

Metodi innovativi per il riconoscimento
automatico dello stile fumettistico

Flavio Neri

19 marzo 2008

Indice

1	Computer Vision	5
1.1	Approccio al riconoscimento dello stile	6
1.2	Tentativi precedenti	6
1.3	Funzioni misuranti	8
2	Funzioni di Taglia: Definizioni e Proprietà	11
2.1	Definizioni delle Funzioni di Taglia Ridotte	12
2.2	Prime proprietà	13
2.3	Rappresentazione algebrica delle Funzioni di Taglia	14
2.4	Distanza tra Funzioni di Taglia	17
2.4.1	Distanza di Hausdorff	18
3	Insiemi Fuzzy	21
3.1	Richiami classici	22
3.2	Il concetto di sottoinsieme fuzzy	23
3.3	Assegnazione	26
4	Esperimenti	27
4.1	Dati e strumenti utilizzati	27
4.2	Esperimenti preliminari	28
4.3	Filtro Gaussiano	32
4.4	Assegnazione fuzzy	33
4.5	Gruppi di funzioni misuranti	35
4.6	Convoluzione	36
4.7	Risultati Finali	37
4.8	Conclusioni	40

Appendice	40
Bibliografia	47

Capitolo 1

Computer Vision

Lo scopo della Computer Vision è sviluppare algoritmi e metodologie per la riproduzione automatica o semi-automatica di meccanismi tipici dei sistemi di visione biologici, come ad esempio il nostro sistema visivo. Questi comprendono il riconoscimento di forme e di colori, l'estrazione di informazioni spaziali e geometriche, o la capacità di segmentare ed inseguire oggetti in movimento. La Computer Vision è forse uno dei problemi delle scienze informatiche di più difficile investigazione poichè si sa che può essere risolto, ma non si sa come risolverlo. Infatti per risolvere un generico problema di computer vision dovremmo prima rispondere alle questioni fondamentali sulla visione e elaborazione di immagini nell'intelligenza umana. I problemi di Computer Vision spaziano in diversi campi: estrapolazione di informazioni da immagini di qualunque genere come ecografie per la diagnosi di tumori o volti umani per il riconoscimento di persone, controllo automatizzato per il rilevamento di difetti di fabbrica in catene industriali, applicazioni militari di rilevamento missili, sistemi di guida automatizzati, e ancora altri. Una piccola ma importante branca della Computer Vision è sicuramente quella che si occupa del riconoscimento delle immagini bidimensionali, o meglio quella che cerca di insegnare al computer a distinguere le varie immagini in base al loro contenuto. Un classico esempio di questo metodo, che prende il nome di Pattern Recognition, è quello di insegnare al computer a distinguere immagini dove sono presenti volti di persone da quelli dove non sono presenti, o in altri casi quello di riconoscere oggetti specifici all'interno di un'immagine

qualsiasi. Tutti questi studi hanno il fine di fornire all'utente un sistema in grado di recuperare immagini da grossi database, come per esempio il Web o raccolte di marchi di fabbrica, senza appoggiarsi alla semplice ricerca per titolo o autore dell'immagine, ma ad una ricerca intuitiva e versatile.

1.1 Approccio al riconoscimento dello stile

Quello che abbiamo fatto in questa tesi è affrontare un nuovo e inesplorato problema di Computer Vision. Se molti studi sono già bene avviati sull'analisi della dissimiglianza tra diverse immagini (Shape Matching) quello che ancora è poco esplorato è studiare la dissimiglianza tra i produttori di queste immagini. Se nello studio delle forme ha un peso fondamentale la struttura topologica e geometrica dell'immagine, nel caso del riconoscimento di autori è fondamentale lo stile che essi adottano. Tuttavia se per discriminare due diverse forme geometriche bidimensionali possiamo affidarci a metodi quantitativi e oggettivi (un quadrato è diverso da un triangolo perchè ha 4 lati e non 3 oppure il cerchio è differente dalla corona circolare perchè quest'ultima ha un buco), per riconoscere lo stile di un autore la situazione è più complicata poichè non ci sono discriminati assoluti per poter fare una classificazione. Ciò che caratterizza lo stile di un fumettista è per esempio la struttura delle tavole, la luminosità, l'uso dei balloons e il littering, la continuità del tratto, le sfumature o il tratteggio e altri ancora. In questo lavoro abbiamo provato a considerare tre diversi aspetti che fossero facilmente quantificabili e quindi digeribili da un calcolatore: la luminosità (quantificabile col numero di pixel scuri), il tratteggio (quantificabile con i salti da pixel chiari a pixel scuri), la geometricità (quantificabile con le linee rette presenti).

1.2 Tentativi precedenti

Questo lavoro consiste nella naturale prosecuzione di un metodo già trattato in [4], pur seguendo il procedimento di fondo useremo nelle varie fasi strumenti diversi, infine faremo un raffronto tra gli esperimenti fatti qui e quelli fatti in [4]. La risoluzione del problema si svolge in tre fasi principali:

trasformazione dell'immagine in un istogramma contenente le informazioni sugli aspetti citati sopra (luminosità, tratteggio e geometricità), costruzione di una metrica (o pseudo-metrica) sulle immagini, assegnazione dell'immagine al suo autore. Le novità di questa tesi rispetto a [4] si concentrano fondamentalmente nelle ultime due fasi, in particolare nella costruzione della metrica abbiamo introdotto le *funzioni di taglia* e nella fase di assegnazione abbiamo introdotto le funzioni di *appartenenza fuzzy*. Il motivo per cui si è pensato di adoperare le funzioni di taglia è che queste sono particolarmente predisposte a misurare l'andamento di un istogramma piuttosto che l'area coperta. D'altraparte si è pensato di unificare i metodi di assegnazione proposti da [4] in un unico metodo, il metodo fuzzy appunto, ritenuto adatto ad estrapolare le informazioni più significative da un insieme di risultati. Definiremo formalmente questi strumenti nei prossimi capitoli. Di seguito invece descriviamo le funzioni usate da [4] e di cui abbiamo fatto uso anche noi nella prima fase del lavoro, relazioneremo in seguito dei problemi pratici incontrati. Precisiamo infine che i dati di partenza su cui abbiamo condotto la sperimentazione sono gli stessi di [4] ovvero un database di 109 immagini digitali di 4 autori diversi (Angelo, Costantini, Sacco, Palumbo collezione Scotellaro). Le immagini sono ridimensionate in un formato standard 800 X 600 e convertite in bianco/nero. La scelta delle operazioni di preprocessing grafico è molto importante in ogni test di Computer Vision tuttavia si è preferito non modificare i dati iniziali di [4] per avere in ultima istanza un confronto verosimile dei risultati.

1.3 Funzioni misuranti

Con funzioni misuranti intendiamo le funzioni che ci permettono di passare dalle immagini agli istogrammi, le prime 4 da cui si è partiti sono:

1. *quantonero* conta il numero di pixel neri lungo una data direzione per fasce di spessore 8 pixel;
2. *quantonero2* analogo a *quantonero* ma lo spessore della fascia è di un solo pixel;
3. *salti* conta il numero di passaggi da pixel bianco a pixel nero (e viceversa) lungo una data direzione;
4. *linee* conta le linee rette di una data lunghezza mediante la trasformata di Hough[6].

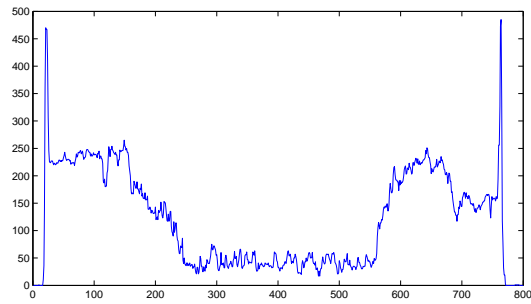
A partire da queste se ne sono create altre 4 considerando le frequenze di queste, ossia misurando con quale occorrenza si ripetono i valori nelle colonne degli istogrammi ottenuti dalle 4 funzioni precedenti, rispettivamente. L'utilizzo di queste ultime funzioni (*fquantonero*, *fquantonero2*, *fsalti*, *flinee*), ha permesso di ricavare informazioni da un punto di vista differente da quello più classico e diretto delle prime 4; i risultati che ne derivano sono in alcuni casi molto interessanti, in altri invece un po' patologici come vedremo.

