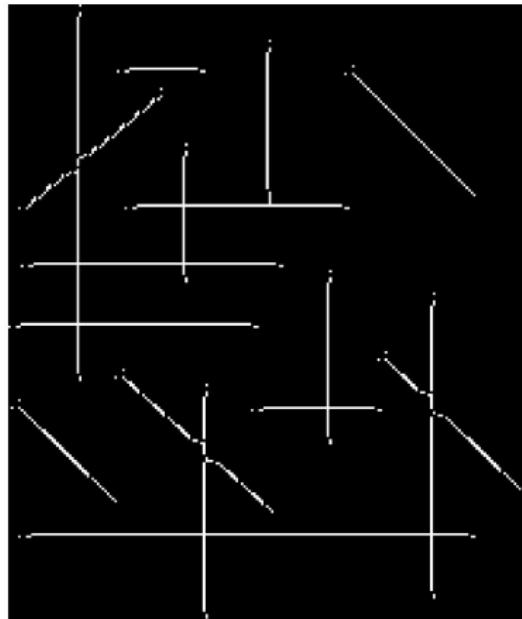
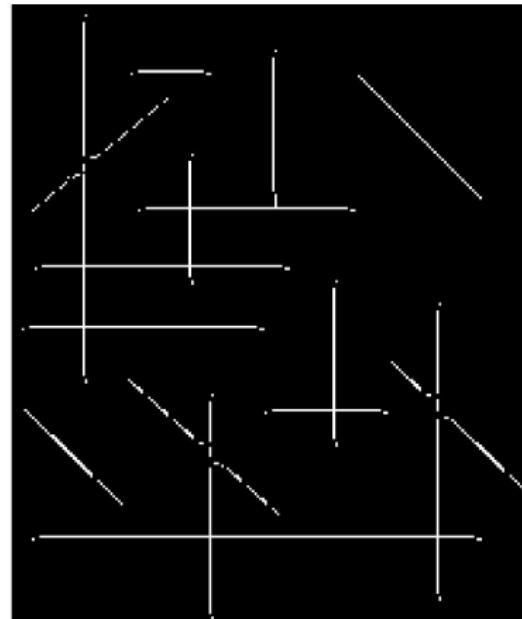


Immagine di input scheletrata



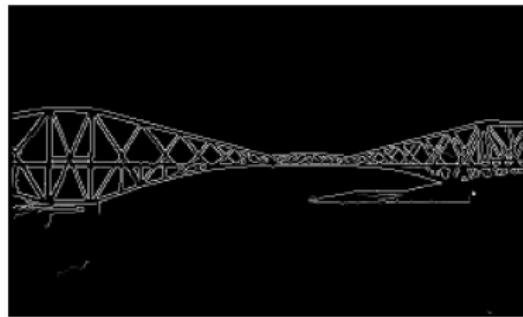
Filtra la risposta



Immagine in ingresso



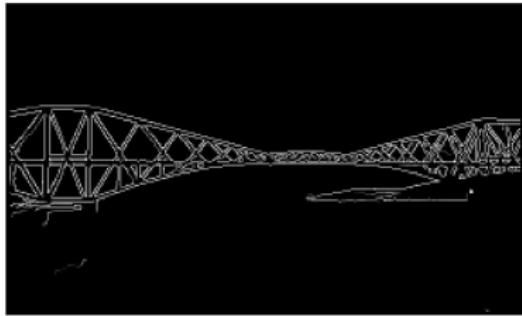
Immagine di uscita



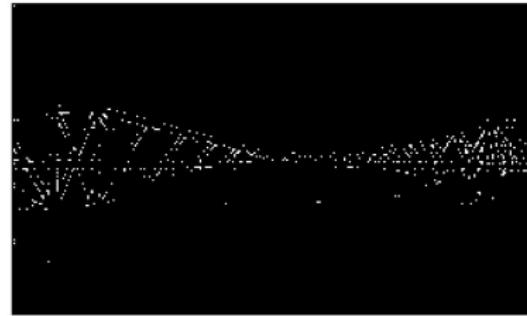
Uscita di Canny



Immagine in uscita



Uscita di Canny



Filtro gaussiano + rilevatore di linea

rilevamento accurato della linea

richiede un ampio set di maschere
e l'intero processo è computazionalmente costoso

se le linee sono più larghe delle maschere

bordi di ritorno delle
linee risultati errati della scheletratura di
due linee parallele riducono la precisione

Rilevamento di linee e curve

Le immagini spesso contengono oggetti creati dall'uomo

I contorni di questi oggetti sono molto regolari

ellisse

del cerchio

della linea

Rilevamento di linee e curve

Le immagini spesso contengono oggetti creati dall'uomo

I contorni di questi oggetti sono molto regolari

ellisse

del cerchio

della linea

Dichiarazione problema

Dato l'output di un rilevatore di bordi, trova tutte le istanze di una data curva o parti di essa.

La trasformata di Hough

È una tecnica utilizzata per isolare le caratteristiche di una forma particolare all'interno di un'immagine.

Richiede che le caratteristiche desiderate siano specificate in qualche forma parametrica

Mappare un difficile problema di rilevamento di pattern in un semplice problema di rilevamento di picchi nello spazio dei parametri della curva.

La trasformata di Hough per le linee

1. Trasformare il rilevamento delle linee in un problema di intersezione di linee

una linea $y = mx + n$ è identificata da (m, n) un

punto $p = (x, y)$ corrisponde alla linea $n = (\ddot{y}x)m + y$ nel parametro spazio

Una linea identificata da

p_1, \dots, p_n è identificata dall'intersezione di n linee.

2. Trasformare l'intersezione delle linee in un problema di rilevamento dei picchi

La trasformata di Hough per le linee

1. Trasformare il rilevamento delle linee in un problema di intersezione di linee

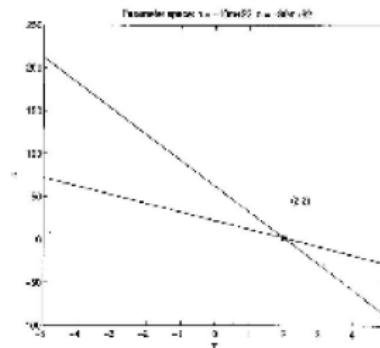
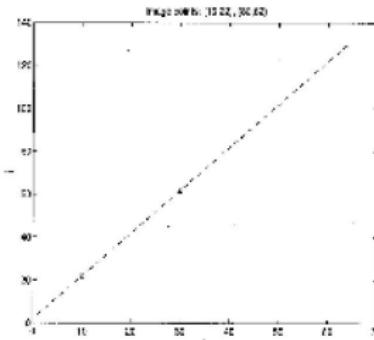
una linea $y = mx + n$ è identificata da (m, n) un

punto $p = (x, y)$ corrisponde alla linea $n = (\hat{y}x)m + y$ nel parametro spazio

Una linea identificata da

p_1, \dots, p_n è identificata dall'intersezione di n linee.

2. Trasformare l'intersezione delle linee in un problema di rilevamento dei picchi



La trasformata di Hough per le linee

1. Trasformare il rilevamento delle linee in un problema di intersezione di linee

una linea $y = mx + n$ è identificata da (m, n) un

punto $p = (x, y)$ corrisponde alla linea $n = (\ddot{y}x)m + y$ nel parametro spazio

Una linea identificata da

p_1, \dots, p_n è identificata dall'intersezione di n linee.

2. Trasformare l'intersezione delle linee in un problema di rilevamento dei picchi

Dividi lo spazio dei parametri in una griglia finita di celle

Associa un contatore ad ogni cella

Incrementa il contatore per ogni linea che passa attraverso la cella

Determinare il picco

Sia m che n variano in $[0, \infty]$

Non possiamo campionare l'intero spazio dei parametri

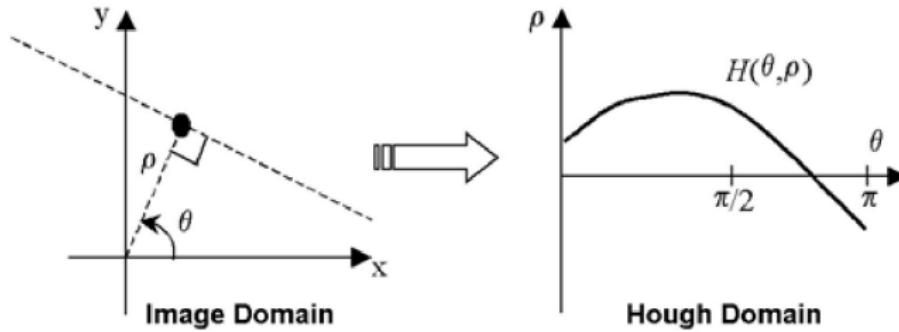
Non possiamo catturare $x = k$

Possiamo usare la rappresentazione polare

$$= x \cos \theta + y \sin \theta$$

intervalli o variazione per θ e r sono finiti $[0, \pi]$

+ $[0, N]$, dove $M \times N$ è la dimensione dell'immagine $[0, \pi]$



Contorni non lineari

Non tutti i punti di bordo appartengono a linee

Diffusione di valori casuali nello spazio dei parametri

Moltitudine di picchi rumorosi

Problema: discriminare tra picchi rumorosi e picchi buoni

Soglia i contatori

Rumore

Pixelizzazione e precisione limitata nel rilevamento dei bordi

non tutte le linee nello spazio dei parametri (punti della stessa linea) si intersecano nello stesso punto

Determinare il picco come massimo locale

Determinare il picco come media ponderata in un intorno dei massimi locali

Algoritmo

funzione [II] = HoughLines(I , θ , ρ , M)

$MH = \sum M_2 + N_2/\rho$, $NH = \sum \rho/\theta$ $A =$

zeri(MH , NH) per ogni

(i, j) tale che $I(i, j) = 1$ per $h=0 \dots NH$ $\theta =$
round($i \cos \theta(h)$

$+ j \sin \theta(h)$

$A(\theta, h) += 1$

Trovare tutti i massimi locali in A , tali che $A(\theta, h) > \theta$



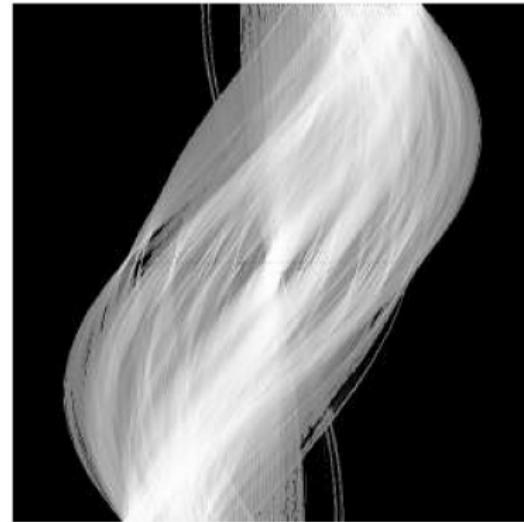
Immagine in ingresso



Uscita di Canny



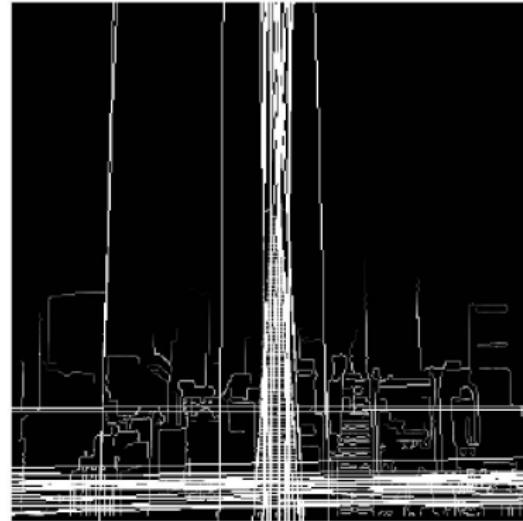
Immagine in ingresso



Hough trasforma



Immagine in ingresso



Linea rilevata

Trasformata di Hough per i cerchi

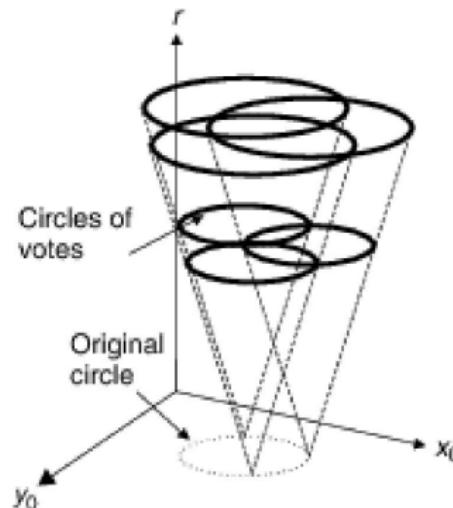
L'equazione di una circonferenza è

$$(x - x_0)^2 + (y - y_0)^2 = r^2$$

Trasformata di Hough per i cerchi

L'equazione di una circonferenza è

$$(x - x_0)^2 + (y - y_0)^2 = r^2$$



(c) 3D accumulator space

L'equazione della circonferenza può essere scritta in forma parametrica come:

$$x = x_0 + r \cos \theta$$

$$y = y_0 + r \sin \theta$$

L'equazione della circonferenza può essere scritta in forma parametrica come:

$$x = x_0 + r \cos \theta$$

$$y = y_0 + r \sin \theta$$

Perciò

$$x_0 = x - r \cos \theta$$

$$y_0 = y - r \sin \theta$$

x_0 e y_0 dipendono dal raggio r .

