

# Template matching



## Il problema del riconoscimento



- Nel contesto dell'interpretazione delle immagini, l'obiettivo centrale è quello di riconoscere gli oggetti presenti all'interno di un'immagine.
- “Riconoscere” un oggetto significa verificare che nell'immagine è presente un'istanza dell'oggetto di cui è memorizzata una rappresentazione.

## Un primo approccio: il template matching



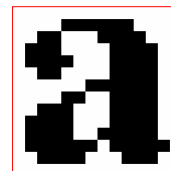
- Il primo e più semplice approccio al problema del riconoscimento nel caso delle immagini è quello di confrontare direttamente l'immagine dell'oggetto cercato con l'immagine in esame.
- Questo tipo di approccio si definisce *template matching*. Si basa sulla misura della similarità esistente tra il prototipo dell'oggetto da riconoscere (*template*) e (una parte dell') immagine.

## Esempio: riconoscimento della lettera 'a' in un testo



**Nel mezzo del cammin di nostra vita  
mi ritrovai in una selva oscura**

Bisogna riconoscere le istanze  
della lettera 'a', di cui si ha a  
disposizione un template



## Realizzazione del template matching



- Siccome non si conoscono a priori le regioni in cui l'istanza può presentarsi, è necessario confrontare il template con tutte le sottoparti dell'immagine che hanno le stesse dimensioni del template.
- A questo scopo, il template viene fatto scorrere sequenzialmente sull'intera immagine, valutando per ogni possibile posizione la similarità tra il template e la regione dell'immagine.

## Realizzazione del template matching

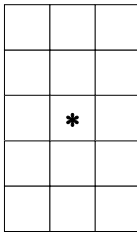


- Si definisca  $T(x,y)$  il template con  $(x,y) \in D_T$ .  
 $D_T$  è il dominio di definizione del template.
- Se  $I(i,j)$  è l'immagine in esame di dimensioni  $NR \times NC$ ,  $T$  è confrontato con tutte le regioni  $I(u+x, v+y)$  con  $(x,y) \in D_T$  tali che:

$$1 \leq u+x \leq NR$$

$$1 \leq v+y \leq NC$$

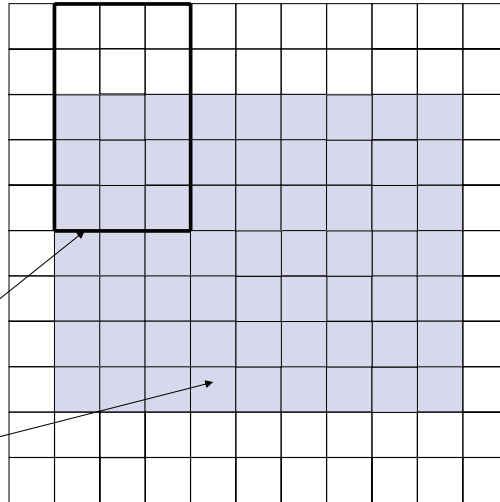
# Realizzazione del template matching



T

Area dell'immagine su cui scorre il centro di T

$I(3+x, 2+y)$



Francesco Tortorella

Teoria e Tecniche di Interpretazione delle Immagini

A.A. 2003/2004

# Confronto template-immagine



- Per ogni shift  $(u,v)$  del template sull'immagine si confronta il template e la regione corrente dell'immagine.
- Per il confronto sono possibili diverse misure di dissimilarità:

- $\max_{D_T} |T(x,y) - I(u+x, v+y)|$

- $\sum_{(x,y) \in D_T} |T(x,y) - I(u+x, v+y)|$

- $\sum_{(x,y) \in D_T} |T(x,y) - I(u+x, v+y)|^2$

Francesco Tortorella

Teoria e Tecniche di Interpretazione delle Immagini

A.A. 2003/2004



## Confronto template-immagine

- Se si utilizza l'ultima espressione, si può ottenere un'importante misura della similarità osservando che:

$$\sum_{(x,y) \in D_T} |T(x,y) - I(u+x, v+y)|^2 =$$
$$\sum_{(x,y) \in D_T} T^2(x,y) + \sum_{(x,y) \in D_T} I^2(u+x, v+y) - 2 \sum_{(x,y) \in D_T} T(x,y)I(u+x, v+y)$$

- Quindi si ha la massima similarità quando si hanno alti valori per  $\sum_{(x,y) \in D_T} T(x,y)I(u+x, v+y)$



## La mutua correlazione

- Il termine  $C_{TI}(u,v) = \sum_{(x,y) \in D_T} T(x,y)I(u+x, v+y)$

si definisce *mutua correlazione* tra T ed I.