Infine l'ultimo gruppo di funzioni utilizzate sono le fasce di frequenza, ottenute analogamente alle precedenti considerando un range di valori invece che un singolo valore di una colonna; si hanno quindi le nuove funzioni *ffquantonero*, *ffquantonero2*, *ffsalti*, *fflinee*.

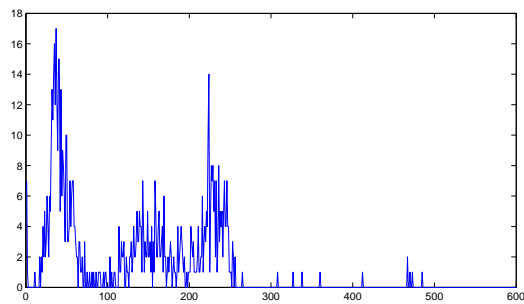
Queste che può sembrare una semplice variante delle precedenti hanno invece portato risultati significativi.



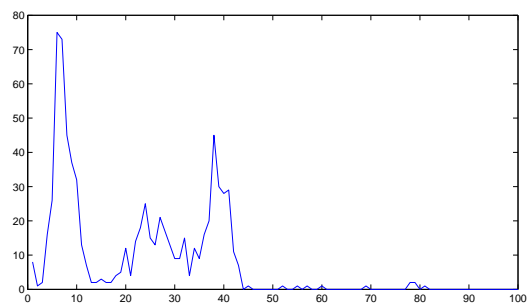
(a) Angelo(1).jpg



(b) Istogramma di *quantonero2*



(c) Istogramma di *fquantonero2*



(d) Istogramma di *ffquantonero2*

Capitolo 2

Funzioni di Taglia: Definizioni e Proprietà

Le Funzioni di Taglia sono uno strumento topologico-geometrico per studiare le forme. Esse ci forniscono informazioni sugli aspetti topologici integrandole con qualche tipo di misura sull'immagine. Più precisamente le misure sono effettuate mediante una funzione reale detta *funzione misurante*, definita su uno spazio topologico \mathcal{M} che si intende studiare (usualmente è un sottospazio di qualche spazio euclideo). La scelta della funzione misurante da associare a \mathcal{M} è arbitraria, pertanto è possibile scegliere la funzione misurante più opportuna al problema. Se si cambiano le funzioni misuranti, le corrispondenti Funzioni di Taglia generano diversi descrittori della data forma che si intende studiare. Un aspetto fondamentale delle Funzioni di Taglia è che ereditano le proprietà di invarianza, se tali proprietà esistono, dalle funzioni misuranti scelte. Queste proprietà delle Funzioni di Taglia può tornare molto utile nelle applicazioni di *shape-matching*, nelle quali si è spesso interessati a confrontare forme indipendentemente dall'azione di determinati gruppi di trasformazioni. Inoltre è stato dimostrato che le Funzioni di Taglia possono essere resistenti al rumore e all'occlusione; riguardo questo aspetto ha una importanza rilevante la scelta della metrica tra Funzioni di Taglia. La natura di \mathcal{M} fa sì che qualunque insieme che può essere modellato mediante uno spazio topologico potrebbe essere rappresentato mediante Funzioni di Taglia, per esempio immagini binarie, in scala di grigio, a colori,

ma anche onde sonore o elettromagnetiche. Diamo ora nelle prossime pagine la definizione formale di Funzioni di Taglia e il riassunto delle principali proprietà.

2.1 Definizioni delle Funzioni di Taglia Ridotte

Fino ad ora abbiamo parlato sempre di Funzioni di Taglia e non di Funzioni di Taglia Ridotte. La differenza tra le due sta nel fatto che le ultime, pur avendo le stesse fondamentali proprietà a noi necessarie, sono più maneggevoli da usare nelle dimostrazioni di certi risultati e hanno un dominio di definizione più semplice e comodo da usare. La definizione di Funzione di Taglia Ridotta è quella più recente e se ne può trovare un approfondimento in [1]. D'ora in poi assumeremo che siano dati uno spazio topologico \mathcal{M} e una funzione $\varphi : \mathcal{M} \rightarrow \mathbb{R}$; (\mathcal{M}, φ) sarà chiamata *coppia di taglia*.

Definizione 2.1.

Per ogni $x \in \mathbb{R}$ definiamo l'insieme $\mathcal{M} \langle \varphi \leq x \rangle = \{P \in \mathcal{M} : \varphi(P) \leq x\}$.

Definizione 2.2.

$\forall y \in \mathbb{R}$ definiamo la relazione $\cong_{\varphi \leq y}$ in \mathcal{M} ponendo $P \cong_{\varphi \leq y} Q \iff P = Q$ oppure P, Q appartengono alla stessa componente connessa di $\mathcal{M} \langle \varphi \leq y \rangle$.

Osservazione 1. È facile mostrare che $\cong_{\varphi \leq y}$ è una relazione di equivalenza su \mathcal{M} , $\forall y \in \mathbb{R}$.

Definizione 2.3. Sia $\Delta^+ = \{(x, y) \in \mathbb{R}^2 : 0 \leq x < y\}$. Consideriamo la funzione $\ell_{(\mathcal{M}, \varphi)} : \Delta^+ \rightarrow \mathbb{N}$ definita ponendo $\ell_{(\mathcal{M}, \varphi)}(x, y)$ uguale al numero di classi di equivalenza in cui $\mathcal{M} \langle \varphi \leq x \rangle$ è divisa dalla relazione di equivalenza $\cong_{\varphi \leq y}$. Tale funzione si chiamerà *Funzione di Taglia Ridotta* associata alla coppia (\mathcal{M}, φ) .

Osservazione 2. L'interpretazione geometrica di $\ell_{(\mathcal{M}, \varphi)}(x, y)$ è il numero di componenti connesse di $\mathcal{M} \langle \varphi \leq y \rangle$ contenenti almeno un punto di $\mathcal{M} \langle \varphi \leq x \rangle$.

Esempio 1. Come esempio di Funzione di Taglia Ridotta consideriamo \mathcal{M} la curva in figura 2.1 e φ è la funzione che associa ad ogni punto la sua distanza dal centro di massa O. A destra è rappresentata $\ell_{(\mathcal{M}, \varphi)}(x, y)$.

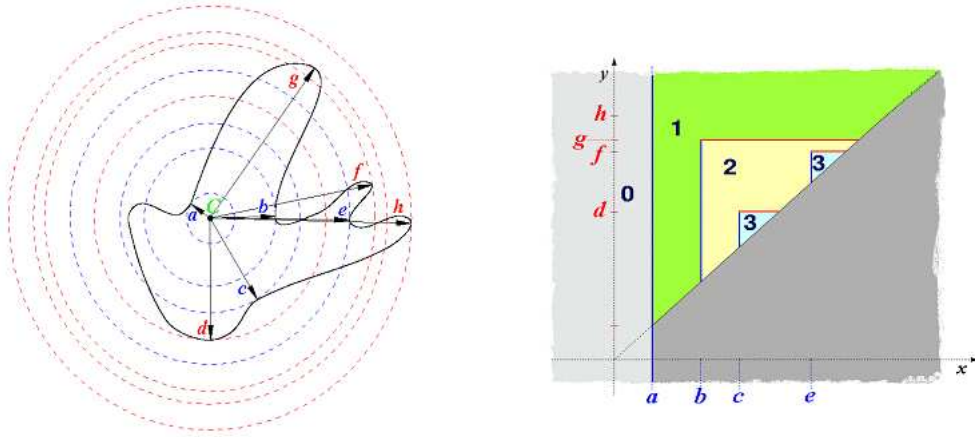


Figura 2.1: Esempio di una Funzione di Taglia

2.2 Prime proprietà

Elenchiamo di seguito alcune osservazioni che seguono dalla definizione delle Funzioni di Taglia Ridotte.

1. $\ell_{(\mathcal{M}, \varphi)}(x, y)$ è non decrescente in x non crescente in y .
2. $\ell_{(\mathcal{M}, \varphi)}(x, y) = 0 \forall x < \min_{P \in \mathcal{M}} \varphi(P)$.
3. $\forall y \geq \max_{P \in \mathcal{M}} \varphi(P)$, $\ell_{(\mathcal{M}, \varphi)}(x, y)$ è uguale al numero di componenti connesse \mathcal{N} di \mathcal{M} tali che $x \geq \min_{P \in \mathcal{N}} \varphi(P)$.

