

Problema

Si vogliono confrontare due immagini stabilendo un certo grado di «similitudine»

Si potrebbe pensare ad un confronto diretto fra immagini: ogni immagine con NR righe e NC colonne potrebbe essere semplicemente considerata come un punto in uno spazio multidimensionale.

Un'immagine a livelli di grigio corrisponderebbe quindi a un vettore in \mathbb{R}^n con $n=NR \times NC$.

Il **confronto diretto di immagini**, mediante distanza fra tali vettori nello spazio, in generale **non funziona** per una serie di ragioni, fra cui:

- ☐ Differenze di traslazione, rotazione, scala e prospettiva
- ☐ Deformazione e variabilità delle caratteristiche geometriche dell'immagine
- ☐ Cambiamenti di illuminazione
- ☐ Presenza di rumore nelle immagini e utilizzo di tecniche di acquisizione diverse
- ☐ ...

$$\begin{array}{c} \left\| \begin{array}{cc} \text{img1} & \text{img2} \end{array} \right\| = \left\| \text{diff} \right\| = 4524.84 \\ \left\| \begin{array}{cc} \text{img3} & \text{img4} \end{array} \right\| = \left\| \text{diff} \right\| = 3990.34 \end{array}$$

*Parte degli esempi da:
Raffaele Cappelli,
Fondamenti di
Elaborazione di
immagini
UniBO*

Problema

Si vogliono confrontare due immagini stabilendo un certo grado di «similitudine»

Un primo approccio: **il template matching**

Anziché tentare di confrontare direttamente due immagini, si costruiscono uno o più pattern modello (template) e li si “ricerca” all’interno dell’immagine, misurandone il grado di “somiglianza” (matching) in tutte le possibili posizioni.

Nel contesto dell’interpretazione delle immagini, si vogliono riconoscere gli oggetti presenti all’interno di un’immagine.

Realizzazione del template matching:

Siccome non si conoscono a priori le regioni in cui l’istanza dell’oggetto può presentarsi, è necessario confrontare il template con tutte le sotto-regioni 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.

Si presentano i problemi delle sotto-regioni che comprendono pixel vicini al bordo dell’immagine: come estenderle?