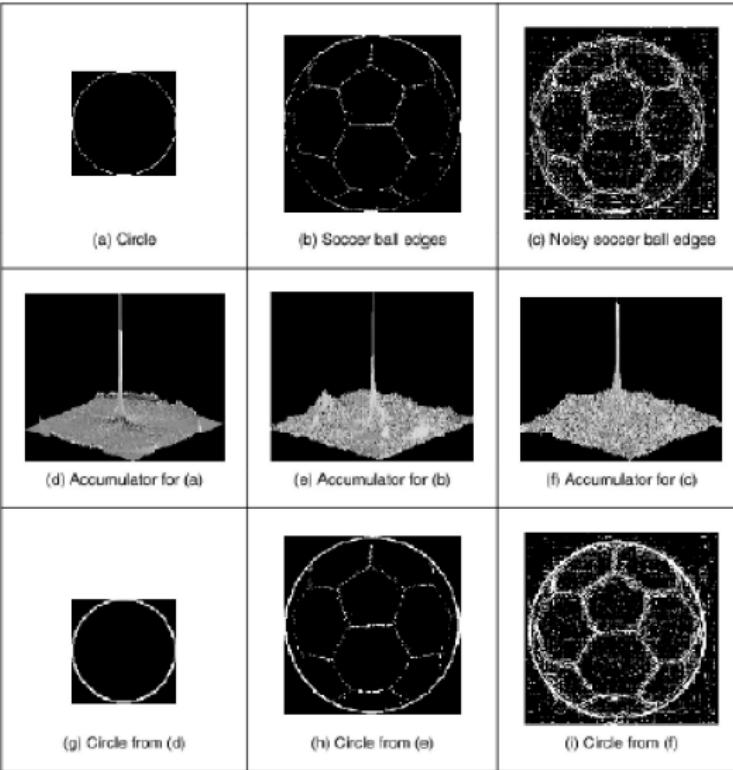
```
%Hough Transform for Circles

function HTCircle(inputimage,r)

%image size
[rows,columns]=size(inputimage);

%accumulator
acc=zeros(rows,columns);

%image
for x=1:columns
    for y=1:rows
        if(inputimage(y,x)==0)
            for ang=0:360
                t=(ang*pi)/180;
                x0=round(x-r*cos(t));
                y0=round(y-r*sin(t));
                if(x0<columns & x0>0 & y0<rows & y0>0)
                    acc(y0,x0)=acc(y0,x0)+1;
                end
            end
        end
    end
end
```





La trasformata di Hough per le curve

L'algoritmo può essere generalizzato a curve generiche

$$\begin{aligned}y &= f(x, a) \\a &= [a_1, \dots, A]^T\end{aligned}$$

Lo spazio dei parametri cresce esponenzialmente con T

Il tempo di calcolo può essere troppo elevato

Soluzione parziale

Multirisoluzione

Aumentare la risoluzione solo dove necessario, ovvero dove sono presenti massimi locali

La trasformata di Hough: punti chiave

L'HT è un algoritmo di votazione I contatori

nello spazio dei parametri possono essere considerati come un istogramma Un
contatore di

coordinate m indica la probabilità relativa dell'ipotesi che esista una curva m nell'immagine

Può far fronte alle occlusioni: i punti vengono elaborati

indipendentemente È robusto da il rumore poiché è improbabile che i
valori anomali

contribuiscano a un singolo contenitore Limitazioni principali: il tempo di ricerca aumenta
con il numero

di parametri le forme non

target possono produrre picchi spuri (ad esempio, il rilevamento
della linea può essere disturbato da cerchi a bassa curvatura).

Caratteristiche puntuali

Due approcci principali

I picchi di curvatura locale sono punti caratteristici

Estrazione degli angoli

Basta con piccoli cambiamenti tra le immagini

Inizia a fallire quando una forte trasformazione affine

Utilizzo dell'analisi basata su regione o patch

Alla ricerca dell'invarianza affine

Può far fronte ai più grandi cambiamenti tra le immagini

Definizione di curvatura

Curvatura

Intuitivamente si tratta del cambiamento della direzione del bordo. I punti in cui cambia rapidamente sono gli angoli.

Una curva in forma parametrica è ($Ux = [1, 0]t$, $Uy = [0, 1]t$)

$$v(t) = x(t)Ux + y(t)Uy$$

Le variazioni del vettore posizione sono date dal vettore tangente

$$\frac{dv(t)}{dt} = \frac{dx(t)}{dt}Ux + \frac{dy(t)}{dt}Uy = \ddot{x}(t)Ux + \ddot{y}(t)Uy$$

Un punto si muove con velocità

$$|v\ddot{y}(t)| = \sqrt{\dot{x}^2(t) + \dot{y}^2(t)}$$

nella direzione

$$\ddot{y}(t) = \tan \ddot{y}_1 \frac{y\ddot{y}(t)}{x\ddot{y}(t)}$$

Curvatura

La curvatura in un punto $v(t)$ descrive il cambiamento nella direzione $\ddot{y}(t)$ rispetto ai cambiamenti nella lunghezza dell'arco s

$$\ddot{y}(t) = ds \frac{d\ddot{y}(t)}{ds}$$

Curvatura

La curvatura in un punto $v(t)$ descrive il cambiamento nella direzione $\ddot{y}(t)$ rispetto ai cambiamenti nella lunghezza dell'arco s

$$\ddot{y}(t) = ds \frac{d\ddot{y}(t)}{ds}$$

$$\ddot{y}(t) = dt \frac{d\ddot{y}(t)}{ds} \frac{dt}{ds}$$

Curvatura

La curvatura in un punto $v(t)$ descrive il cambiamento nella direzione $\ddot{y}(t)$ rispetto ai cambiamenti nella lunghezza dell'arco s

$$\ddot{y}(t) = ds \frac{d\ddot{y}(t)}{dt}$$

$$\ddot{y}(t) = dt \frac{d\ddot{y}(t)}{ds} \frac{dt}{ds}$$

$$\frac{ds}{dt} = |v\ddot{y}(t)| = x\ddot{y}(t) \sqrt{\dot{x}^2 + \dot{y}^2}$$

Curvatura

La curvatura in un punto $v(t)$ descrive il cambiamento nella direzione $\ddot{y}(t)$ rispetto ai cambiamenti nella lunghezza dell'arco s

$$\ddot{y}(t) = ds \frac{d\ddot{y}(t)}{dt}$$

$$\ddot{y}(t) = dt \frac{d\ddot{y}(t)}{ds} \frac{dt}{ds}$$

$$\frac{ds}{dt} = |v(t)| = \sqrt{x^2(t) + y^2(t)}$$

$$= + \ddot{y}(t) \frac{x\ddot{y}(t)y(t) - y\ddot{y}(t)x(t)}{\sqrt{x^2(t) + y^2(t)}}^2$$

Per una linea retta allora è $x(t) = y(t) = 0$

Per una linea retta allora è $\ddot{x}(t) = \ddot{y}(t) = 0$

$$\ddot{y}(t) = 0$$

Per una linea retta allora è $\ddot{x}(t) = \ddot{y}(t) = 0$

$$\ddot{y}(t) = 0$$

Per un cerchio di raggio r è

$$\begin{aligned}y\ddot{y}(t) &= r \cos(t) \quad \ddot{x}(t) = \ddot{y}r \sin(t) \quad \ddot{y}(t) \\&= \ddot{y}r \sin(t) \quad \ddot{x}(t) = \ddot{y}r \cos(t)\end{aligned}$$

Per una linea retta allora è $\ddot{x}(t) = \ddot{y}(t) = 0$

$$\ddot{y}(t) = 0$$

Per un cerchio di raggio r è

$$y\ddot{y}(t) = r \cos(t) \quad \ddot{x}(t) = \ddot{y}r \sin(t) \quad \ddot{y}(t) = \ddot{y}r \sin(t)$$

$$\ddot{x}(t) = \ddot{y}r \cos(t)$$

$$1\ddot{y}(t) = -\frac{r}{R}$$

Nelle immagini digitali le curve sono dati discreti. Esistono tre modi per calcolare $\hat{y}(t)$

Calcola direttamente la differenza tra le direzioni angolari dei pixel del bordo consecutivi

Derivare una misura dei cambiamenti di curvatura nell'intensità dell'immagine
Ottenere una misura della curvatura mediante correlazione

Il rilevatore di Moravec

Rilevatore di Harris

Calcolo delle differenze nella direzione dei bordi

Approccio precoce (anni '70)

Richiede bordi precalcolati

Utilizza la direzione del gradiente $\hat{y}(t)$

Per ogni bordo i pixel sono

$$\hat{y}(t) = \hat{y}(t + 1) - \hat{y}(t - 1)$$

I valori di $\ddot{y}(t)$ definiti in
punti discreti potrebbero
non variare in modo uniforme

I valori di $\ddot{y}(t)$ definiti in punti

discreti potrebbero non variare
in modo uniforme

Può essere livellato considerando la differenza nella direzione angolare media di n pixel

$$1\ddot{y}(t) = \frac{1}{N} \sum_{i=1}^N \ddot{y}(t + i)$$

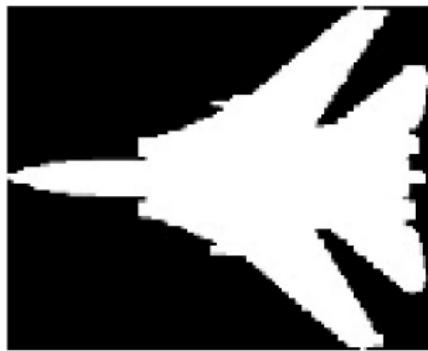
Può essere livellato considerando una media ponderata (es
livellamento gaussiano) di n pixel

$$1\ddot{y}(t) = \frac{1}{N} \sum_{i=1}^N w(i) \ddot{y}(t + i)$$

```
%Curvature detection
function outputimage=CurvConnect(inputimage)

[rows,columns]=size(inputimage); %Image size
outputimage=zeros(rows,columns); %Result image
[Mag,Ang]=Edges(inputimage); %Edge Detection
Mag=MaxSupr(Mag,Ang); %Maximal Suppression
Next=Cont(Mag,Ang); %Next connected pixels

%Compute curvature in each pixel
for x=1:columns-1
    for y=1:rows-1
        if Mag(y,x)~=0
            n=Next(y,x,1); m=Next(y,x,2);
            if(n~=-1 & m~-1)
                [px,py]=NextPixel(x,y,n);
                [qx,qy]=NextPixel(x,y,m);
                outputimage(y,x)=abs(Ang(py,px)-Ang(qy,qx));
            end
        end
    end
end
end
```



(a) Image



(b) Detected corners

Questo approccio non fornisce risultati affidabili

Si presuppone che le informazioni sugli angoli siano presenti nella mappa dei bordi rilevati

Non utilizza alcuna struttura angolare

Le misurazioni possono essere influenzate da errori di quantizzazione

La precisione è limitata

Misurazione della curvatura mediante variazioni di intensità

La curvatura viene derivata in funzione dei cambiamenti nell'intensità dell'immagine

Non è necessario calcolare i bordi

Misurazione della curvatura mediante variazioni di intensità

La curvatura viene derivata in funzione dei cambiamenti nell'intensità dell'immagine

Non è necessario calcolare i bordi

Supponiamo

(x, y) è e il punto del bordo $\delta(x,$

$y)$ è la funzione che rappresenta la direzione del bordo

Misurazione della curvatura mediante variazioni di intensità

La curvatura viene derivata in funzione dei cambiamenti nell'intensità dell'immagine

Non è necessario calcolare i bordi

Supponiamo

(x, y) è e il punto del bordo

$\hat{y}(x, y)$ è la funzione che rappresenta la direzione del bordo

Intorno a (x, y) la curva del bordo può essere approssimata in forma parametrica come

$$\begin{aligned}x(t) &= x + t \cos(\hat{y}(x, y)) \\y(t) &= y + t \sin(\hat{y}(x, y))\end{aligned}$$

$[\cos(\ddot{\gamma}(x, y)), \sin(\ddot{\gamma}(x, y))]$ è il vettore unitario nella direzione del bordo,
quindi t è la lunghezza dell'arco

$$\ddot{\gamma}(x, y) = dt \quad \frac{d\ddot{\gamma}(x, y)}{dt} = \frac{\ddot{\gamma}_x(x, y)}{\ddot{\gamma}_y} \frac{dx}{dt} + \frac{\ddot{\gamma}_y(x, y)}{\ddot{\gamma}_y} \frac{dy}{dt}$$

Dove

$$\frac{dx}{dt} = \cos(\ddot{\gamma}(x, y)) \qquad \frac{dy}{dt} = \operatorname{seno}(\ddot{\gamma}(x, y))$$

$[\cos(\ddot{y}(x, y)), \sin(\ddot{y}(x, y))]$ è il vettore unitario nella direzione del bordo,
quindi t è la lunghezza dell'arco

$$\ddot{y}(x, y) = dt \quad \frac{d\ddot{y}(x, y)}{\ddot{y}x} = \frac{\ddot{y}\ddot{y}(x, y)}{\ddot{y}x} \frac{dx}{dt} + \frac{\ddot{y}\ddot{y}(x, y)}{\ddot{y}y} \frac{dy}{dt}$$

Dove

$$\frac{dx}{dt} = \cos(\ddot{y}(x, y)) \qquad \frac{dy}{dt} = \operatorname{seno}(\ddot{y}(x, y))$$

Cos'è $\ddot{y}(x, y)$?

$[\cos(\ddot{y}(x, y)), \sin(\ddot{y}(x, y))]^T$ è il vettore unitario nella direzione del bordo, quindi t è la lunghezza dell'arco

$$\ddot{y}(x, y) = dt \quad \frac{d\ddot{y}(x, y)}{dt} = \frac{\ddot{y}\ddot{y}(x, y) dx + \ddot{y}x}{dt} \quad \frac{\ddot{y}\ddot{y}(x, y) dy + \ddot{y}y}{dt}$$

Dove

$$\frac{dx}{dt} = \cos(\ddot{y}(x, y)) \quad \frac{dy}{dt} = \operatorname{peccato}(\ddot{y}(x, y))$$

Cos'è $\ddot{y}(x, y)$? $\ddot{y}(x, y)$

è ortogonale al gradiente dell'immagine $[Ix, ly]^T$

$$\ddot{y}(x, y) = \operatorname{arcotan} \frac{ly}{Ix}$$

Sì

Pertanto lo è

Pertanto lo è

$$\frac{dx}{dt} = \cos(\ddot{y}(x, y)) = \frac{\ddot{y}ly}{io2x + \frac{lx}{y}} \quad \frac{dy}{dt} = \operatorname{peccato}(\ddot{y}(x, y)) = \frac{lx}{io2x + \frac{ly}{x}}$$

Pertanto lo è

$$\frac{dx}{dt} = \cos(\ddot{\gamma}(x, y)) = \frac{\ddot{y}ly}{\sqrt{x^2 + y^2}} \quad \frac{dy}{dt} = \text{peccato}(\ddot{\gamma}(x, y)) = \frac{lx}{\sqrt{x^2 + y^2}}$$

$$\ddot{\gamma}(x, y) = \frac{1}{\sqrt{x^2 + y^2}} (I_2_{\text{si}} l_{xx} - l_x l_y l_{yx} + I_2 x l_{yy} - l_x l_y l_{xy})$$

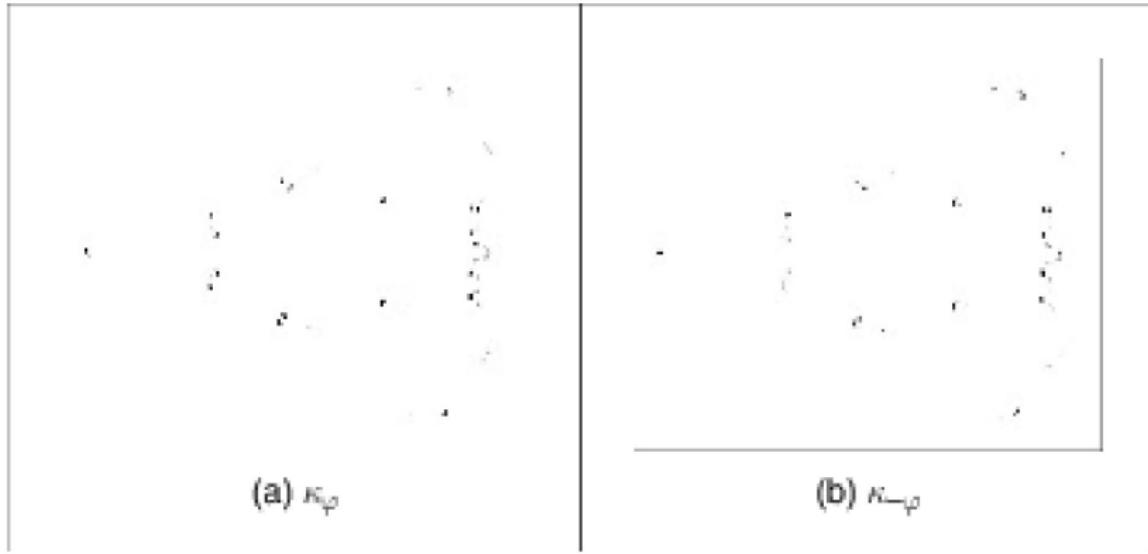
Questa è una misura diretta $\ddot{y}(x, y)$ di \ddot{y} .