Un'altra importante proprietà delle Funzioni di Taglia è quella della resistenza al rumore. Da un punto di vista teorico esistono alcuni teoremi che ci danno stime degli errori date da alcuni tipi di perturbazioni; qui ci limitiamo a mostrare (fig 2.2 e 2.3) in che forma appare il rumore in una Funzione di Taglia e osservare come sia facilmente localizzabile.

Osserviamo come i due punti di discontinuità più lontani dalla diagonale rimangono pressoché invariati mentre nel grafico relativo all'immagine con rumore appaiono altri punti localizzati vicino alla diagonale. Possiamo dire che maggiore è la distanza di un punto di discontinuità dalla diagonale, più macroscopica è la caratteristica dell'immagine che lo ha generato, viceversa

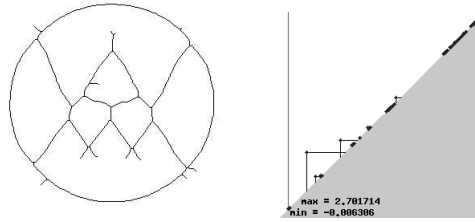


Figura 2.2: A sinistra l'immagine in assenza di rumore e a destra la relativa Funzione di Taglia

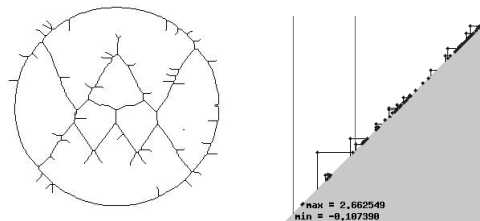


Figura 2.3: A sinistra l'immagine con presenza di rumore e a destra la relativa Funzione di Taglia

tanto più è vicino un punto alla diagonale tanto più fine è il dettaglio che lo ha generato. La retta verticale che appare nella seconda Funzione di Taglia ci informa sulla presenza di una seconda componente connessa che non è altro che un punto (o meglio pixel) isolato.

2.3 Rappresentazione algebrica delle Funzioni di Taglia

Una Funzione di Taglia Ridotta per sua natura è sempre rappresentabile come una funzione da $\Delta^+ \subseteq \mathbb{R}^2$ a valori in \mathbb{N} , tuttavia abbiamo già fatto notare in precedenza come i punti di discontinuità contengano la maggior parte delle informazioni di una Funzione di Taglia Ridotta. Di seguito mostreremo come è possibile rappresentare una Funzione di Taglia Ridotta mediante un oggetto matematico più maneggevole come le serie formali. Questo può

essere fatto perchè nel caso in cui \mathcal{M} sia sufficientemente regolare, per esempio compatto, allora i punti di discontinuità che appartengono alla regione $\Delta^+ = \{(x, y) \in \mathbb{R}^2 : x < y\}$ dividono Δ^+ in regioni triangolari sovrapposte aventi un lato sulla diagonale $\{(x, y) \in \mathbb{R}^2 : x = y\}$. In generale i triangoli possono avere area finita o infinita. Convenzionalmente identifichiamo ogni triangolo finito con un punto e ogni triangolo infinito con una retta verticale. Più precisamente diamo le seguenti definizioni.

Definizione 2.4. Per ogni punto $p = (x, y) \in \Delta^+$, definiamo il numero $\mu(p)$ come il minimo, su tutti gli ϵ positivi tali che $x + \epsilon < y - \epsilon$, di

$$\ell_{(\mathcal{M}, \varphi)}(x + \epsilon, y - \epsilon) - \ell_{(\mathcal{M}, \varphi)}(x - \epsilon, y - \epsilon) - \ell_{(\mathcal{M}, \varphi)}(x + \epsilon, y + \epsilon) + \ell_{(\mathcal{M}, \varphi)}(x - \epsilon, y + \epsilon).$$

Il numero finito $\mu(p)$ è chiamata molteplicità di p per $\ell_{(\mathcal{M}, \varphi)}$. Se $\mu(p)$ è strettamente positivo allora p si dice *punto angolare*.

Definizione 2.5. Per ogni retta verticale $r : x = k$ con $k \in \mathbb{R}$, definiamo il numero $\mu(k)$ come il minimo, su tutti i numeri reali ϵ tali che $k + \epsilon < 1/\epsilon$, di

$$\ell_{(\mathcal{M}, \varphi)}(k + \epsilon, 1/\epsilon) - \ell_{(\mathcal{M}, \varphi)}(k - \epsilon, 1/\epsilon).$$

Quando questo numero finito, chiamato *molteplicità di r* per $\ell_{(\mathcal{M}, \varphi)}$, è strettamente positivo allora r si dice *retta angolare*.

L'importanza dei punti e rette angolari sta nel fatto che il valore della funzione di taglia è, su quasi tutti i punti sopra la diagonale, la somma delle molteplicità $\mu(p_i)$ e $\mu(r_i)$ dove p_i e r_i identificano i triangoli, finiti o infiniti, contenenti questi punti. Pertanto, essendo possibile descrivere una funzione di taglia con la collezione dei suoi punti angolari e delle linee angolari con le rispettive molteplicità, possiamo rappresentare una funzione di taglia con una serie formale di punti e rette del piano. Questa rappresentazione riduce la complessità di operare con funzioni di taglia traducendo il problema in operazioni tra serie formali, cioè un problema algebrico. Vediamo come sono fatte queste serie formali; innanzitutto enunciamo il seguente teorema che ci da il valore di $\ell_{(\mathcal{M}, \varphi)}(x, y)$.

Teorema 2.3.1. Per ogni $(\bar{x}, \bar{y}) \in \Delta^+$ si ha:

$$\ell_{(\mathcal{M}, \varphi)}(\bar{x}, \bar{y}) = \sum_{(x, y) \in \Delta^+, x \leq \bar{x}, y > \bar{y}} \mu((x, y)).$$

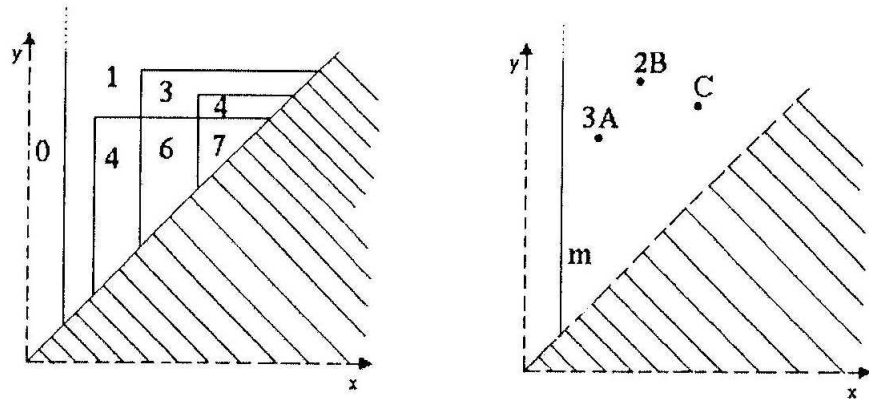


Figura 2.4: Punti e rette angolari di una funzione di taglia ridotta: A,B,C sono punti angolari, rispettivamente di molteplicità 3, 2, 1. m è l'unica retta angolare.

Potrebbe succedere che la funzione di taglia abbia una infinità numerabile di punti di discontinuità; tuttavia questi punti dovranno necessariamente accumularsi sulla diagonale Δ . D'altra parte se la varietà \mathcal{M} è connessa, allora ogni funzione di taglia avrà una e una sola retta angolare. Descriviamo ora l'oggetto che conterrà tutte le informazioni di una Funzione di Taglia.

Definizione 2.6. Sia ℓ una Funzione di Taglia ridotta. Chiamiamo *Successione di Rappresentazione* di ℓ la successione $a : \mathbb{N} \rightarrow \Delta^+$ che gode delle seguenti proprietà:

1. a_0 è una retta angolare per ℓ ;
2. $\forall i > 0$, o a_i è un punto angolare per ℓ , o a_i sta sulla diagonale Δ ;
3. Se p è un punto angolare per ℓ di molteplicità $\mu(p)$, allora la cardinalità dell'insieme $\{i \in \mathbb{N} : a_i = p\}$ è uguale a $\mu(p)$;
4. L'insieme di indici per cui $a_i \in \Delta$ ha una cardinalità al più numerabile.