I



T₁



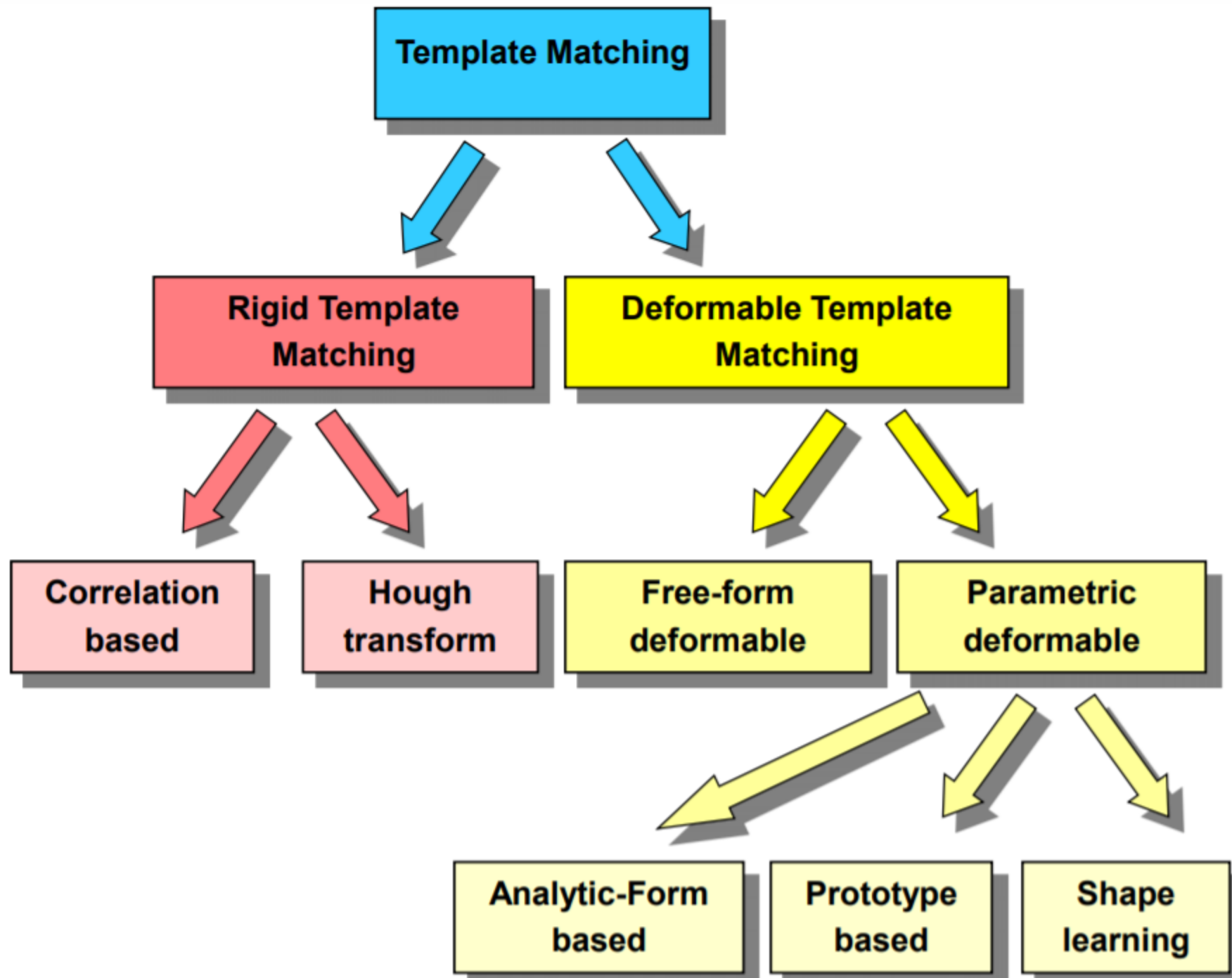
T₂

2 template

Immagine di riferimento

la posizione di massima somiglianza per i
due template all’interno dell’immagine





Esempio.

Si definisca $T(x,y)$ il template con $(x,y) \in D_T$, D_T è il dominio di definizione del template. Se $A(i,j)$ è l'immagine in esame di dimensioni $NR \times NC$, T è confrontato con tutte le regioni $A(u+x, v+y)$ con $(x,y) \in D_T$ tali che: $1 \leq u+x \leq NR$, $1 \leq v+y \leq NC$.

Utilizziamo la mutua correlazione tra T ed A :

$$C_{T,A} = \sum_{(x,y) \in D_T} T(x,y) A(u+x, v+y)$$

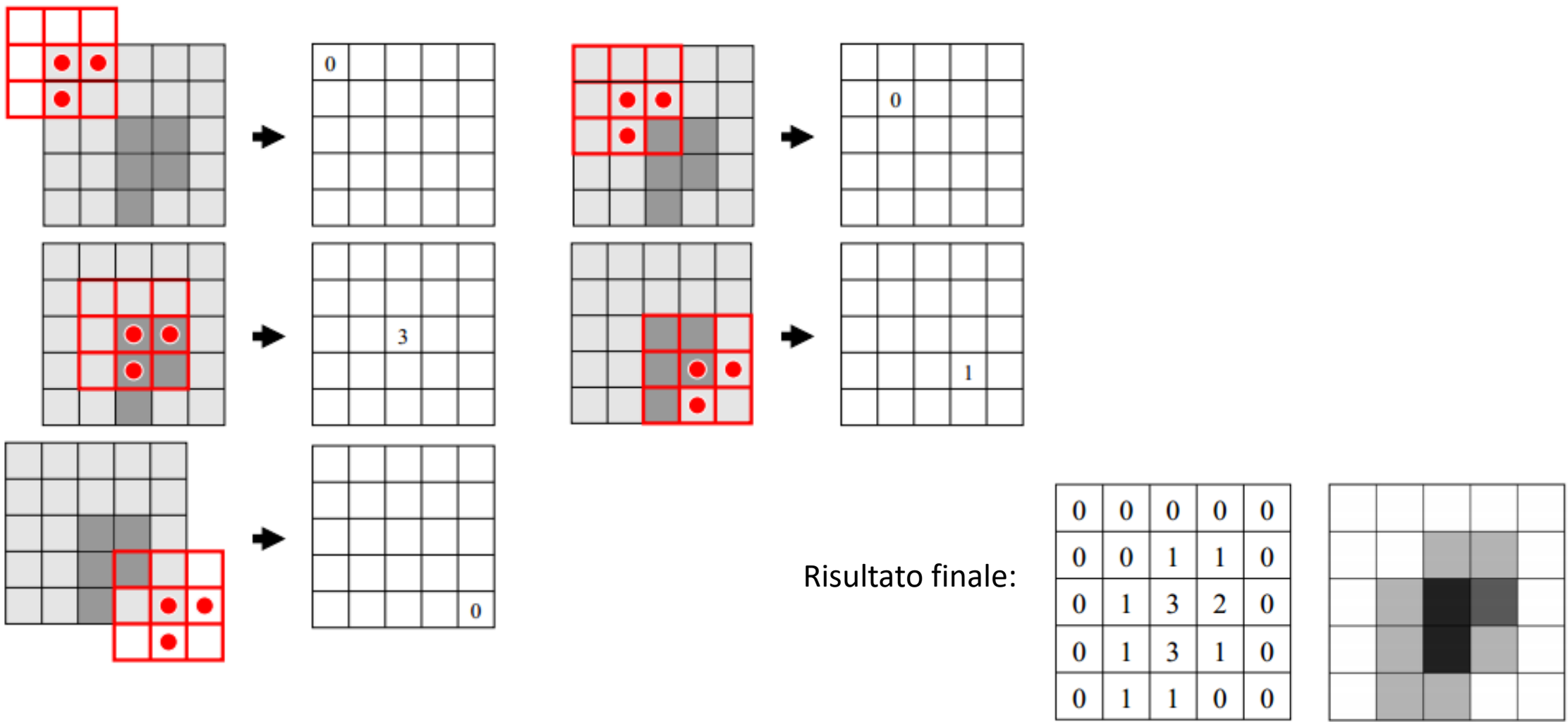
Le posizioni in cui risulta massima la mutua correlazione sono probabili occorrenze del template T .