Possiamo misurare la curvatura lungo altre direzioni, come $\ddot{y}\dot{y}(x, y)$

$$\begin{aligned}x(t) &= x + t \cos(\ddot{y}\dot{y}(x, y)) \quad y(t) \\&= y + t \sin(\ddot{y}\dot{y}(x, y))\end{aligned}$$

con

$$\ddot{y}(x, y) = \frac{1}{I_2^2 + I_1^2} (I_2^2 |_{\text{si}} l_{xx} \ddot{y} |_{\text{si}} l_x l_y l_{xy} \ddot{y} |_{\text{si}}^2 - x l_{yy} + l_x l_y l_{xy})$$



Non è un metodo totalmente affidabile per ottenere la curvatura.

In parte a causa di

differenziazione di ordine superiore,

approssimazione locale della curva

Rivelatore Moravec

Idea base

Una misura della curvatura può essere ottenuta considerando i cambiamenti lungo una particolare direzione nell'immagine.

Rivelatore Moravec

Idea base

Una misura della curvatura può essere ottenuta considerando i cambiamenti lungo una particolare direzione nell'immagine.

Calcola i cambiamenti medi nell'intensità dell'immagine
sposta una finestra $2a + 1 \times 2a + 1$ in più direzioni

Risolto il problema con il calcolo di un pixel (x, y) .

UN UN

$$E_{uv}(x, y) = \frac{(I(x+i, y+j) - I(x+i+u, y+j+v))^2}{i=j=u=v=a}$$

Una misura della curvatura è data da

$$\ddot{y}(x, y) = \min_{tu, v} U_{e,v}(x, y)$$

dove (u, v) varia in $\{(1, 0), (1, 1), (0, 1), (\ddot{y}_1, 1), (\ddot{y}_1, 0), (\ddot{y}_1, \ddot{y}_1), (0, \ddot{y}_1), (1, \ddot{y}_1)\}$

Una misura della curvatura è data da

$$\ddot{y}(x, y) = \min_{tu, v} U_{e,v}(x, y)$$

dove (u, v) varia in $\{(1,$

$0), (1, 1), (0, 1), (\ddot{y}1, 1), (\ddot{y}1, 0), (\ddot{y}1, \ddot{y}1), (0, \ddot{y}1), (1, \ddot{y}1)\}$

Il minimo perché

Una misura della curvatura è data da

$$\ddot{y}(x, y) = \min_{tu, v} U_{u,v}(x, y)$$

dove (u, v) varia in $\{(1, 0), (1, 1), (0, 1), (\ddot{y}_1, 1), (\ddot{y}_1, 0), (\ddot{y}_1, \ddot{y}_1), (0, \ddot{y}_1), (1, \ddot{y}_1)\}$
Il minimo perché

se (x, y) è in una zona pianeggiante

$U_{u,v}$ è sempre piccolo

Una misura della curvatura è data da

$$\ddot{y}(x, y) = \min_{tu, v} U_{u,v}(x, y)$$

dove (u, v) varia in $\{(1,$

$0), (1, 1), (0, 1), (\ddot{y}_1, 1), (\ddot{y}_1, 0), (\ddot{y}_1, \ddot{y}_1), (0, \ddot{y}_1), (1, \ddot{y}_1)\}$

Il minimo perché

se (x, y) è in una zona pianeggiante

$E_{u,v}$ è sempre

piccolo se (x, y) è un arco proprio

la curvatura è piccola

$E_{u,v}$ è piccolo per uno spostamento lungo un bordo dritto

$E_{u,v}$ è grande per uno spostamento lungo la normale

Una misura della curvatura è data da

$$\ddot{y}(x, y) = \min_{tu, v} U_{u,v}(x, y)$$

dove (u, v) varia in $\{(1,$

$0), (1, 1), (0, 1), (\ddot{y}_1, 1), (\ddot{y}_1, 0), (\ddot{y}_1, \ddot{y}_1), (0, \ddot{y}_1), (1, \ddot{y}_1)\}$

Il minimo perché

se (x, y) è in una zona pianeggiante

$U_{u,v}$ è sempre

piccolo se (x, y) è un arco proprio

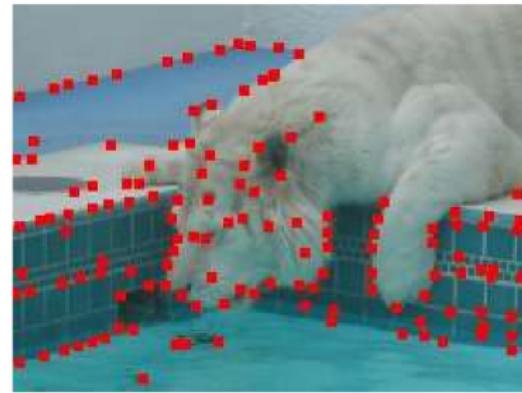
la curvatura è piccola

$U_{u,v}$ è piccolo per uno spostamento lungo un bordo dritto

$U_{u,v}$ è grande per uno spostamento lungo la

normale se (x, y) è un punto

d'angolo tutti gli spostamenti producono un valore grande



Il problema principale con il rilevatore di Moravec è

La risposta è anisotropa: considera solo un piccolo insieme discreto di spostamenti

La risposta è rumorosa: utilizza una finestra binaria rettangolare

Considera solo il minimo di E, quindi risponde troppo prontamente ai bordi

La maggior

parte dei problemi sono risolti con il rilevatore Harris.

Rilevatore d'angolo Harris

dà un'espressione analitica per
 $(I(x + i, y + j) - I(x + i + u, y + j + v))$ usa una
maschera di ponderazione
gaussiana cambia la misura dell'angolo $\hat{\gamma}(x, y)$

Rilevatore d'angolo Harris

dà un'espressione analitica per
 $(I(x + i, y + j) - I(x + i + u, y + j + v))$ usa una
maschera di ponderazione gaussiana
cambia la misura dell'angolo $\hat{\gamma}(x, y)$

Considerando lo sviluppo di Taylor del primo ordine attorno a (x, y)

$$I(x + u, y + v) = I(x, y) + I_x(x, y)u + I_y(x, y)v$$

Rilevatore d'angolo Harris

dà un'espressione analitica per
 $I(x + i, y + j) \circ I(x + i + u, y + j + v)$ usa una
maschera di ponderazione gaussiana
cambia la misura dell'angolo $\theta(x, y)$

Considerando lo sviluppo di Taylor del primo ordine attorno a (x, y)

$$I(x + u, y + v) = I(x, y) + I_x(x, y)u + I_y(x, y)v$$

Ciò vale anche per $(x + i, y + j)$

$$I(x+i+u, y+j+v) = I(x+i, y+j) + I_x(x+i, y+j)u + I_y(x+i, y+j)v$$

$$\begin{array}{ccc} \text{UN} & & \text{UN} \\ \\ \text{E} u, v(x, y) = & w(i, j)(lx(x+i, y+j)u + ly(x+i, y+j)v) & ^2 \\ & i=\ddot{y}aj=\ddot{y}a & \end{array}$$

dove $w(i, j)$ è una maschera gaussiana.

$$\text{Eu,v}(x, y) = \sum_{i=0}^{a} \sum_{j=0}^{b} w(i, j) (lx(x+i, y+j)u + ly(x+i, y+j)v)^2$$

dove $w(i, j)$ è una maschera gaussiana.

Dopo un po' di algebra

$$\text{Eu,v} = A(x, y)u^2 + 2C(x, y)uv + B(x, y)v^2$$

Dove

$$A(x, y) = \sum_{i,j=0}^{a,b} lx(x+i, y+j)^2 \quad B(x, y) = \sum_{i,j=0}^{a,b} ly(x+i, y+j)^2$$

$$C(x, y) = \sum_{i,j=0}^{a,b} lx(x+i, y+j)ly(x+i, y+j)$$

Il problema ora è trovare quando $E_{u,v}$ è minimo.

Il problema ora è trovare quando $E_{u,v}$ è minimo.

$E_{u,v}$ è una funzione quadratica e quindi ha 2 assi principali
(notare che le isocurve sono coniche)

Il problema ora è trovare quando $E_{u,v}$ è minimo.

$E_{u,v}$ è una funzione quadratica e quindi ha 2 assi principali
(notare che le isocurve sono coniche)

Possiamo ruotare la funzione in modo che gli assi siano paralleli al sistema di coordinate

$$\begin{matrix} R & \stackrel{tu}{\sim} & = & \mu \\ v & & & \ddot{y} \end{matrix}$$

Il problema ora è trovare quando $E_{u,v}$ è minimo.

$E_{u,v}$ è una funzione quadratica e quindi ha 2 assi principali
(notare che le isocurve sono coniche)

Possiamo ruotare la funzione in modo che gli assi siano paralleli al sistema di coordinate

$$R \begin{matrix} u \\ v \end{matrix} = \begin{matrix} \mu \\ \ddot{y} \end{matrix}$$

$$E_{\ddot{y},\ddot{y}} = \ddot{y}(x, y) 2 \ddot{y}^2 + \ddot{y}(x, y) 2 \ddot{y}^2$$

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}, \ddot{y}(x, y)$

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}$, $\ddot{y}(x, y)$ se (x, y) è in una regione piatta

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}$, se (x, y) è in una regione piatta

$E\ddot{y}$ è piccolo per ogni (\ddot{x}, \ddot{y}) sia $\ddot{y}(x, y)$ che $\ddot{y}(x, y)$ sono piccoli

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}$, $\ddot{y}(x, y)$ se (x, y) è in una regione piatta

$E\ddot{y}$, \ddot{y} è piccolo per ogni (\ddot{x}, \ddot{y}) sia $\ddot{y}(x, y)$ che $\ddot{y}(x, y)$ sono piccoli se (x, y) è un arco

I valori di $\ddot{y}(x, y)$ e $\ddot{\bar{y}}(x, y)$ sono proporzionali a $E\ddot{y}, \ddot{\bar{y}}(x, y)$ se (x, y) è in una regione piatta

$E\ddot{y}, \ddot{\bar{y}}$ è piccolo per ogni (\ddot{x}, \ddot{y})

sia $\ddot{y}(x, y)$ che $\ddot{\bar{y}}(x, y)$ sono piccoli se $(x,$

$y)$ è una variazione

più piccola del bordo (E più piccola) lungo la direzione del bordo

variazione più grande (E più grande) lungo la normale

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}, \ddot{y}(x, y)$ se (x, y) è in una regione piatta

$E\ddot{y}, \ddot{y}$ è piccolo per ogni (\ddot{x}, \ddot{y})
sia $\ddot{y}(x, y)$ che $\ddot{y}(x, y)$ sono piccoli se

(x, y) è una variazione

più piccola del bordo (E più piccola) lungo la direzione del
bordo variazione più grande (maggiore E) lungo
la direzione normale del bordo e normali sono gli
assi principali $\ddot{y}(x, y)$ grande e $\ddot{y}(x, y)$ piccolo, o viceversa

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}$, se (x, y) è in una regione piatta

$E\ddot{y}$ è piccolo per ogni (\ddot{x}, \ddot{y})
sia $\ddot{y}(x, y)$ che $\ddot{y}(x, y)$ sono piccoli se
 (x, y) è una variazione più piccola del bordo (E più piccola) lungo la direzione del bordo variazione più grande
(maggiore E) lungo la direzione normale del bordo e normali sono gli assi principali $\ddot{y}(x, y)$ grande e $\ddot{y}(x, y)$ piccolo, o viceversa se (x, y) è un angolo

I valori di $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono proporzionali a $E\ddot{y}, \ddot{y}(x, y)$ se (x, y) è in una regione piatta

$E\ddot{y}, \ddot{y}$ è piccolo per ogni (\ddot{y}, \ddot{y})

sia $\ddot{y}(x, y)$ che $\ddot{y}(x, y)$ sono piccoli se

(x, y) è una variazione

più piccola del bordo (E più piccola) lungo la direzione del

bordo variazione più grande (maggiore E) lungo

la direzione normale del bordo e normale sono gli assi

principali $\ddot{y}(x, y)$ grande e $\ddot{y}(x, y)$ piccolo, o

viceversa se (x, y) è

una variazione d'angolo grande lungo

tutte le direzioni entrambe $\ddot{y}(x, y)$ e $\ddot{y}(x, y)$ sono grandi

La curvatura (risposta d'angolo) è definita come

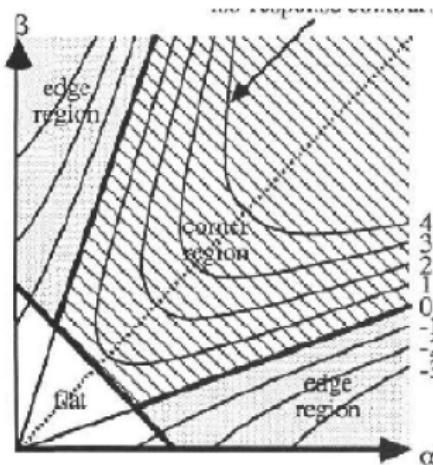
$$\ddot{y}(x, y) = \ddot{y}(x, y)\ddot{y}(x, y) \ddot{y} K(\ddot{y}(x, y) + \ddot{y}(x, y))^2$$

\ddot{y} è

piccolo nelle zone pianeggianti

positivo per gli angoli

negativo per i bordi



Come otteniamo la risposta?

$$M = \begin{matrix} A(x, y) & C(x, y) \\ C(x, y) & B(x, y) \end{matrix}$$

$$D = [u, v]^T$$

$$Euv(x, y) = D tMD$$

$$E\ddot{y}, \ddot{y}(x, y) = (RD) tMRD = D tQD$$

$$Q = \begin{matrix} \ddot{y}(x, y) & 0 \\ 0 & \ddot{y}(x, y) \end{matrix}$$

Come otteniamo la risposta?

$$M = \begin{pmatrix} A(x, y) & C(x, y) \\ C(x, y) & B(x, y) \end{pmatrix}$$

$$D = [u, v]^T$$

$$Euv(x, y) = D tMD$$

$$E\ddot{y}, \ddot{y}(x, y) = (RD) tMRD = D tQD$$

$$Q = \begin{pmatrix} \ddot{y}(x, y) & 0 \\ 0 & \ddot{y}(x, y) \end{pmatrix}$$

\ddot{y} e \ddot{y} sono gli autovalori di M

Noi possiamo

Calcolare esplicitamente gli autovalori
Determina $\ddot{y}\ddot{y}$ e $\ddot{y} + \ddot{y}$

Harris segue il secondo approccio che è più veloce

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^t M R) = \det(R^t) \det(M) \det(R)$$