Nell'esempio in figura 2.4, la successione di rappresentazione di ℓ è:
 $a_0 = m, a_1 = A, a_2 = A, a_3 = A, a_4 = B, a_5 = B, a_6 = C$ e $a_i \in \Delta \forall i > 6$.

2.4 Distanza tra Funzioni di Taglia

In questo paragrafo descriveremo le due metriche sullo spazio delle Funzioni di Taglia più standard, ricordando che esiste già una vasta antologia di metriche approfonditamente studiate e utilizzate come quella di *p-matching*. Le metriche sullo spazio delle Funzioni di Taglia si distinguono a seconda della rappresentazione usata:

- funzioni da $\Delta^+ \subseteq \mathbb{R}^2 \rightarrow \mathbb{N}$,
- sottoinsieme di punti di Δ^+ .

Due proprietà importanti per una metrica sullo spazio delle Funzioni di Taglia sono:

- resistenza al rumore,
- consistenza.

La prima è verificata quando perturbando la funzione misurante φ è possibile controllare la variazione della Funzione di Taglia sotto l'azione delle perturbazione. La consistenza è in un certo senso la proprietà inversa: data la distanza tra due Funzioni di Taglia è possibile modificare le relative funzioni misuranti in maniera che le nuove Funzioni di Taglia coincidano.

Formalmente:

Definizione 2.7. Sia d una metrica nello spazio delle Funzioni di Taglia. Diremo che d è resistente al rumore se esiste c in \mathbb{R}^+ tale che per ogni varietà \mathcal{M} e per ogni funzione misurante $\tilde{\varphi}$ risulta

$$d(\ell_{(\mathcal{M},\varphi)}, \ell_{(\mathcal{M},\tilde{\varphi})}) \leq c \max_{m \in \mathcal{M}} |\varphi(m) - \tilde{\varphi}(m)|$$

Se pensiamo le Funzioni di Taglia come usuali funzioni da $\Delta^+ \subseteq \mathbb{R}^2 \rightarrow \mathbb{N}$ allora ogni metrica o pseudometrica sullo spazio delle funzioni da $\Delta^+ \rightarrow \mathbb{N}$ induce una metrica sullo spazio delle Funzioni di Taglia. L'esempio più antico è la *norma- \mathcal{L}^p* . Tuttavia queste metriche non si sono mai rivelate estremamente valide, infatti non godono di nessuna delle due proprietà citate. Descriveremo ora più in dettaglio un tipo di metrica resistente al rumore, basata sulla rappresentazione delle Funzioni di Taglia come sottoinsieme di punti di Δ^+ .

2.4.1 Distanza di Hausdorff

Faremo uso di questa distanza per definire una metrica tra Funzioni di Taglia. Abbiamo già visto come è possibile associare ad ogni Funzione di Taglia una successione finita di punti contenuti in Δ^+ pertanto tali punti formano un sottoinsieme in Δ^+ . D'altra parte la diagonale estesa $\bar{\Delta}$ può essere interpretata come un deposito infinito di punti. Allora è ben definita una funzione σ :

$$\ell \xrightarrow{\sigma} P \cup \bar{\Delta}$$

dove $p \in P \Leftrightarrow p$ appartiene alla successione associata a ℓ . Ora come si può notare da semplici esempi la σ non è iniettiva, vale a dire che più Funzioni di Taglia possono essere associate allo stesso insieme di punti. Come conseguenza immediata di questa osservazione abbiamo che ogni metrica definita su un sottoinsieme di punti di Δ^+ indurrà una metrica sullo spazio delle Funzioni di Taglia che non potrà soddisfare la condizione di separabilità di una metrica ($d(x, y) = 0 \Rightarrow x = y$). Pertanto si tratterà di una pseudometrica. Useremo la distanza di Hausdorff per definire sullo spazio delle Funzioni di Taglia una pseudodistanza che chiameremo pseudodistanza di Hausdorff. Diamo di seguito le definizioni formali.

Definizione 2.8 (Distanza di Hausdorff). Sia (X, δ) uno spazio metrico e $A, B \subset X$ compatti. La *distanza di Hausdorff* d_H tra A e B è così definita:

$$d_H(A, B) = \max \{h(A, B), h(B, A)\}$$

dove

$$h(A, B) = \sup_{a \in A} \inf_{b \in B} \delta(a, b)$$

Nelle applicazioni si ha a che fare con sottoinsiemi finiti di punti quindi è ben posta la seguente definizione.

Definizione 2.9 (Pseudodistanza di Hausdorff). Siano

$$\{p_1, \dots, p_n\} = \sigma(\ell_1) - \bar{\Delta} \text{ e } \{q_1, \dots, q_m\} = \sigma(\ell_2) - \bar{\Delta},$$

definiamo pseudodistanza di Hausdorff:

$$\vec{d}_H(\ell_1, \ell_2) = \max_{i=1, \dots, n} \min_{j=1, \dots, m} \min \{\delta(p_i, q_j), \delta_\Delta(p_i)\},$$

dove $\delta_\Delta(p_i)$ è la distanza di p_i dalla diagonale Δ .

La non consistenza della pseudodistanza di Hausdorff tra Funzioni di Taglia è una conseguenza della non iniettività della funzione σ . D'altra parte il grande vantaggio di questa pseudometrica è che soddisfa la proprietà di resistenza al rumore con un costo computazionale basso, di $O(nm)$.

Capitolo 3

Insiemi Fuzzy

In questo capitolo presenteremo una breve introduzione della teoria degli insiemi fuzzy finalizzata a illustrare il concetto di fondo che ci ha portato a scegliere e poi implementare questa teoria in fase di assegnazione durante il nostro esperimento. La teoria degli insiemi fuzzy è, in effetti, uno strumento che cerca di conciliare la precisione della matematica con la pervasiva imprecisione del mondo reale. Il concetto primitivo fondamentale che sta alla base di tutta la matematica è quello di insieme, cioè di una collezione di oggetti; la teoria insiemistica classica fa uso del principio del terzo escluso, ovvero: dato un elemento, esso o appartiene o non appartiene a un certo insieme e non c'è una terza possibilità. L'idea della teoria fuzzy è proprio quella di costruire una teoria insiemistica che non si appoggia su questo principio ma sul fatto che vi sono più gradi (valori) possibili di appartenenza per un elemento ad un dato insieme che sfumano da un valore massimo (l'elemento appartiene all'insieme come nel senso classico) a un valore minimo (l'elemento non appartiene all'insieme). Si può per un attimo accantonare la logica booleana per adottarne un'altra, la logica fuzzy appunto più adatta a questa teoria [5]. Ora diamo qualche accenno e qualche definizione formale sugli insiemi fuzzy, alla fine del capitolo spiegheremo come abbiamo implementato questa teoria per il nostro esperimento mentre nel successivo evidenzieremo come l'adozione di questo strumento ha portato notevoli miglioramenti dei risultati.