Quanto vale il massimo della mutua correlazione ?

Quanto dipende dalla scalatura (illuminazione) o dal rumore?

Esempio.

i prodotto di mutua correlazione per immagini 2D, condizioni al bordo nulle (cropping)
Per semplicità si è scelta una immagine binaria di partenza, area scura, valore pixel = 1;
area chiara, valore pixel = 0.



Il valore della mutua correlazione in dipende dal comportamento locale dell'immagine A. Infatti, per la disuguaglianza di Cauchy-Schwartz, si ottiene (si utilizza l'usuale prodotto scalare e norma euclidea):

$$\left| \sum_{(x,y) \in D_T} T(x,y) A(u+x, v+y) \right| \leq \sqrt{\sum_{(x,y) \in D_T} T^2(x,y)} \sqrt{\sum_{(x,y) \in D_T} A^2(u+x, v+y)}$$

Quindi, se le norme a destra della disuguaglianza non sono nulle, una misura più corretta della similarità è data dalla mutua correlazione normalizzata:

$$0 \leq \frac{\left| \sum_{(x,y) \in D_T} T(x,y) A(u+x, v+y) \right|}{\sqrt{\sum_{(x,y) \in D_T} T^2(x,y)} \sqrt{\sum_{(x,y) \in D_T} A^2(u+x, v+y)}} \leq 1$$

Abbiamo inoltre che il valore è massimo (=1) nel caso in cui $T(x,y)$ e $A(u+x, v+y)$ (con $(x,y) \in D_T$) sono, pensati vettori, paralleli: hanno la stessa forma.

NOTA

Esistono diverse proposte di misure di similarità basate sia sui valori dei pixel che del gradiente (discreto).

Martin, J. and Crowley, J. (1995). Experimental comparison of correlation techniques. In Proc. Int. Conf. On Intelligent Autonomous Systems, volume 4, pages 86–93.

Fitch, A. J., Kadyrov, A., Christmas, W. J., and J, K. (2002). Orientation correlation. In Rosin, P. and Marshall, D., editors, British Machine Vision Conference, volume 1, pages 133–142.

Crouzil, A., Massip-Pailhes, L., and Castan, S. (1996). A new correlation criterion based on gradient fields similarity. In Proc. Int. Conf. Pattern Recognition (ICPR), pages 632–636.

.....

La mutua correlazione è utilizzata per il **Template matching** che è la forma più semplice per la ricerca di oggetti in una immagine.

Si parte da una data immagine (**template**) di un oggetto e la si “sposta” su un’altra immagine calcolando quanto è simile l’oggetto alla porzione di immagine selezionata. Massimi locali di una misura di similarità sono i candidate migliori per il **match**.

Una variazione che produce un’altra misura di similarità, la **Normalized cross correlation**

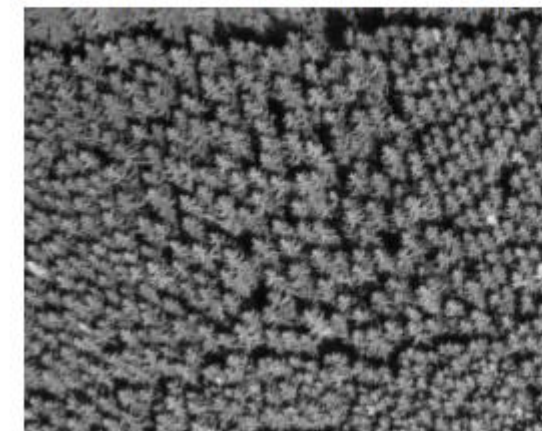
$$NCC = \frac{|\sum_{(x,y) \in D_T} (T(x,y) - \mu(T)) (A(u+x, v+y) - \mu_{u,v}(A))|}{\sqrt{\sum_{(x,y) \in D_T} (T(x,y) - \mu(T))^2} \sqrt{\sum_{(x,y) \in D_T} (A(u+x, v+y) - \mu_{u,v}(A))^2}}$$