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^t M R) = \det(R^t) \det(M) \det(R)$$

$$\det(R) = 1$$

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^T M R) = \det(R^T) \det(M) \det(R)$$

$$\det(R) = 1$$

$$\ddot{y}y = \det(Q) = \det(M)$$

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^t M R) = \det(R^t) \det(M) \det(R)$$

$$\det(R) = 1$$

$$\|y\| = \det(Q) = \det(M) \|y\|$$

$$= AB \|C\|^2$$

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^t M R) = \det(R^t) \det(M) \det(R)$$

$$\det(R) = 1$$

$$\sqrt{\det(Q)} = \sqrt{\det(M)} \sqrt{\det(R)}$$

$$= \sqrt{AB} \sqrt{C}^2$$

La traccia è invariante quindi $\sqrt{A} + \sqrt{B} = \text{traccia}(M)$

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^t M R) = \det(R^t) \det(M) \det(R)$$

$$\det(R) = 1$$

$$\|y\| = \det(Q) = \det(M) \|y\|$$

$$= \|A\| \|B\|^2 \|C\|^2$$

La traccia è invariante quindi $\|y\| + \|y\| = \text{traccia}(M)$

$$\|y(x, y)\| = \|A(x, y)\| \|B(x, y)\| \|C(x, y)\|^2 \leq K(\|A(x, y)\| + \|B(x, y)\|)^2$$

Harris segue il secondo approccio che è più veloce

$$\det(Q) = \det(R^t M R) = \det(R^t) \det(M) \det(R) \det(R) = 1$$

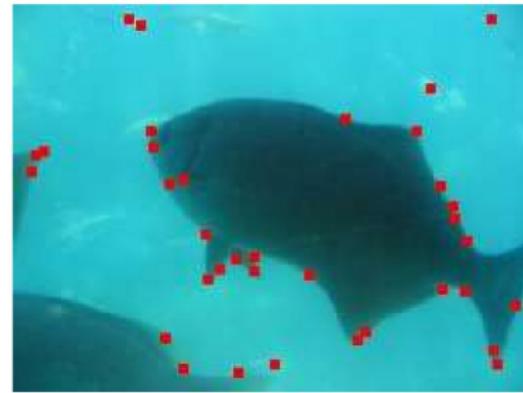
$$\ddot{y} \ddot{y} = \det(Q)$$

$$= \det(M) \ddot{y} \ddot{y} = AB \ddot{y} C$$

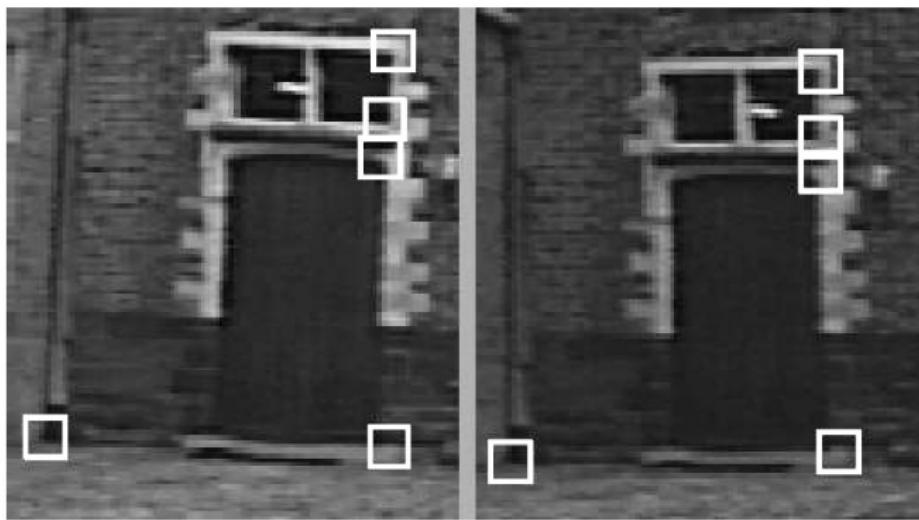
La traccia è

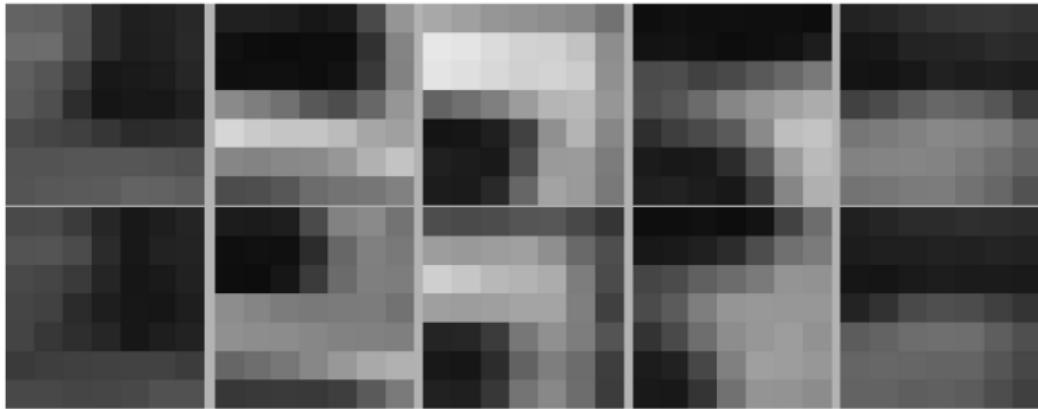
$$\text{invariante quindi } \ddot{y} + \ddot{y} = \text{traccia}(M) \ddot{y}(x, y) = A(x, y)B(x, y) \ddot{y} C(x, y) \quad 2 \quad \ddot{y} K(A(x, y) + B(x, y))2$$

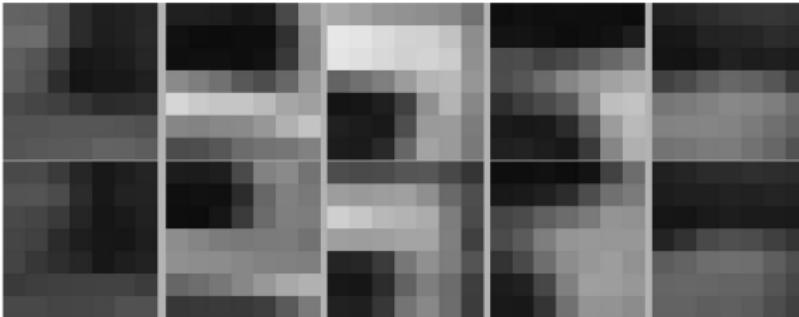
K è empiricamente impostato su valori compresi tra 0,004 e 0,15











0.9639	-0.3994	-0.1627	-0.3868	0.1914
-0.0533	0.7503	-0.4677	0.5115	0.7193
-0.1826	-0.3905	0.7730	0.1475	-0.7457
-0.2724	0.4878	0.1640	0.7862	0.2077
0.0835	0.5044	-0.4541	0.2802	0.9876

Valori di correlazione incrociata dell'intensità per tutte le possibili combinazioni
dei 5 angoli

Rivelatore Tomasi-Kanade

Tomasi e Kanade lo hanno dimostrato

quando le deformazioni affini tra le immagini Harris $\ddot{y}(x, y)$ non sono una buona misura

mantengono lo stesso algoritmo $\ddot{y}(x, y)$

$$= \min(\ddot{y}(x, y), \ddot{y}(x, y))$$

Lo spazio della scala

Finora il contenuto di un'immagine è descritto da

bordo

angoli

Lo spazio della scala

Finora il contenuto di un'immagine è descritto da

bordo

angoli

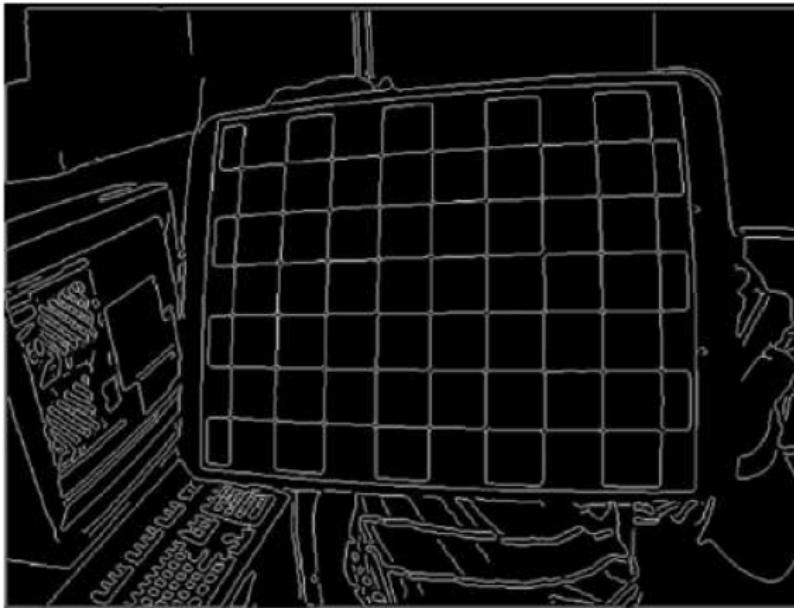


Lo spazio della scala

Finora il contenuto di un'immagine è descritto da

bordo

angoli

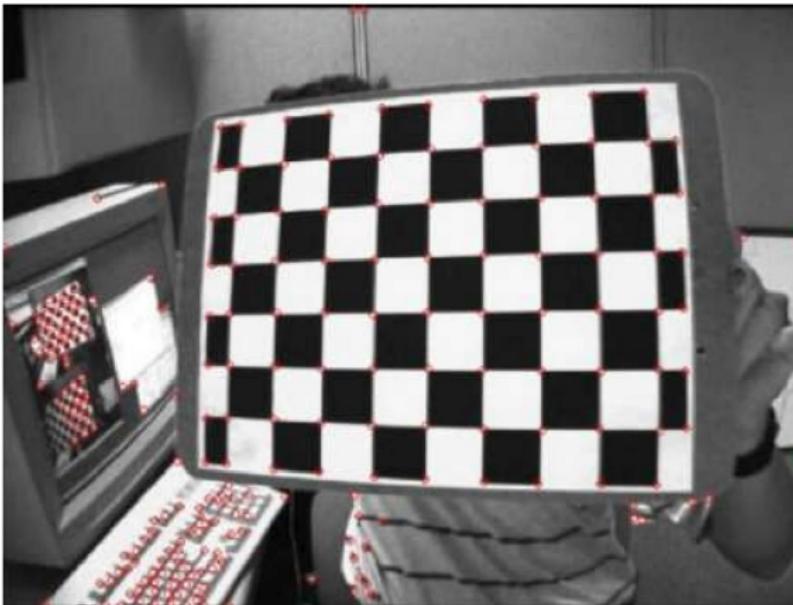


Lo spazio della scala

Finora il contenuto di un'immagine è descritto da

bordo

angoli



Percepiamo alcuni
dettagli a seconda della
nostra distanza dagli oggetti

Gli oggetti in un'immagine
hanno significato solo se
osservati ad una certa scala

La descrizione che
associamo ad an
oggetti dipende
fortemente dalla scala
utilizzata per modellare
l'oggetto



Percepiamo alcuni
dettagli a seconda della
nostra distanza dagli oggetti

Gli oggetti in un'immagine
hanno significato solo se
osservati ad una certa scala

La descrizione che
associamo ad an
oggetti dipende
fortemente dalla scala
utilizzata per modellare
l'oggetto



Percepiamo alcuni
dettagli a seconda della
nostra distanza dagli oggetti

Gli oggetti in un'immagine
hanno significato solo se
osservati ad una certa scala

La descrizione che
associamo ad an
oggetti dipende
fortemente dalla scala
utilizzata per modellare
l'oggetto



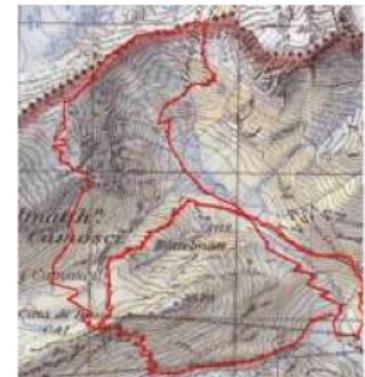
Esempi



Esempi



Esempi





Cos'è la scala?

Cattura le informazioni sull'estensione spaziale di una struttura nell'immagine.

Cos'è la scala?

Cattura le informazioni sull'estensione spaziale di una struttura nell'immagine.



Cos'è la scala?

Cattura le informazioni sull'estensione spaziale di una struttura nell'immagine.



Cos'è la scala?

Cattura le informazioni sull'estensione spaziale di una struttura nell'immagine.



Rappresentazione multiscala

In generale non possiamo dire quale sia la scala interessante

Non sappiamo cosa c'è nell'immagine.

E' necessario un metodo per la selezione automatica della scala

Soluzioni diverse

Spazio su scala gaussiana [Lindeberg 1994]

DoG e SIFT [Lowe 1999]

Harris multiscala [Mikolajczyk 2002]

Spazio di scala: definizione

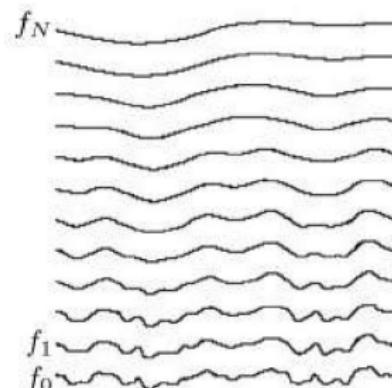
Il segnale originale è
rappresentato da una famiglia
parametrica di segnali

Tutti i segnali derivano da
quello originale

Le strutture più fini
vengono soppresse

Non devono esistere nuove strutture
creato

Domanda: come si passa da f_i
a f_{i+1} ?



Dato un segnale

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

La sua rappresentazione spaziale in scala è data da

$$L(x,t) = f(x \cdot y) g(y) dy$$

Dove

Dato un segnale

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

La sua rappresentazione spaziale in scala è data da

$$L(x,t) = f(x - y)g(y) dy$$

Dove

t è un parametro di scala

Dato un segnale

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

La sua rappresentazione spaziale in scala è data da

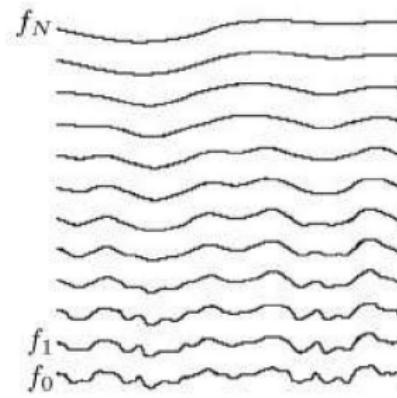
$$L(x,t) = f(x - y)g(y) dy$$

Dove

t è un parametro di scala

g è una gaussiana con $y = x - t$

$$f_i = f_0 g(y_i)$$



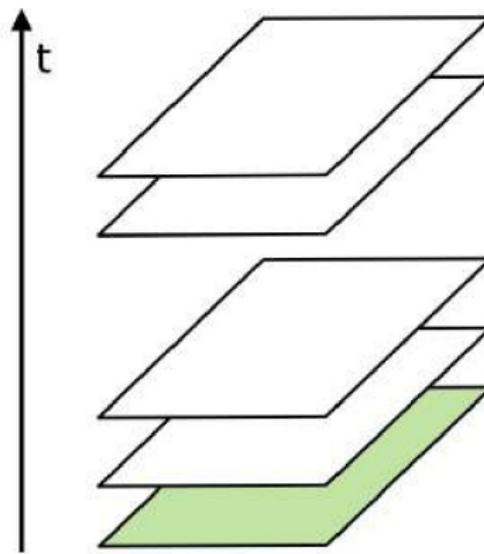
E per un'immagine?