3.1 Richiami classici

Sia \mathbf{E} un insieme e \mathbf{A} un suo sottoinsieme:

$$A \subset E.$$

Usiamo come consuetudine indicare che un elemento x di \mathbf{E} appartiene ad A con il simbolo \in ,

$$x \in A$$

Allo stesso modo possiamo indicare che un elemento x di \mathbf{E} appartiene ad A utilizzando la funzione caratteristica $\mu_A(x)$:

$$\mu_A(x) = 1 \quad \text{se } x \in A$$

$$\mu_A(x) = 0 \quad \text{se } x \notin A.$$

Esempio 2.

$$E = \{x_1, x_2, x_3, x_4, x_5\}$$

$$A = \{x_1, x_3, x_5\}.$$

Possiamo scrivere nella nuova notazione

$$A = \{(x_1, 1), (x_2, 0), (x_3, 1), (x_4, 0), (x_5, 1)\}.$$

E' facile vedere come si può riscrivere il concetto di intersezione, unione, complementazione in termini della funzione caratteristica.

$$\mu_{A \cap B}(x) = \mu_A(x) \mu_B(x)$$

$$\mu_{A \cup B}(x) = \max \{ \mu_A(x), \mu_B(x) \}$$

$$\mu_{\overline{A}} = 1 - \mu_A$$

Esempio 3. Sia

$$B = \{(x_1, 0), (x_2, 1), (x_3, 0), (x_4, 1), (x_5, 1)\}$$

$$\overline{A \cap B} = \{(x_1, 1), (x_2, 1), (x_3, 1), (x_4, 1), (x_5, 0)\}$$

3.2 Il concetto di sottoinsieme fuzzy

Abbandoniamo ora il principio del terzo escluso che assegna alla funzione caratteristica μ_A i soli valori dell'insieme $\{0, 1\}$ assumiamo invece che μ_A possa prendere valori sull'intervallo $[0, 1]$ cosicchè si possa dire che un elemento x di \mathbf{E} non appartiene ad A se $\mu_A(x) = 0$, appartiene 'poco' ad A se $\mu_A(x)$ è vicina a 0, appartiene 'più o meno' se $\mu_A(x)$ è vicino a 0.5, appartiene 'quasi del tutto' se $\mu_A(x)$ è vicina a 1 oppure appartiene ad A se $\mu_A(x) = 1$.

Definizione 3.1. Sia \mathbf{E} un insieme, numerabile o non numerabile, e sia x un suo elemento. Si definisce un *sottoinsieme fuzzy* \tilde{A} di \mathbf{E} l'insieme delle coppie ordinate

$$\{(x, \mu_A(x)), \forall x \in \mathbf{E}\}$$

dove $\mu_A(x)$ è detta *funzione di appartenenza* ed è definita su tutto \mathbf{E} a valori in un insieme totalmente ordinato \mathbf{M} detto *insieme di appartenenza*.

Osservazione 3. Se $\mathbf{M} = \{0, 1\}$, allora l'insieme fuzzy \tilde{A} si può vedere come un insieme ordinario A .

Esempio 4. Prendiamo un numero reale $a \in \mathbb{R}$ e definiamo il sottoinsieme fuzzy degli elementi di \mathbb{R} vicino ad a : $\tilde{A} = \{(x, \mu_A(x))\}$ dove la funzione di appartenenza assume valori nell'intervallo $[0, 1]$ con $\mu_A(a) = 1$ per $x = a$ e

$$\mu_A(x) = 1 - e^{-\frac{1}{(a-x)^2}}$$

per $x \neq a$.

Occorre definire le operazioni tra sottoinsiemi fuzzy nella nuova ottica delle funzioni di appartenenza.

Uguaglianza.

Diremo che

$$\tilde{A} = \tilde{B} \Leftrightarrow \mu_A(x) = \mu_B(x) \quad \forall x \in \mathbf{E}$$

altrimenti scriveremo $\tilde{A} \neq \tilde{B}$.

Inclusione.

Diremo che

$$\tilde{A} \subseteq \tilde{B} \Leftrightarrow \mu_A \leq \mu_B \quad \forall x \in \mathbf{E}$$

Complementazione.

Diremo che

$$\tilde{B} = \overline{\tilde{A}} \Leftrightarrow \mu_B(x) = 1 - \mu_A(x) \quad \forall x \in \mathbf{E}.$$

Intersezione e unione.

Siano \tilde{A}, \tilde{B} due sottoinsiemi fuzzy definiamo la funzione di appartenenza relativa all'intersezione come:

$$\mu_{A \cap B}(x) = \min\{\mu_A, \mu_B\}$$

Analogamente per l'unione

$$\mu_{A \cup B}(x) = \max\{\mu_A, \mu_B\}$$

Differenza.

Definiamo la differenza $\tilde{A} - \tilde{B} = \tilde{A} \cap \overline{\tilde{B}}$ dove la funzione di appartenenza si esprimerà

$$\mu_{\tilde{A} - \tilde{B}} = \min\{\mu_A(x), 1 - \mu_B(x)\} \quad \forall x \in \mathbf{E}.$$

E' divertente chiedersi a questo punto se valgono le consuete proprietà degli insiemi. In generale sarà sì [5] ma con alcune eccezioni. Chiaramente continuano a valere le proprietà commutative dell'unione e dell'intersezione, l'associatività, l'idempotenza, la distributività rispetto a intersezione e unione, l'involutività della complementazione e il teorema di De Morgan, tuttavia come è presumibile aspettarsi cadono le proprietà classiche $\tilde{A} \cap \overline{\tilde{A}} = \emptyset$ e $\tilde{A} \cup \overline{\tilde{A}} = \tilde{A}$ basate sul principio del terzo escluso.

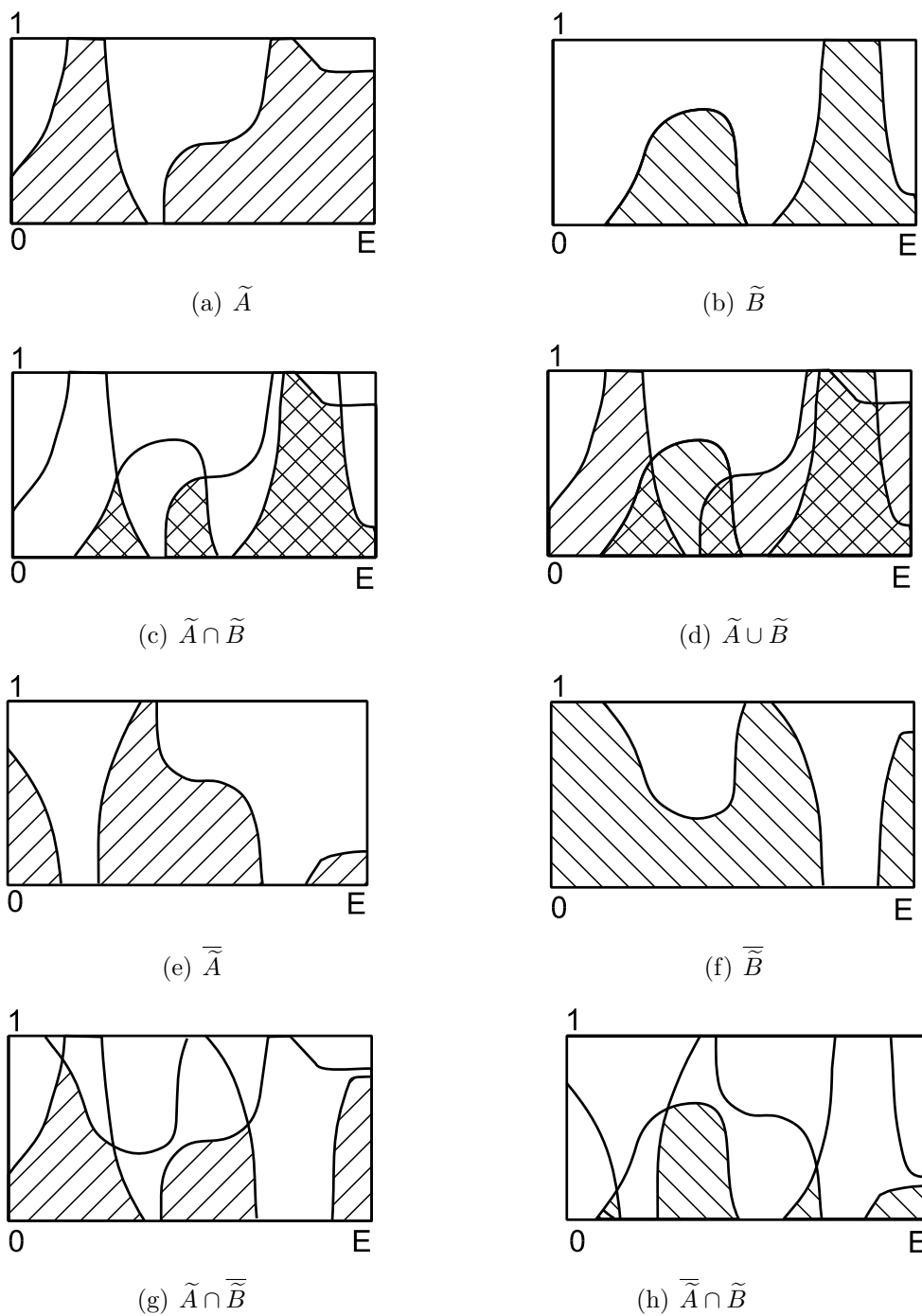


Figura 3.1: Rappresentazione grafica di alcune semplici operazioni tra sottoinsiemi fuzzy.