- Le posizioni in cui risulta massima la mutua correlazione sono probabili occorrenze del template T.
- Problema: quanto vale il massimo della mutua correlazione ?



## La mutua correlazione

- Il valore della mutua correlazione in effetti dipende dal comportamento locale di  $I$ .
- Infatti, per la disuguaglianza di Schwartz, si ottiene:

$$\sum_{(x,y) \in D_T} T(x,y)I(u+x,v+y) \leq \sqrt{\sum_{(x,y) \in D_T} T^2(x,y)} \cdot \sqrt{\sum_{(x,y) \in D_T} I^2(u+x,v+y)}$$



## La mutua correlazione

- Di conseguenza, una misura più corretta della similarità è data dalla *mutua correlazione normalizzata*:

$$\frac{\sum_{(x,y) \in D_T} T(x,y)I(u+x,v+y)}{\sqrt{\sum_{(x,y) \in D_T} I^2(u+x,v+y)}}$$

## Immagine e template binari



- Quando si considerano immagini e template binari, una misura più efficiente si può ottenere considerando per la correlazione una versione modificata del template  $T'$

$$T'(x,y) = \begin{cases} 1 & \text{se } T(x,y) = 1 \\ -1 & \text{se } T(x,y) = 0 \end{cases}$$

- In questo caso la correlazione raggiunge il massimo valore quando  $T(x,y)$  e  $I(u+x,v+y)$  combaciano perfettamente.

## Immagine e template binari



- Il valore massimo (template e porzione dell'immagine perfettamente coincidenti) è pari al numero  $N_1$  di pixel uguali a 1 nel template.
- Il valore minimo (template e porzione dell'immagine perfettamente speculari) è pari a  $-N_0$ , dove  $N_0$  è il numero di pixel uguali a 0 nel template.
- Se quindi si sommasse  $N_0$  al risultato della correlazione, questa varierebbe tra 0 ed  $A$ , dove  $A = N_0 + N_1$  è il numero di pixel contenuti nel template.



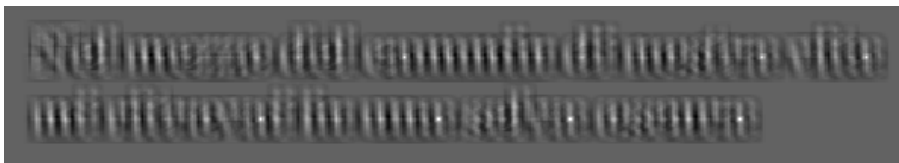
## Template modificato

1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	-1	-1	-1	-1	-1	-1	-1	1	1
1	1	-1	-1	1	1	1	-1	-1	-1	-1	-1	1
1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	-1
1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1
1	1	-1	-1	1	1	1	1	-1	-1	-1	-1	-1
1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1
1	1	1	1	-1	-1	1	1	-1	-1	-1	-1	-1
1	1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1
1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1
1	-1	-1	-1	-1	1	1	-1	-1	-1	-1	-1	-1
1	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	-1	-1
1	1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1
1	1	1	1	1	1	1	1	1	1	1	1	1

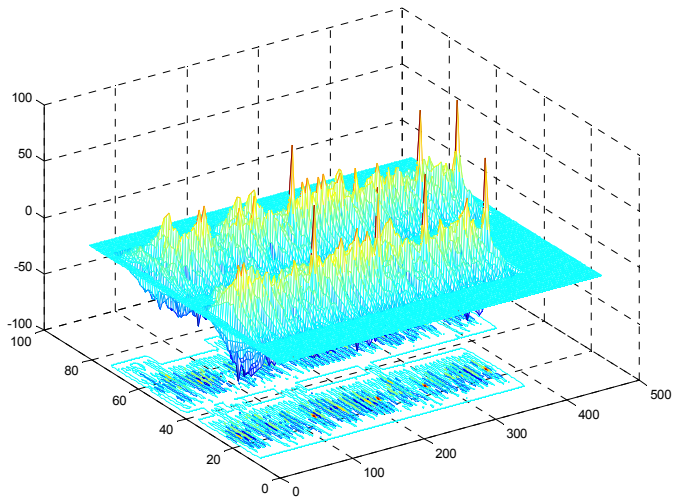


## Esempio

**Nel mezzo del cammin di nostra vita  
mi ritrovai in una selva oscura**







## Problemi del template matching

- scale dependent
- rotation dependent
- non robusto al rumore ed alle variazioni delle istanze
- come si definisce il template ?
  - a mano
  - automaticamente
- come si valuta la soglia ?