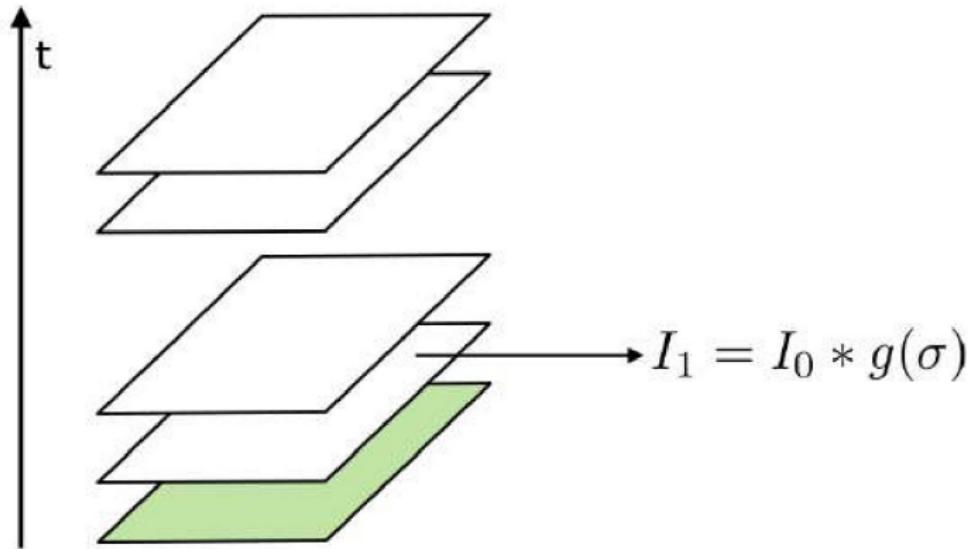


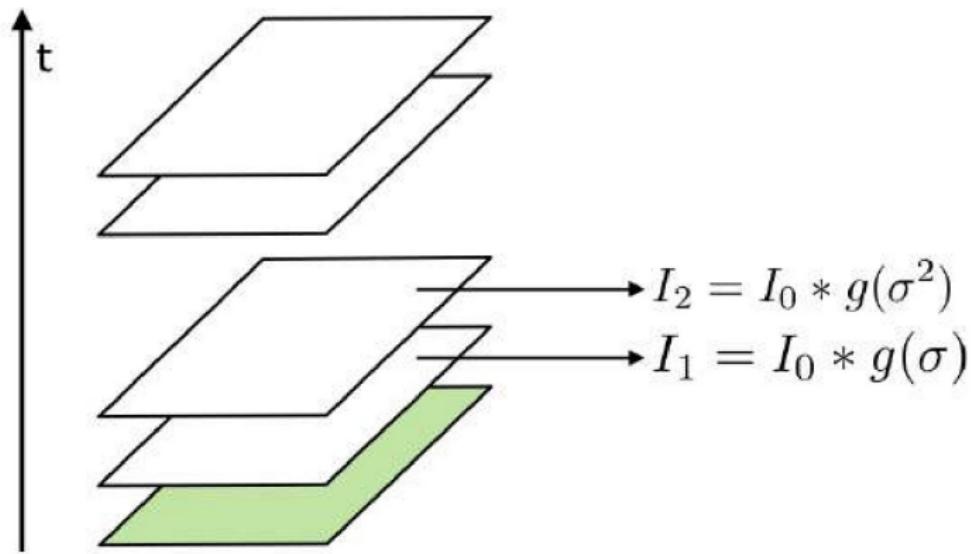
le strutture iniziano a scomparire con la scala
non vengono create nuove strutture

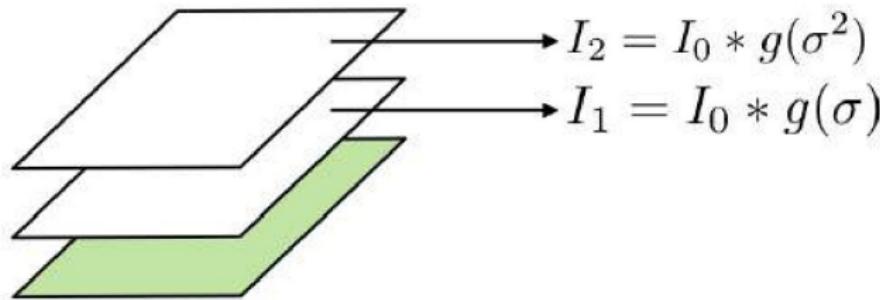
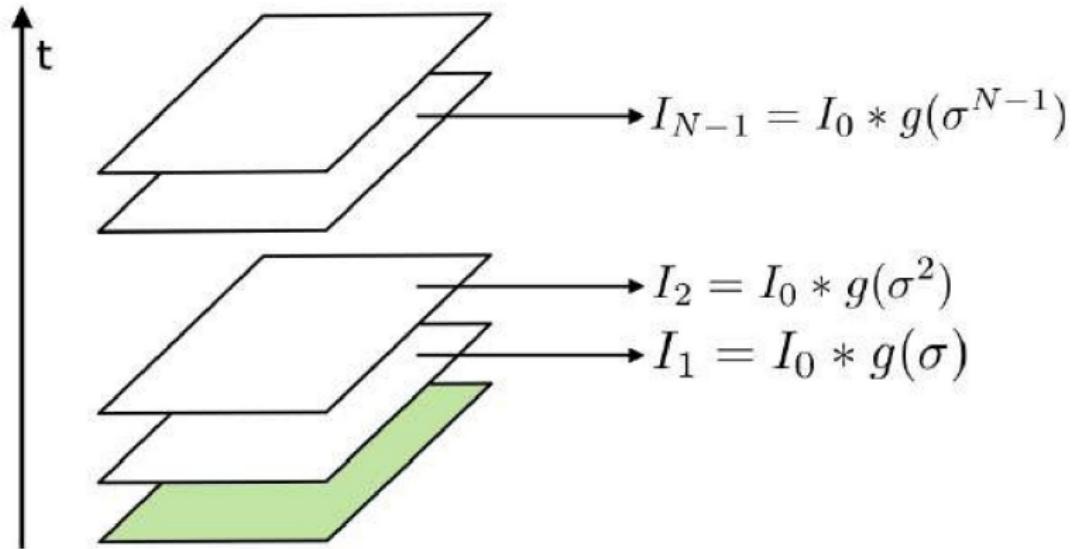
I dettagli più piccoli vengono identificati solo su scala più fine (ovvero, risoluzione più elevata)

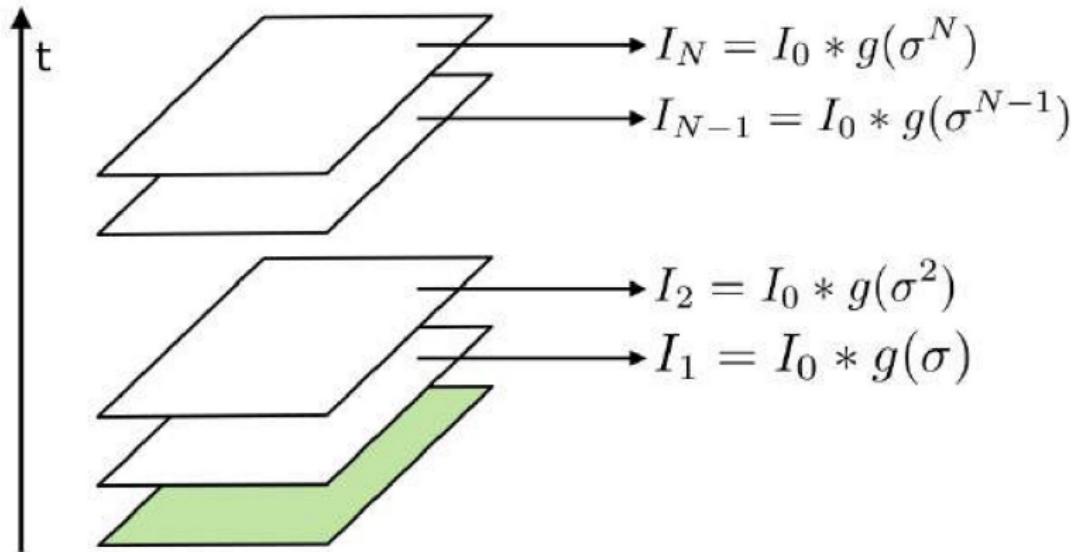
Non li vedremmo da lontano Tendono a scomparire man mano che il parametro di scala aumenta Le strutture più grandi sono identificate alla scala più grossolana (cioè a bassa risoluzione)









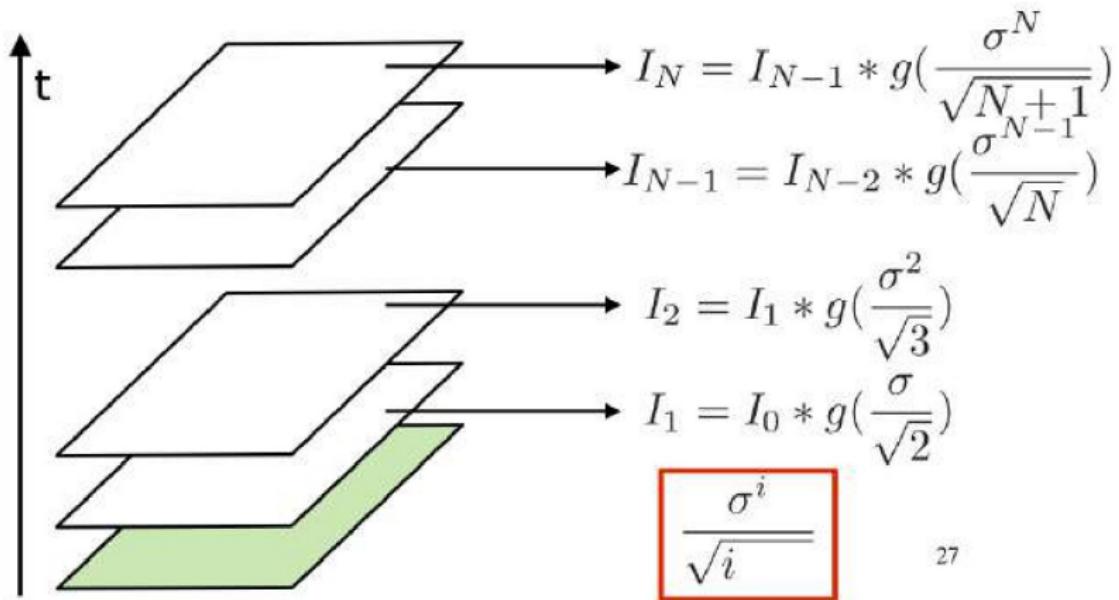


$\ddot{y} = \ddot{y}_2$ è considerato un buon valore per la scala iniziale

$\hat{y} = \hat{y}^2$ è considerato un buon valore per la scala iniziale. Dal punto di vista computazionale è una buona idea filtrare l'immagine originale per ciascuna scala?

$\hat{y} = \hat{y}^2$ è considerato un buon valore per la scala iniziale. Dal punto di vista computazionale è una buona idea filtrare l'immagine originale per ciascuna scala?

Possiamo filtrare l'immagine alla scala precedente.



Rappresentazione pyramidale

Tutte le immagini nello spazio scala hanno la stessa dimensione

Un altro approccio utilizza una rappresentazione pyramidale

Ad ogni livello la dimensione viene ridotta (tipicamente della metà $\sqrt{2}$)

Per il downsampling è possibile utilizzare un filtro medio

Rilevamento delle caratteristiche nello spazio di scala

Abbiamo a che fare con N immagini filtrate

L'idea principale è

Estrai punti interessanti su ciascuna scala

Selezione la scala più adeguata

Caratteristiche diverse:

bordo

angoli

macchia

...

L'algoritmo dipende dalla funzionalità

Ci concentriamo sugli angoli/punti

Tomasi-Kanade nello spazio della scala

L'estrazione è, mutatis mutandis, la stessa del caso a scala singola.

Calcola il gradiente

Per ogni punto

Calcola la matrice M

Calcolare il più piccolo autovalore $\hat{\gamma}$

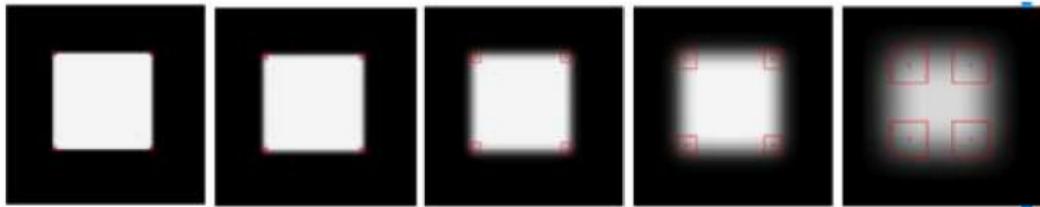
Se $\hat{\gamma}$ è grande, segna il punto

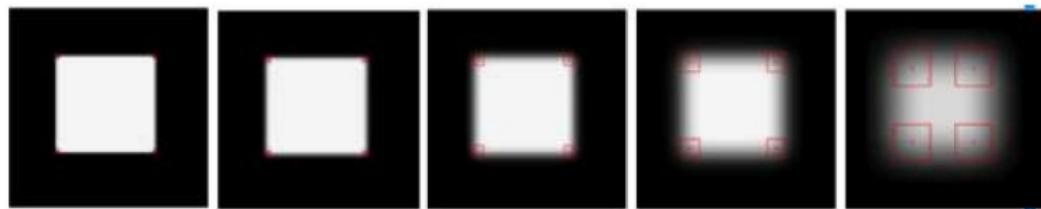
Soppressione dei non massimi

Ogni punto rilevato sarà identificato da

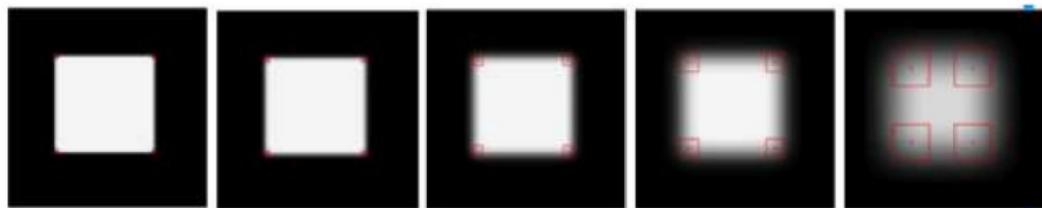
coordinate

scala





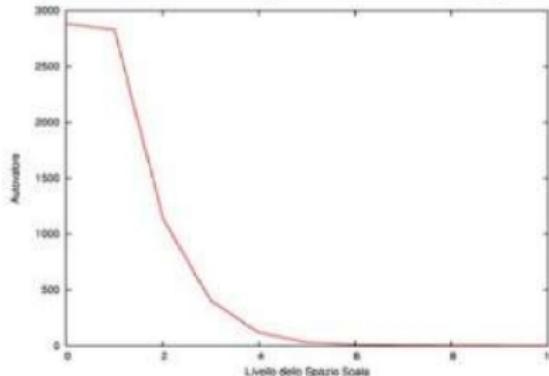
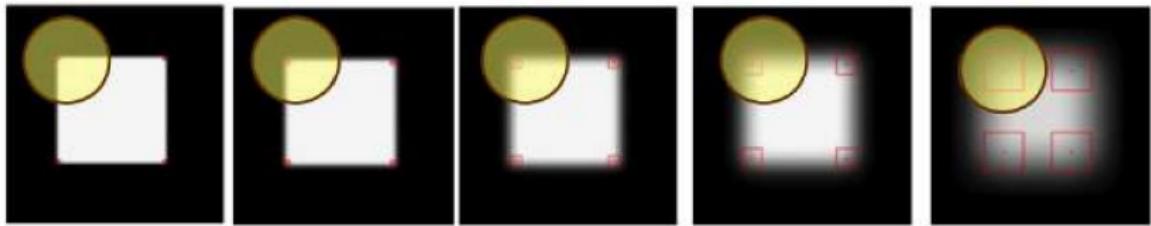
La localizzazione diventa meno accurata all'aumentare della scala