3.3 Assegnazione

Questa breve introduzione ci basta per comprendere come abbiamo agito durante l'esperimento nella fase di assegnazione. Chiamiamo il nostro database contenente tutte le immagini con \mathbf{E} e indichiamo con x un suo elemento. Definiamo ora i sottoinsiemi fuzzy $A_1 = \{x \mid x \text{ è opera di Angelo}\}$, $A_2 = \{x \mid x \text{ è opera di Costantini}\}$, $A_3 = \{x \mid x \text{ è opera di Sacco}\}$, $A_4 = \{x \mid x \text{ è un opera Scotellaro}\}$ e di conseguenza le relative funzioni di appartenenza che indicheremo per semplicità con μ_n $n = 1, 2, 3, 4$. Siano $\varphi^1, \dots, \varphi^m$ le funzioni misuranti considerate (quantonero, quantonero2, etc...- cfr cap.1). Per ogni φ^i si definisca ℓ^i la funzione di taglia associata, sia C_j^i la minima distanza di ℓ^i dalle rispettive funzioni di taglia del training set relativo al j -esimo autore. Poniamo $\mu_j^{\prime i} = 1/(C_j^i)^2$ e definiamo la funzione di appartenenza di x , rispetto φ^i , la normalizzazione

$$\mu_j^i = \frac{\mu_j^{\prime i}}{\sum_{k=1}^n \mu_k^{\prime i}},$$

in seguito si è fatta la media sulle funzione misuranti usate:

$$\bar{\mu}_j = \frac{1}{m} \sum_{i=1}^m \mu_j^i, \quad \forall j = 1, \dots, n,$$

si è poi traslato tutto nello 0:

$$\tilde{\mu}_j = \bar{\mu}_j - \min\{\bar{\mu}_k, k = 1, \dots, n\},$$

infine normalizzando un'ultima volta abbiamo ottenuto la funzione di appartenenza di x rispetto A_j :

$$\mu_j = \frac{\tilde{\mu}_j}{\sum_{k=1}^n \tilde{\mu}_k}.$$

In fase di assegnazione x verrà associata all'autore per il quale la funzione di appartenenza è più alta. Nel prossimo capitolo vedremo i risultati in relazione ai gruppi di funzioni misuranti scelte.

Capitolo 4

Esperimenti

4.1 Dati e strumenti utilizzati

I dati iniziali da cui siamo partiti sono gli stessi di [4] cioè un database di 109 immagini suddiviso in 4 classi così composte: 30 immagini di A.Mennillo, 23 di G.Costantini, 26 di J.Sacco, 30 della collezione Scotellaro (di G.Palumbo). Le immagini si presentano in formato standard 800 X 600 in bianco e nero (2 bit). Gli strumenti software usati sono stati Matlab e il linguaggio C con cui si sono implementati diversi programmi utili. Per il calcolo delle funzioni di taglia abbiamo usato il programma ST4Comics, appositamente modificato a partire da un pacchetto già in uso (*Size Tool For Linux*) a cui sono state aggiunte le funzioni *ad hoc* per il nostro esperimento. In particolare abbiamo aggiunto tutte le nuove funzioni misuranti, un applet per calcolare direttamente il baricentro, il filtro gaussiano e un applet per la lettura diretta degli istogrammi da un file testo. Per la fase di assegnazione abbiamo scritto altri programmi in linguaggio C implementando sia il metodo di [4] che quello fuzzy. Infine si è deciso di ripetere gli esperimenti di [4] per uniformare il settaggio dei parametri ed eliminare qualche piccola imperfezione così da rendere il confronto tra i due metodi più preciso possibile. In questa fase si è usato il Matlab e si ringrazia N.Cavazza per la disponibilità e collaborazione. Si è fatto uso del Matlab anche per calcolare gli istogrammi delle funzioni linee, flinee, flinee mediante trasformata di Hough [6].

4.2 Esperimenti preliminari

Il primo approccio alla fase sperimentale è stato piuttosto cauto, si è pensato prima a calcolare bene le funzioni di taglia delle immagini e costruire la metrica mediante la distanza di Hausdorff. L'output di questa fase sono state le classifiche per ogni immagine delle distanze dalle altre (vedi Appendice 4.3). Sulla base di queste classifiche si è passati alla fase di assegnazione. Inizialmente si è adottato lo stesso metodo di [4] poichè molto intuitivo e ci ha permesso di avere subito un termine di confronto. Il metodo è composto da 4 metodi differenti:

1. *Primovicino*: assegna all'immagine test l'autore dell'immagine più vicina in classifica.
2. *Punteggi*: assegna prima un punteggio ai primi 5 in classifica (10 al primo, 5 al secondo, 3 al terzo, 2 al quarto, 1 al quinto) e poi prende l'autore che ha il maggior punteggio.
3. *Baricentro*: prima calcola il baricentro di ogni classe, cioè la funzione di taglia della media degli istogrammi, infine assegna l'autore il cui baricentro è più vicino all'immagine test.
4. *Media*: prima calcola la media delle distanze di taglia su ogni classe, poi assegna l'autore la cui media è più vicina all'immagine test; a causa della diversa cardinalità di ogni classe si è fatta la media solo sui primi 15 classificati di ognuna di esse.

In principio le funzioni misuranti salti, quantonero, quantonero2 sono state calcolate solo sulla direzione verticale, ma si è osservato subito che, a causa dei valori interi degli istogrammi, i risultati non erano molto discriminanti, così si è optato per integrare l'informazione con anche le direzioni orizzontali, obliqua destra e obliqua sinistra. Si è usato come *merge method* la somma, vale a dire che si è presa come distanza la somma delle singole distanze ottenute per ogni direzione. La funzione che tiene conto di tutte e 4 le direzioni la chiameremo 'completa'. Vediamo ora i risultati e un primo confronto con quelli medi ottenuti ripercorrendo [4]. Nota: per risultati medi si intende la

media delle percentuali di successo rispetto alle 3 distanze usate in [4].

Per prima cosa occorre fare una considerazione sulla funzione, un po'

PRIMOVICINO

Funzione misurante	Proiezioni Verticali	Complete	N.Cavazza
quantonero	41.28	53.21	68.19
quantonero2	62.39	71.56	69.41
salti	51.37	69.72	62.69
linee	—	46.29	54.13
fquantonero	—	27.52	28.14
fquantonero2	45.87	54.13	57.83
fsalti	40.37	54.13	69.72
flinee	—	47.70	45.87
ffquantonero	43.11	46.78	59.33
ffquantonero2	52.39	48.62	59.63
ffsalti	66.97	64.22	67.31
fflinee	—	48.62	50.15

patologica, fquantonero. L'istogramma di questa funzione è infatti sostanzialmente piatto (in quanto i valori non superano quasi mai il valore 4) e quindi le funzioni di taglia risultano prevalentemente tutte uguali. Tuttavia si osserva una patologia analoga anche negli esperimenti di N.Cavazza, è quindi plausibile pensare che tale funzione non sia appropriata per questo studio. Un discorso a parte va fatto per la funzione linee e sue relative frequenze: ricordiamo che essa agisce in modo globale sulla figura e non dipende quindi dalla scelta di una direzione, tuttavia benchè nel complesso contengano un piccolo contributo di informazione (come vedremo in seguito), prese singolarmente non sono certo all'altezza delle altre. Facciamo un'ulteriore osservazione sulle funzioni di taglia, notiamo come mediamente le proiezioni verticali siano decisamente meno efficaci della funzione completa, eccezione fatta per i salti e ffquantonero2. Infine facendo un confronto con i risultati medi ottenuti come N.Cavazza osserviamo un'alternanza di successi a favore dell'uno o dell'altro a seconda delle funzioni, benchè in media le percentuali

siano più a favore di [4].