Dove $\mu(T)$ è la media aritmetica dei valori di T e $\mu_{u,v}(A)$ è la media aritmetica dei valori dell’immagine individuati dagli indici $(u+x, v+y)$.
In questo modo ho invarianza per variazioni affini dell’intensità
 $\text{luminosa} \propto T + \beta$

Esempio, localizzare automaticamente degli alberi
In una foto aerea (da *Morten Larsen*, Department of Basic Sciences and Environment Mathematics and computer science group)

Template matching example: Trees

Locating tree tops in aerial photography.

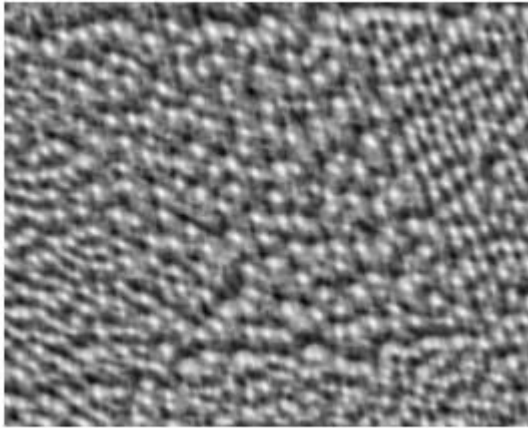


Model



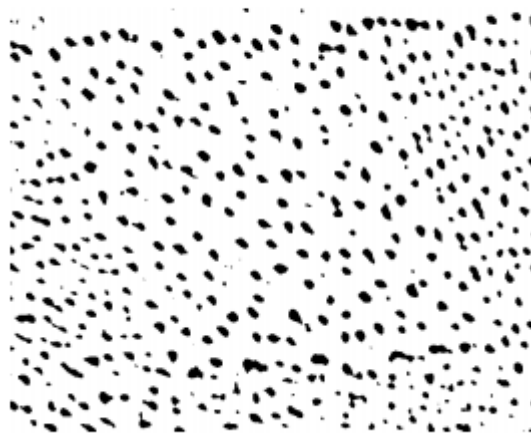
Sampled

Step 1: Similarity to template



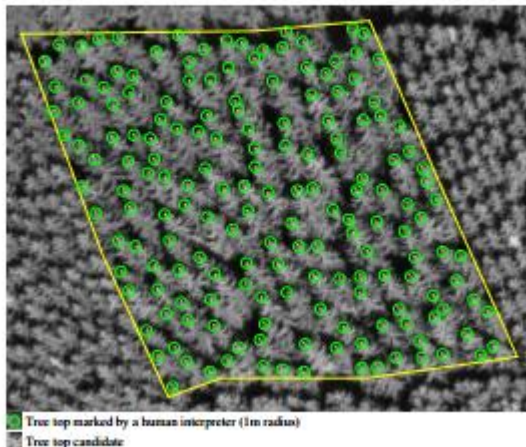
(Normalized cross correlation)

Step 2: Threshold for high similarity

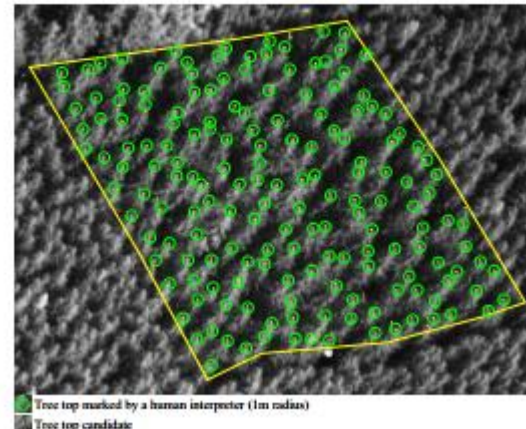


Step 3: Local maxima

Results:



Other angles (II)



Alcune limitazioni:

- Non posso considerare scaling o rotazioni del template.
- Devo utilizzare una libreria di template.
- Templates possono essere difficili da generare automaticamente

Esercizio.

Utilizzando la stessa struttura del programma di convoluzione implementare il calcolo della Normalized cross correlation.

Attenzione: il template può avere dimensioni qualsiasi, per cui per l'estensione dell'immagine al bordo occorre decidere come valutare il pixel centrale e quindi come estendere nella direzione orizzontale e verticale.

Si applichi per la ricerca automatica del numero di volte che appare il simbolo quadri in una carta da gioco.