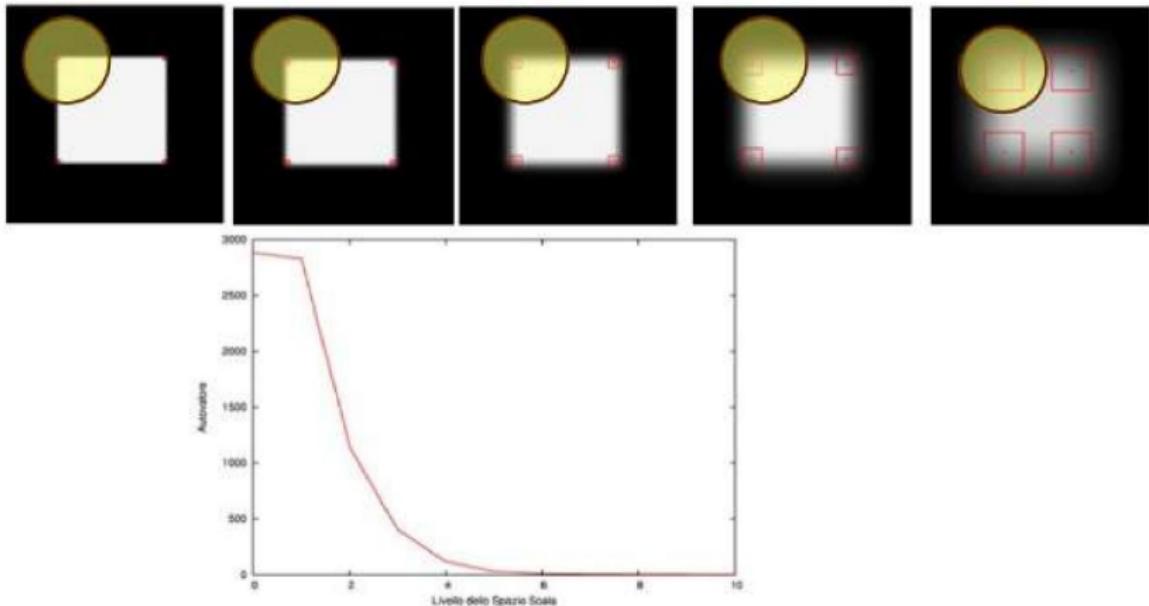
La localizzazione diventa meno accurata all'aumentare della scala

Gli angoli si allontanano nella direzione del gradiente

Come cambia \hat{y} con la scala?



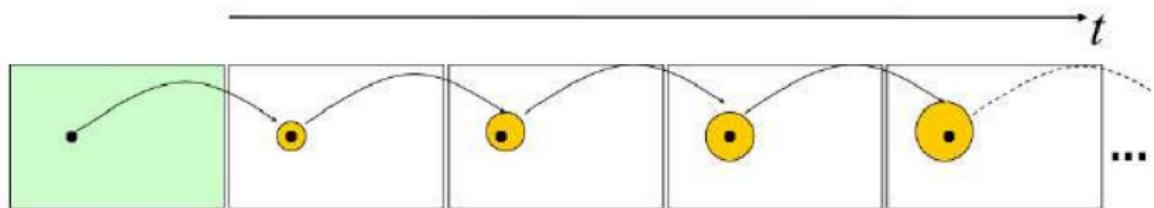
Come cambia \hat{y} con la scala?



Pertanto \hat{y} non può aiutare nella scelta della scala migliore

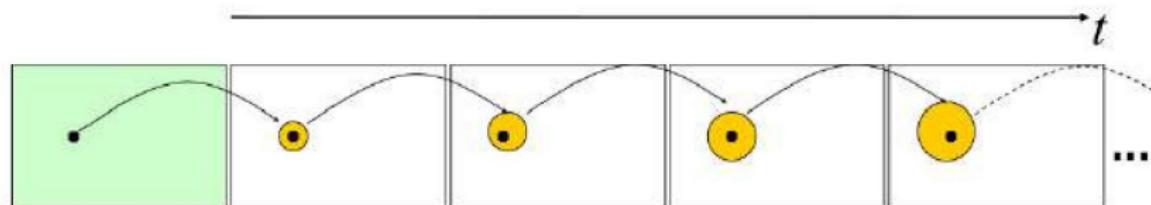
Selezione automatica della scala

Possiamo ricostruire il percorso di un angolo nello spazio in scala.



Selezione automatica della scala

Possiamo ricostruire il percorso di un angolo nello spazio in scala.



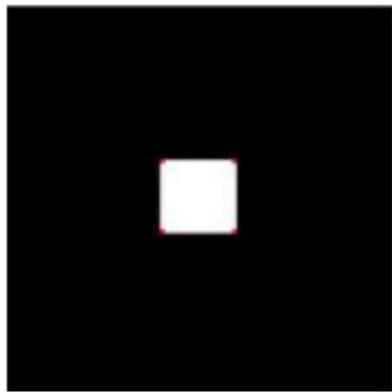
Usiamo una misura d'angolo, definita da Lindeberg
Dato
un percorso d'angolo c_1, \dots, c_k

$$= t \times (c_i) l_{yy}^2 + \sum_{i=1}^k (c_i) l_{xx}(c_i) - 2l_x(c_i) l_y(c_i) l_{xy}(c_i) + l_z^2 s(c_i)$$

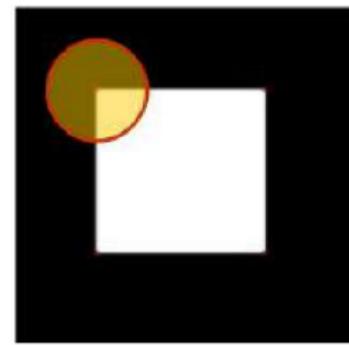
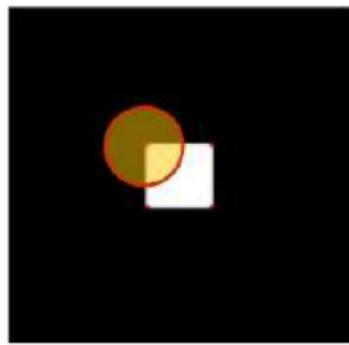
La scala giusta è quella che massimizza le s sul percorso.

Viene introdotto un secondo passaggio iterativo per migliorare
la localizzazione [Lindeberg 1998]

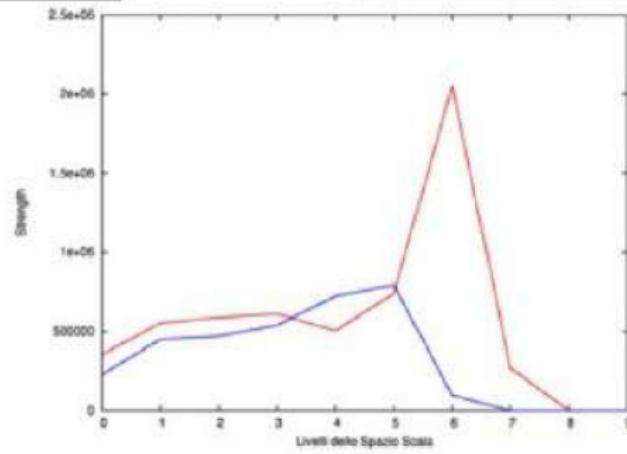
Esempio



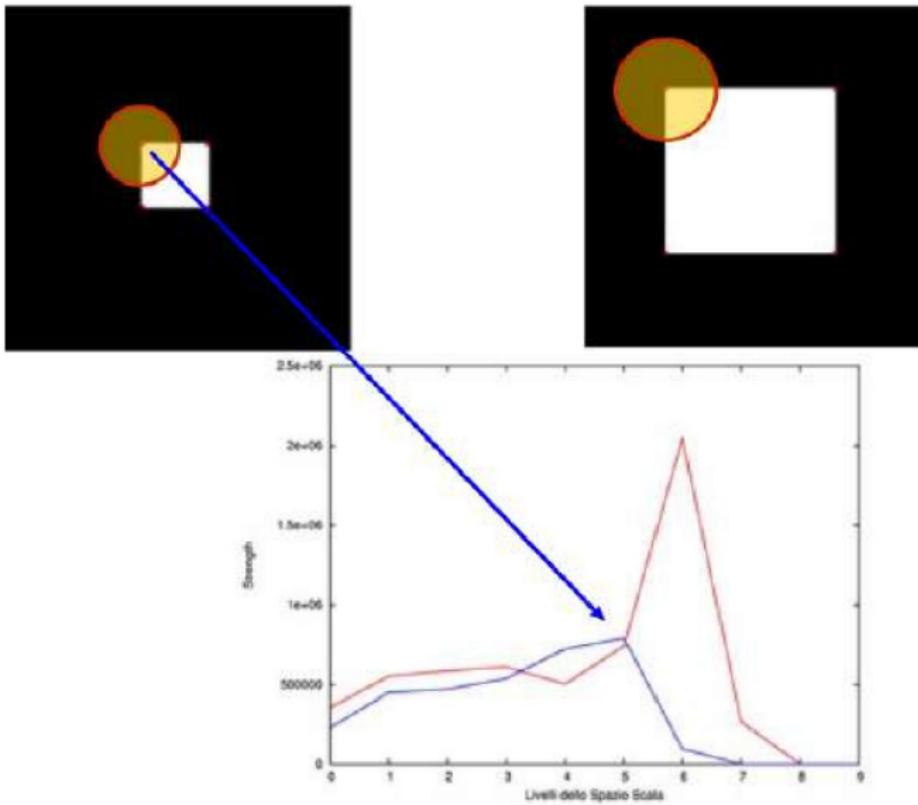
Esempio



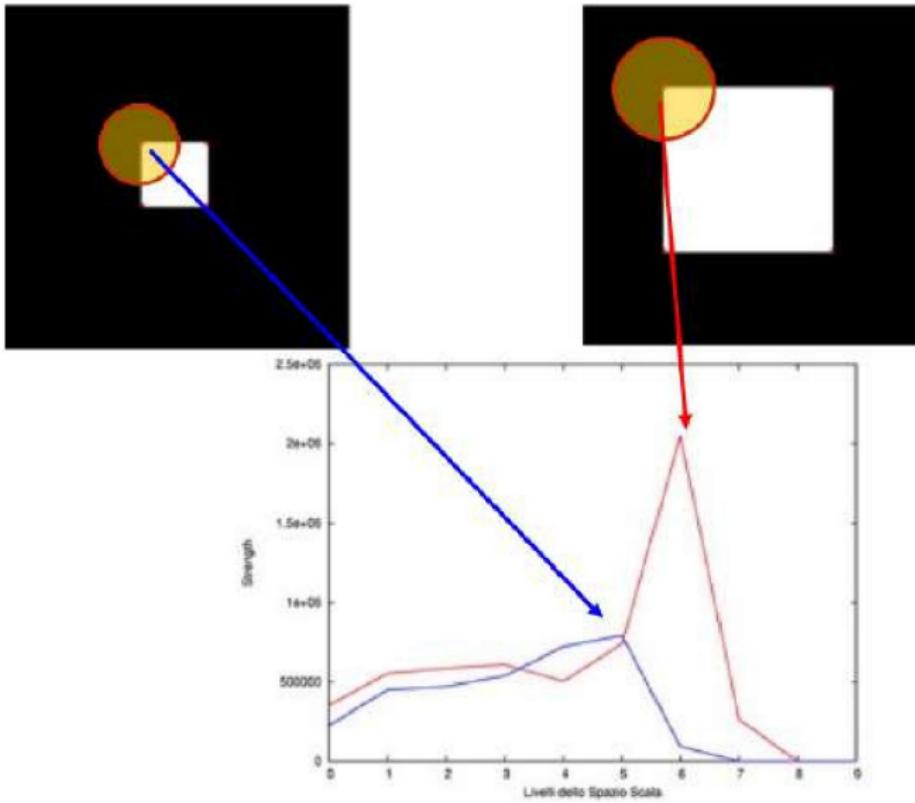
Esempio



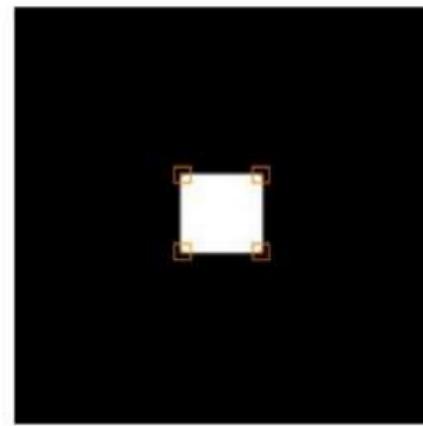
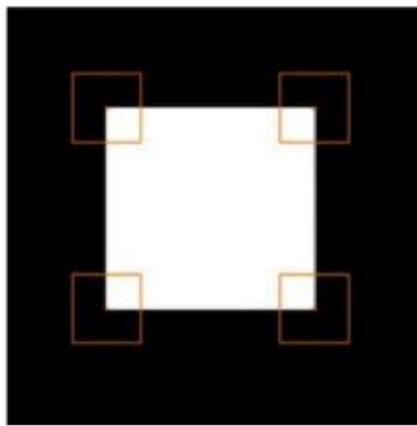
Esempio