Vediamo ora quanto ottenuto col metodo Punteggi:

PUNTEGGI

Funzione misurante	Proiezioni Verticali	Complete	N.Cavazza
quantonero	42.20	55.97	69.72
quantonero2	63.30	72.48	71.25
salti	51.38	72.47	64.83
linee	—	46.79	57.79
fquantonero	—	27.52	28.13
fquantonero2	43.11	54.12	59.33
fsalti	47.70	55.05	69.11
flinee	—	45.87	47.08
ffquantonero	42.20	54.13	59.93
ffquantonero2	52.39	48.62	57.49
ffsalti	69.72	66.05	67.31
fflinee	—	45.87	49.53

Non osserviamo sostanziali differenze con il metodo Primovicino, solo nel caso dei salti si ha un percettibile miglioramento ma in media i risultati non variano. Constatiamo ancora una volta il fatto che la funzione completa dà risultati migliori della sola proiezione verticale pertanto faremo riferimento solo alla prima nelle prossime tabelle.

Vediamo ora i risultati per i metodi Baricentro e Media, da questi trarremo alcune conclusioni.

BARICENTRO e MEDIA

Funzione misurante	BARICENTRO		MEDIA	
	F. di Taglia	N.Cavazza	F. di taglia	N.Cavazza
quantonero	41.28	48.01	50.46	54.12
quantonero2	46.79	46.79	61.47	56.27
salto	44.04	71.86	60.56	60.24
linee	27.52	54.74	51.38	53.20
fquantonero	—	81.35	29.35	32.71
fquantonero2	34.00	66.06	53.21	55.35
fsalto	53.21	67.28	55.96	62.69
flinee	27.52	47.09	50.46	47.70
ffquantonero	38.54	63.00	59.63	53.21
ffquantonero2	45.87	60.86	46.79	61.16
ffsalto	52.29	64.53	61.46	62.07
fflinee	27.52	54.74	46.79	51.98

La prima riflessione da fare riguarda il metodo Baricentro, osserviamo che per le funzioni di taglia si ha un calo brusco dei successi a dispetto di un quasi miglioramento in [4] dove si leggono percentuali anomale per la funzione fquantonero. Questa situazione potrebbe dipendere dal fatto che quando si esegue la media degli istogrammi su una classe, inevitabilmente si ha un appiattimento dell'istogramma su un range più stretto di valori. Ciò potrebbe essere la spiegazione della divergenza dei due metodi: da una parte le metriche di [4] tengono conto della differenza di altezza sui singoli punti dell'istogramma e quindi istogrammi piatti, ma di altezza diversa, sono visti molto lontani; d'altra parte le funzioni di taglia sono indice più dell'andamento (o della forma) che dell'altezza assoluta di ogni valore, quindi evidenzieranno meglio la differenza tra un istogramma sinuoso e uno piatto piuttosto che tra due piatti. Ben diverso dalla media degli istogrammi è invece fare la media delle distanze tra le funzioni di taglia, non si presenta

più in questo caso quel fenomeno di appiattimento mal digerito dalle funzioni di taglia. Osserviamo infatti dalla tabella che le percentuali di successo relative a Media risalgono alla normalità mentre quelle di N.Cavazza calano lievemente rimanendo in linea con le precedenti.

Da questi metodi si intravede già che ci sono funzioni che tendenzialmente hanno percentuali di successo maggiori (salti vs linee) ovvero portano più informazione utile di altre. Sembrano esserci infatti caratteristiche dello stile che sono più discriminanti di altre nell'ottica di un riconoscimento automatico; se da una parte la geometricità sembra non essere un fattore molto discriminante, dall'altra il tratteggio (i salti da bianco a nero) sembra esserlo molto di più, come si può notare confrontando le funzioni relative ai salti con quelle relative alle linee.

Ciò non significa che lo siano in generale da un punto di vista artistico ma solo per un riconoscimento effettuato con un calcolatore.

4.3 Filtro Gaussiano

Come avviene per molti problemi di Computer Vision anche in questo caso occorre tenere conto del disturbo dovuto al rumore cioè alle piccole imperfezioni presenti inevitabilmente nelle immagini. Abbiamo perciò rifatto gli stessi esperimenti dopo aver applicato un filtro gaussiano alle immagini. I risultati mostrati nella tabella successiva ci indicano come il filtro non abbia globalmente portato miglioramenti sostanziali anzi, nel caso di fquantonero e fquantonero2 ha portato persino dei peggioramenti; tuttavia questo test ci ha rassicurato sulla robustezza al rumore delle funzioni di taglia. I risultati ottenuti da N.Cavazza con lo stesso filtro gaussiano confermano quanto da noi registrato e portano alle medesime conclusioni.

FILTRO GAUSSIANO

Funzione misurante	PrimoVicino	Punteggi	Media
quantonero	55.05	56.89	47.70
quantonero2	74.31	73.40	62.39
salti	63.30	61.46	66.97
fquantonero	—	—	50.46
fquantonero2	43.12	44.04	52.29
fsalti	59.63	61.47	54.13
ffquantonero	49.54	49.54	58.72
ffquantonero2	46.79	47.70	43.12
ffsalti	67.88	68.80	67.89

4.4 Assegnazione fuzzy

Giunti a questo punto, cercare l'accoppiamento ottimale tra metodi e funzioni per massimizzare i risultati non porterebbe a una conclusione solida ma sarebbe anzi dispersivo. Basta osservare per esempio che se da un lato la funzione `ffquantonero` ha successi migliori nel metodo `media` rispetto a `Primovicino`, la funzione `fquantonero2` si rivela all'opposto. E' difficile quindi dai risultati ottenuti estrarre un metodo unico ed efficiente per il riconoscimento automatico dell'autore. Occorre dunque trovare un metodo generale di assegnazione che riesca a tenere conto simultaneamente dell'informazione di ogni funzione misurante. Il metodo che abbiamo ritenuto più adatto a questa operazione è appunto basato sulle funzioni di appartenenza fuzzy (vedi Cap. 3.3).

Per cominciare abbiamo usato il metodo fuzzy su ogni singola funzione misurante completa. Questo procedimento potrebbe considerarsi come una variante più fine della sintesi tra la fase di *merge* delle quattro sotto-funzioni mediante somma con quella di assegnazione `Primovicino`. Anche qui abbiamo ripetuto l'esperimento dopo aver filtrato l'immagine:

SINGOLE FUNZIONI

Funzione Misurante	Senza Filtro	Con Filtro	N.Cavazza
quantonero	54.13	55.05	78.90
quantonero2	73.39	73.39	73.39
salti	65.14	65.14	67.89
linee	46.29	47.70	54.13
fquantonero	35.78	34.86	29.35
fquantonero2	51.38	52.29	61.47
fsalti	50.46	50.46	74.31
flinee	47.70	46.29	45.87
ffquantonero	51.38	63.30	61.47
ffquantonero2	57.80	52.29	59.33
ffsalti	65.14	63.30	73.39
fflinee	48.62	47.70	50.15

Facendo una prima analisi di questi dati appare che il metodo fuzzy ha sempre esito migliore rispetto alle proiezioni verticali, eccezion fatta per la funzione ffsalti; d'altra parte se paragonata al merge method somma più Primovicino si osserva una lieve superiorità pari all' 1,6 % (in media) con alcuni piccoli sbalzi; in particolare ffquantonero e ffquantonero2 sono a vantaggio di fuzzy mentre fquantonero2 e fsalti a vantaggio di merge method somma. Anche col metodo fuzzy i dati relativi a fquantonero non sono attendibili e ciò sancisce una volta di più l'inadeguatezza di questa funzione. Analizzando i dati relativi al filtro gaussiano si osserva lo stesso andamento che si ha avuto nei casi precedenti, cioè nessuna globale differenza tra filtro e senza filtro. Da segnalare solo un visibile incremento per ffsalti con filtro e un calo per ffquantonero2.

Commentiamo ora i valori dell'ultima colonna. Non si può che constatare una netta superiorità sugli altri, questo porterebbe alla conclusione immediata che il metodo di [4] sia decisamente più performante del nostro, tuttavia proseguendo nel raffinamento della fase di assegnazione il divario si è un po'

colmato come vedremo nella prossimo paragrafo.