Esempio



Esempio



Un altro esempio

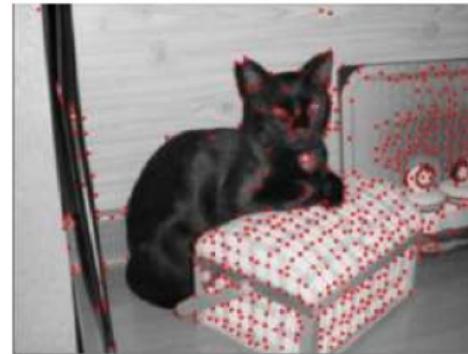


Immagine originale

Un altro esempio



Immagine originale

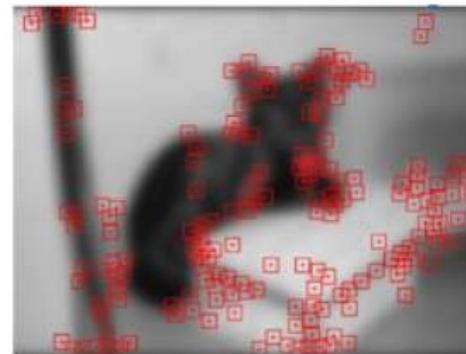


Angoli a livello 0

Un altro esempio



Immagine originale



Angoli al livello 5

Un altro esempio



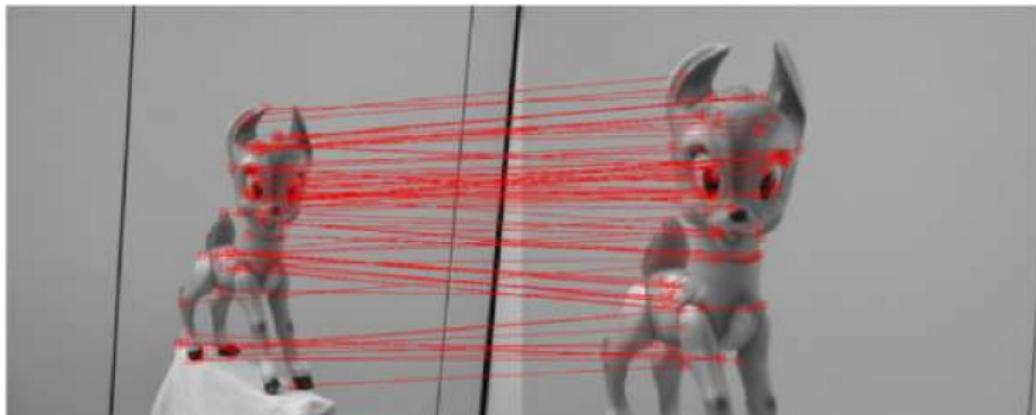
Immagine originale



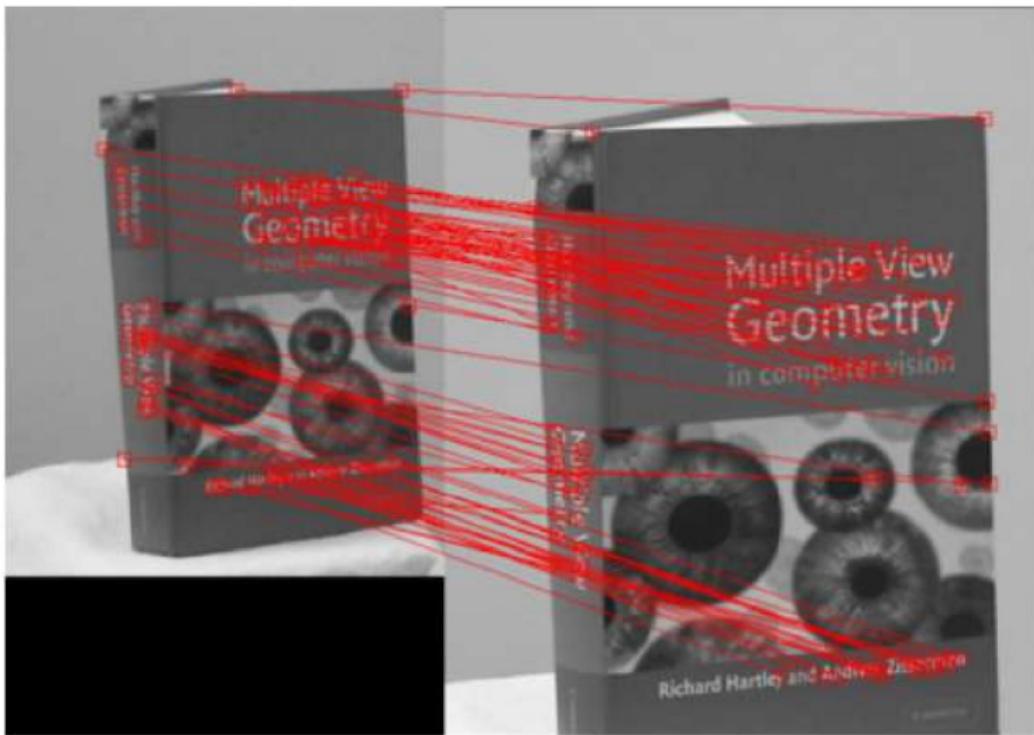
Angoli estratti

Applicazione: abbinamento di immagini a diversa risoluzione

Applicazione: abbinamento di immagini a diversa risoluzione



Applicazione: abbinamento di immagini a diversa risoluzione



Caratteristiche del SIFT

SIFT [Lowe 1999] sono caratteristiche locali

Creata per attività di riconoscimento degli oggetti

Mostrano una certa invarianza rispetto ai

cambiamenti di
scala,

rotazione, trasformazione

affine, cambiamenti di
illuminazione, cambiamenti del punto di vista

Passaggi principali

1. Determinare gli estremi nello spazio di scala 2.

Localizzazione dei punti caratteristici 3.

Associare una direzione a ciascun punto caratteristico

4. Calcolare un vettore di descrizione per ciascun punto caratteristico

La piramide del cane

La differenza delle gaussiane è definita come

$$D(x, y, \ddot{y}) = L(x, y, k\ddot{y}) - L(x, y, \ddot{y}) = (G(x, y, k\ddot{y}) - G(x, y, \ddot{y}))l(x, s)$$

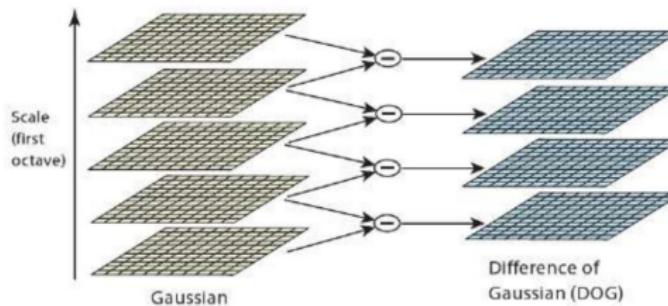
dove k è un fattore moltiplicativo costante.

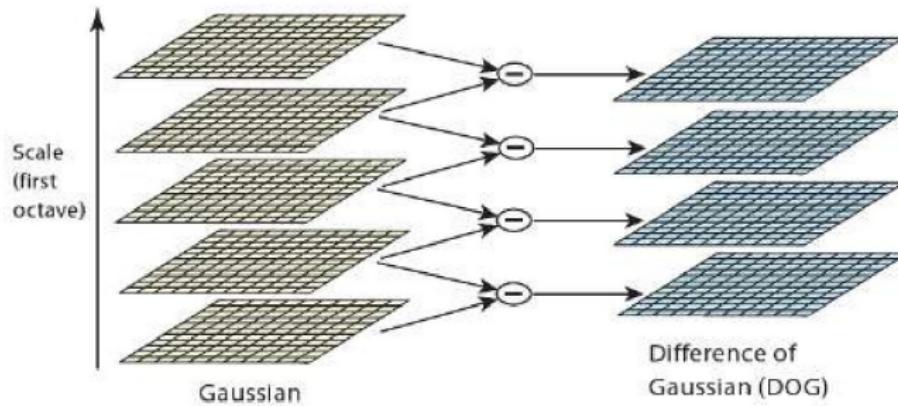
La piramide del cane

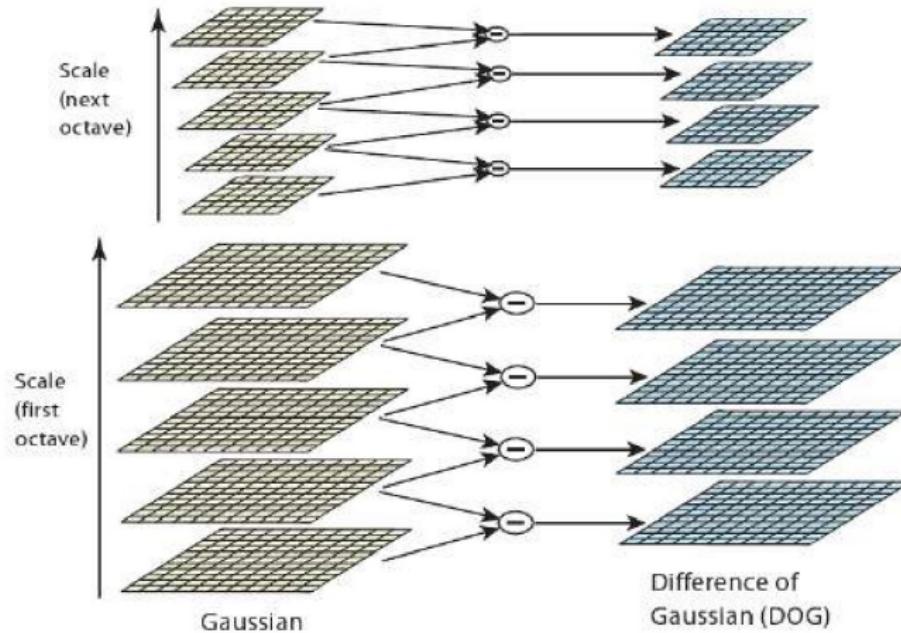
La differenza delle gaussiane è definita come

$$D(x, y, \ddot{y}) = L(x, y, k\ddot{y}) - L(x, y, \ddot{y}) = (G(x, y, k\ddot{y}) - G(x, y, \ddot{y}))I(x, s)$$

dove k è un fattore moltiplicativo costante.







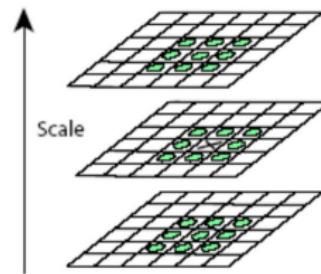
Determina gli estremi nella piramide del cane

Determinare i massimi e i minimi

locali in $D(x, y, \ddot{y})$

Deve essere un'estrema per un

quartiere 3D



Localizzazione dei punti caratteristici

Gli estremi locali sono punti caratteristici candidati

Punti

di scarto a basso contrasto

localizzati lungo un bordo (instabili al rumore e non invarianti)

Localizzazione più accurata dei punti adattando una funzione quadratica 3D

Eliminazione dei bordi

Autovalori λ e λ' della M dall'algoritmo dell'angolo di Harris

$$M = \begin{matrix} A(x, y) & C(x, y) \\ C(x, y) & B(x, y) \end{matrix}$$

se (x, y) è uno spigolo: λ grande e λ' piccolo, o viceversa

$\lambda + \lambda' = \text{trace}(M)$ e $\lambda\lambda' = \det(M)$ se $\lambda = r\lambda'$

$(\lambda > \lambda')$

$$\frac{(\lambda+\lambda')\lambda\lambda'}{r\lambda'}^2 = \frac{(r\lambda' + \lambda')^2}{r\lambda'}^2 = \frac{(r+1)2}{r}$$

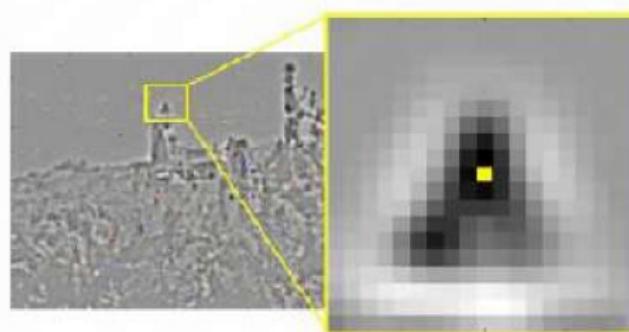
$\frac{(r+1)2}{r}$ minimo quando $\lambda = \lambda'$ e aumenta con r

Vogliamo verificare quando r è grande, ovvero fissando una soglia r_0

$$\frac{\text{traccia}(M)}{\det(M)} < \frac{(r_0 + 1)2}{r_0}$$

Direzione principale

Assegnare una direzione principale a ciascun punto caratteristico consente di avere descrittori invarianti alla rotazione

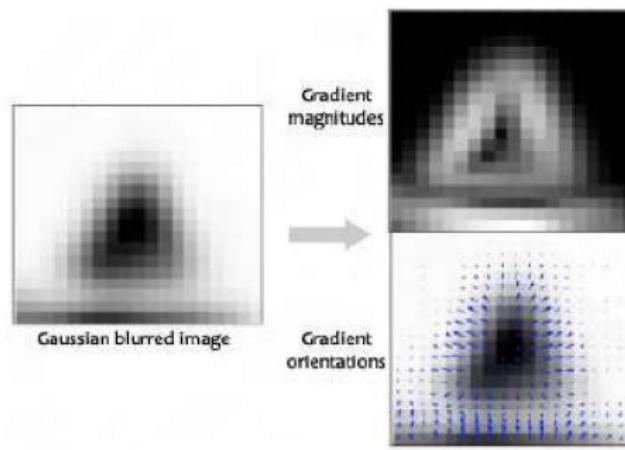


A keypoint

Direzione principale

Assegnare una direzione principale a ciascun punto caratteristico consente di avere descrittori invarianti alla rotazione

Calcola magnitudo e direzione per ciascun punto nello spazio della scala (non nel DoG)

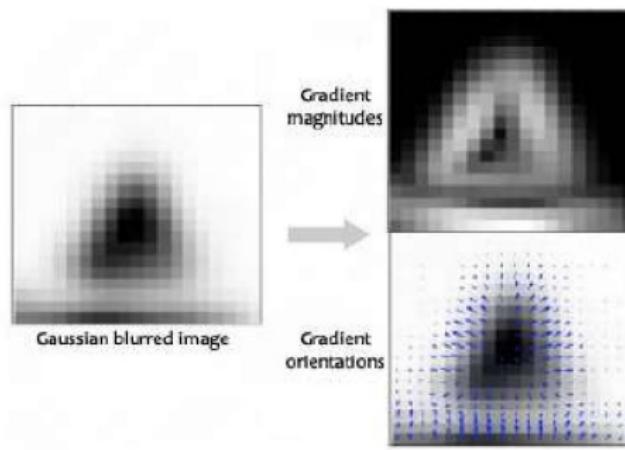


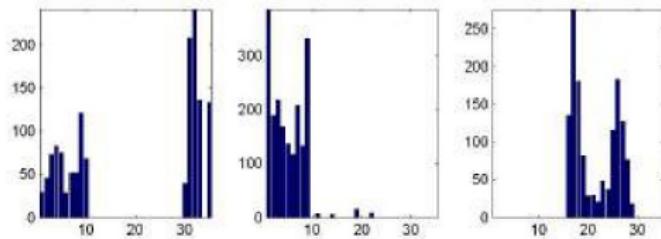
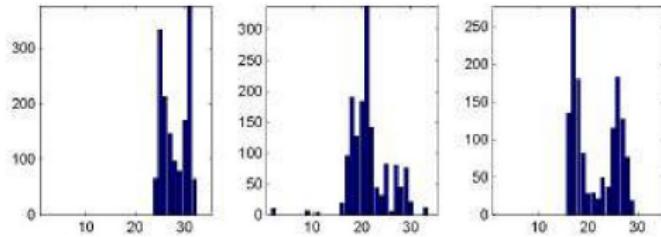
Direzione principale

Assegnare una direzione principale a ciascun punto caratteristico consente di avere descrittori invarianti alla rotazione

Calcola magnitudo e direzione per ciascun punto nello spazio della scala (non nel DoG)

La direzione principale di un punto caratteristico viene determinata come picco nell'istogramma degli orientamenti calcolato per tutti i pixel attorno al punto chiave





Cosa succede se c'è più di un picco?

Cosa succede se c'è più di un picco?

tutti i picchi superiori all'80% del picco più alto vengono convertiti in un nuovo punto chiave

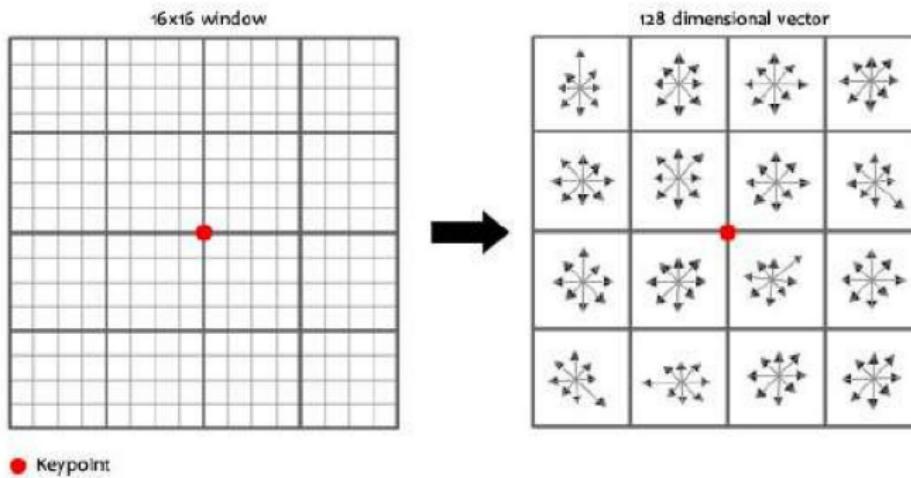
questo nuovo punto chiave ha la stessa posizione e scala dell'originale
il suo

orientamento è uguale all'orientamento dell'altro

picco può dividere un punto chiave in più punti chiave

Descrittore

Finestra 16×16 attorno al punto chiave. Questa finestra è suddivisa in 16 finestre 4×4 .



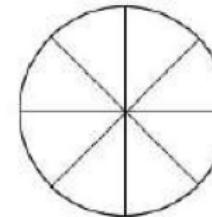
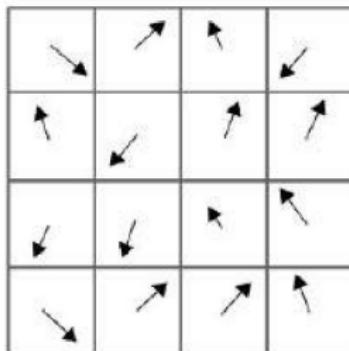
Descrittore

All'interno di ciascuna finestra 4×4 vengono calcolati l'ampiezza e l'orientamento del gradiente.

Questi orientamenti vengono inseriti in un istogramma a 8 bin.

Qualsiasi orientamento del gradiente compreso tra 0 e 44 gradi viene aggiunto al primo contenitore. 45-89 aggiungi al contenitore successivo. E così via.

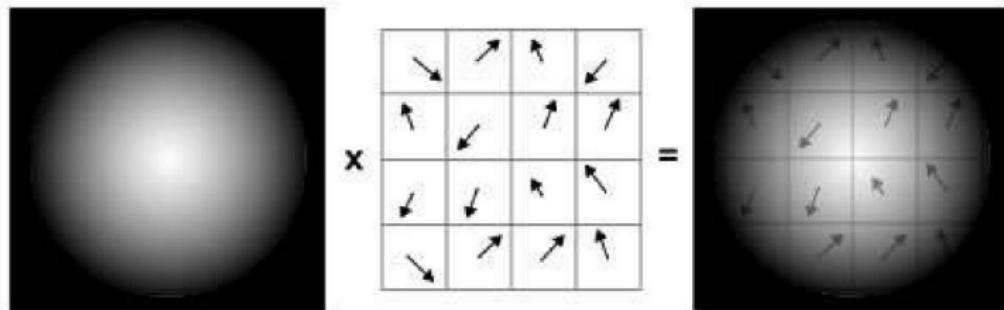
La quantità aggiunta al contenitore dipende dall'entità del gradiente.



Descrittore

La quantità aggiunta dipende anche dalla distanza dal punto chiave.
I gradienti lontani dal punto chiave aggiungeranno valori più piccoli all'istogramma.

Questo viene fatto utilizzando una funzione di ponderazione gaussiana.



Descrittore

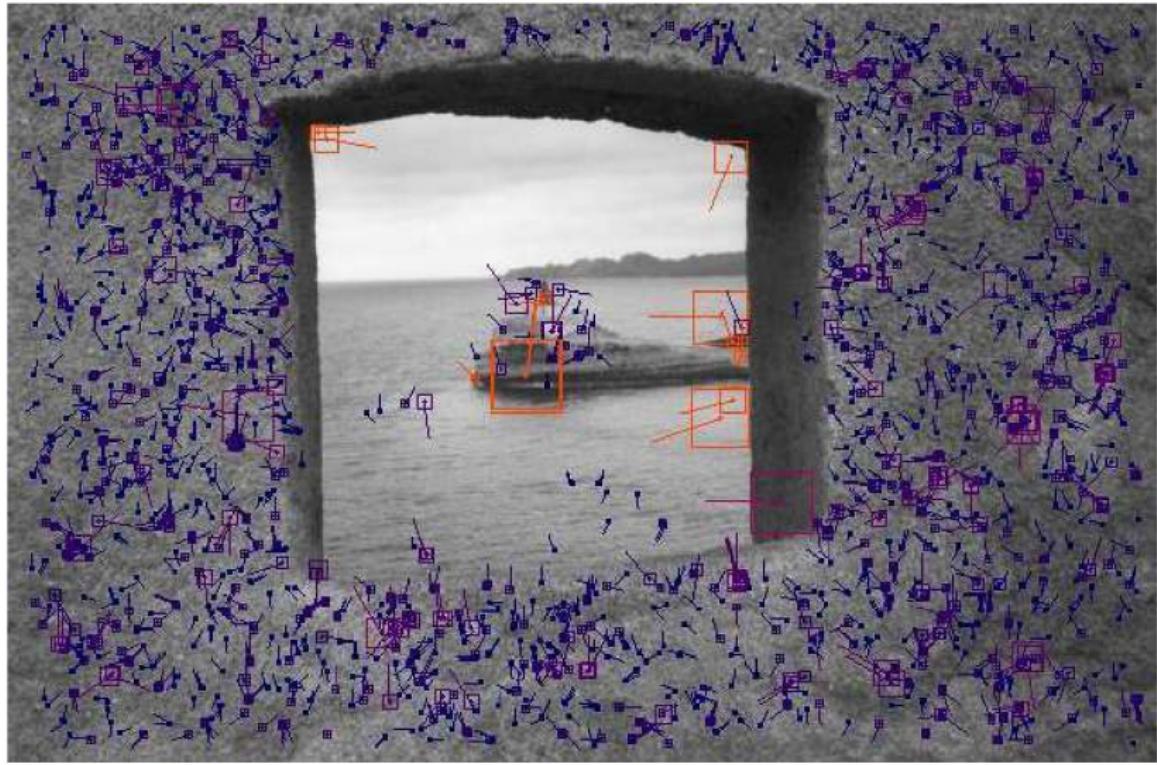
Questo viene fatto per tutte le sedici regioni 4×4 .

Alla fine otteniamo un vettore di caratteristiche di $4 \cdot 4 \cdot 8 = 128$ componenti.

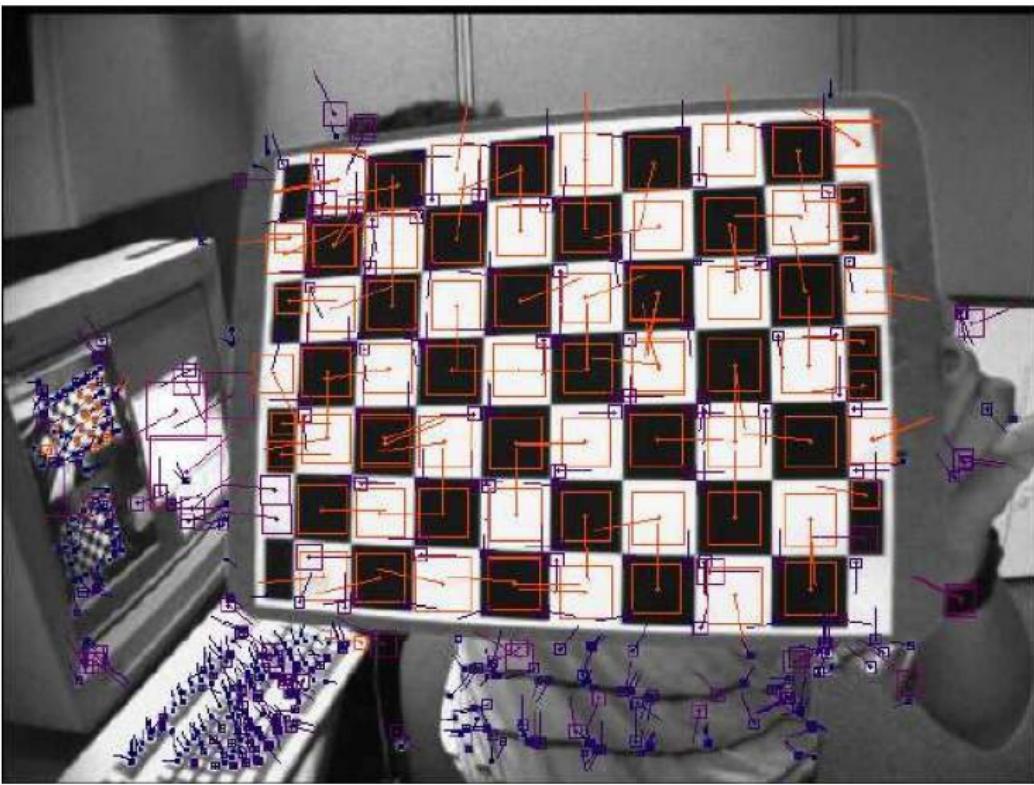
Il vettore delle caratteristiche è normalizzato.

Il punto chiave è identificato in modo univoco da questo vettore di funzionalità.

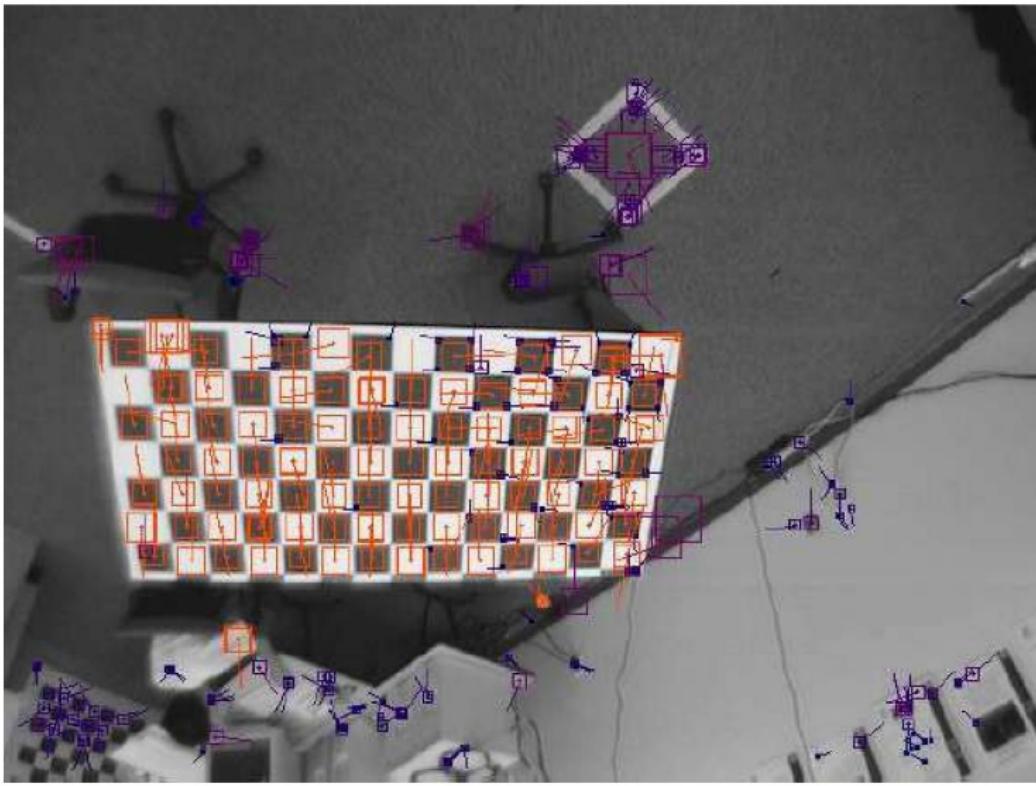
Esempi



Esempi



Esempi



Applicazioni

Localizzazione di robot

per il riconoscimento di oggetti e mosaicatura

della mappatura

Modellazione, riconoscimento e tracciamento di scene 3D

analizzando il cervello umano in immagini di risonanza magnetica 3D

Descrittori 3D SIFT per il riconoscimento delle azioni umane

Riconoscimento degli oggetti

1. Le caratteristiche SIFT vengono ottenute dall'immagine di input 2.

Le caratteristiche vengono abbinate al database delle caratteristiche SIFT

l'abbinamento delle caratteristiche viene effettuato attraverso un approccio
del vicino più vicino basato sulla
distanza euclidea, per cui il rapporto tra la distanza del vicino più vicino e la
distanza del secondo vicino più vicino è maggiore di 0,8 viene rifiutato 3.
raggruppare

le caratteristiche che appartengono allo stesso oggetto e rifiutare le corrispondenze
che vengono escluse le caratteristiche dei

cluster di
trasformazione di Hough che votano per la stessa posa dell'oggetto

4. Determinare l'oggetto utilizzando la trasformazione affine e una probabilità
metodo.







Localizzazione e mappatura dei robot

soluzione al problema della localizzazione dei robot in ambienti sconosciuti

Utilizzare un sistema trinoculare

La posizione 3D dei punti SIFT viene calcolata e il robot si

localizza utilizzando le corrispondenze delle caratteristiche con un 3D esistente

carta geografica

aggiunge in modo incrementale elementi alla mappa

Considerazioni finali sul SIFT

SIFT sono funzionalità robuste.

Relativamente facile da calcolare

Molto lento

Nuove funzionalità simili a SIFT proposte

PCA-SIFT

Funzionalità robusta accelerata (SURF)

Velocità migliorata sfruttando le GPU