4.5 Gruppi di funzioni misuranti

A questo punto abbiamo usato il metodo fuzzy sfruttandone al pieno le sue proprietà di mediatore dei risultati. Abbiamo allargato il numero di funzioni misuranti coinvolte da fuzzy considerando gruppi di funzioni complete così formati:

Base: quantonero, quantonero2, salti.

Frequenze: fquantonero, fquantonero2, fsalti.

Fascie: ffquantonero, ffquantonero2, ffsalti.

Quantonero: quantonero, fquantonero, ffquantonero.

Quantonero2: quantonero2, fquantonero2, ffquantonero2.

Salti: salti, fsalti, ffsalti.

TUTTE: tutte le funzioni.

Abbiamo escluso per un attimo le linee a causa delle avvisaglie di inaffidabilità che hanno dato nelle fasi precedenti. Da notare che abbiamo pensato anche all'esclusione, per lo stesso motivo, di fquantonero ma per ragioni di simmetria l'abbiamo poi riammesso. Abbiamo inoltre applicato il metodo fuzzy a [4] per sintetizzare i risultati delle tre diverse distanze (Assoluta, Euclidea, Geometrica) in modo da fare un confronto globale con i nostri. Ecco i risultati:

Come anticipato prima notiamo un livellamento dei due metodi ma quel che in realtà più conta è il notevole aumento generalizzato delle percentuali di successo, salite in media del 10-15 %. Un ottimo risultato! Il passo verso

FUZZY GLOBALE

Funzione Misurante	F.Taglia	N.Cavazza
Base	77.98	82.57
Frequenze	66.06	75.23
Fascie	77.06	76.15
Salti	73.39	77.06
Quantonero	69.72	78.90
Quantonero2	76.15	70.64
TUTTE	84.40	82.57

la costruzione di un procedimento generale automatico per il riconoscimento dell'autore sembra essere compiuto; la percentuale del 84.40% relativa a tutte le funzioni è soddisfacente. Tuttavia abbiamo ancora qualche cartuccia da sparare per migliorare questa percentuale e per giudicare in ultima istanza quale dei due metodi ([4] o funzioni di taglia) sia il più performante. Una piccola osservazione, vediamo come le percentuali di frequenze e quantonero siano le uniche sotto il 70%, ricordiamo che al loro interno c'è la funzione patologica fquantonero che potrebbe esserne la causa, tuttavia abbiamo effettuato un test rimuovendo questa funzione ma il risultato è ulteriormente peggiorato.

4.6 Convoluzione

Facciamo una parentesi riportando un tentativo fatto per cercare di raffinare la fase di costruzione della metrica. Abbiamo applicato una maschera di convoluzione a tutti gli istogrammi prima di calcolarne la funzione di taglia, i parametri scelti per la gaussiana di convoluzione tendono a non appiattire oltremodo l'istogramma ma solo a smussare i picchi di discontinuità. Sugli istogrammi così convoluti sono state inoltre ricalcolate le distanze alla maniera di [4] per avere un raffronto.

CONVOLUZIONE

Gruppi di funzioni	Funzioni di taglia	N.Cavazza
Base	85.32	82.57
Frequenze	71.56	69.72
Fascie	66.97	77.06
Quantonero	60.55	69.72
Quantonero2	72.47	69.72
Salti	73.39	75.23
TUTTE	83.48	82.56

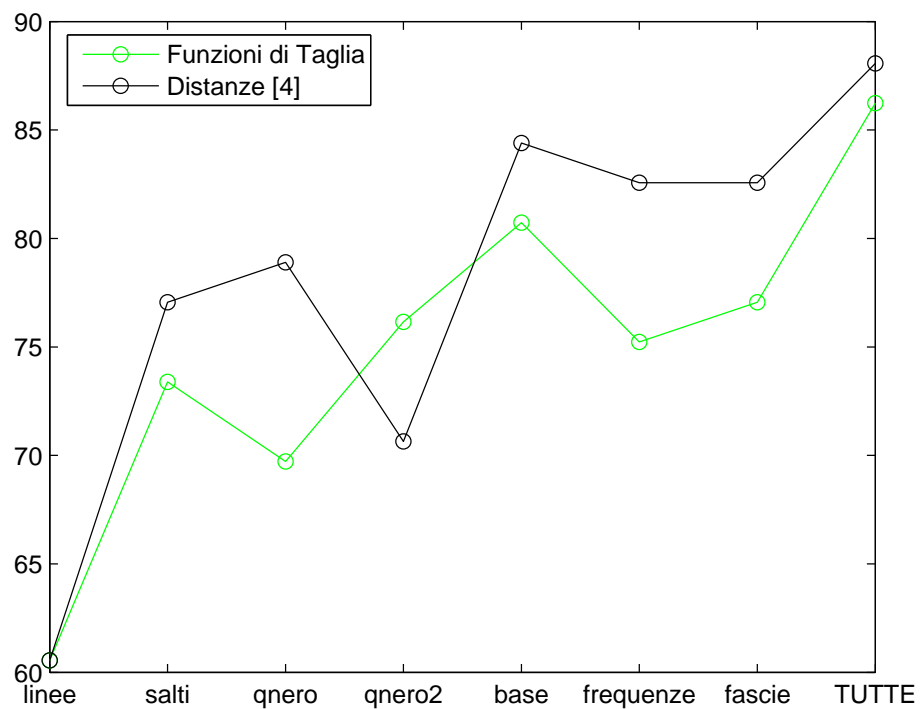
I risultati non sono molto significativi, allo stesso del filtro gaussiano è difficile esprimere un giudizio favorevole tanto è vero che le percentuali sono pesantemente calate nei casi delle frequenze a fasce e quantonero. La convoluzione non si è rivelato uno strumento adatto al nostro studio.

4.7 Risultati Finali

Dopo aver scoperto che la rimozione della funzione quantonero, benchè ritenuta patologica, ha peggiorato il risultato invece che sollevarlo, abbiamo pensato di recuperare le funzioni linee precedentemente accantonate, sperando che potessero fornire qualche informazione in più. E così è stato! L'inclusione nei primi tre gruppi Base, Frequenze e Fascie rispettivamente con le funzioni *linee*, *flinee*, *fflinee* costituenti a loro volta il gruppo Linee ha portato i seguenti nuovi risultati.

RISULTATI FINALI

Gruppi di funzioni	FUNZIONI DI TAGLIA	ISTOGRAMMI
Base	80.73	84.40
Frequenze	75.23	82.57
Fascie	76.15	82.57
Quantonero	69.72	78.90
Quantonero2	76.15	70.64
Salti	73.39	77.06
Linee	60.55	66.05
TUTTE	86.24	88.07



Notiamo come effettivamente il recupero delle funzioni linee sia stato positivo, ancora una volta le percentuali sono globalmente aumentate. E' aumentata inoltre la percentuale massima di successo all' 88.07% che è ot-

tima in un contesto applicativo. Inoltre come si può leggere nel grafico, le funzioni linee sono state l'ago della bilancia che assegna la miglior performance al metodo di [4] per pochi punti percentuali sul metodo delle funzioni di taglia. Solo in un caso su 8, quello di quantonero2, le funzioni di taglia hanno un successo superiore a [4]. Si osserva come il divario fra le due sia molto più accentuato per i gruppi Quantonero, Fascie, Frequenze che per gli altri dove non si notano differenze sostanziali.

Come ultima mossa, spinti dal successo dell'assegnazione fuzzy abbiamo fuso insieme i due metodi considerando l'unione delle funzioni misuranti dell'uno e dell'altro.

UNIONE TAGLIA E ISTOGRAMMI

Gruppi di funzioni	TAGLIA + ISTOGRAMMI
Base	88.07
Frequenze	84.40
Fascie	87.16
Quantonero	79.81
Quantonero2	80.73
Salti	76.15
Linee	73.39
TUTTE	91.74

Abbiamo preso per ogni funzione misurante le 4 sotto-funzioni (una per ogni direzione di proiezione) relative alle funzioni di taglia e le abbiamo unite alle 3 sotto-funzioni di [4] (una per ogni distanza usata) per un totale di 7 sotto-funzioni per ogni funzione. Le linee contano invece solo 4 sotto-funzioni poichè non c'è distinzione tra direzioni quando si applicano le funzioni di taglia. Abbiamo quindi ottenuto un totale di 75 funzioni su cui abbiamo applicato l'assegnazione fuzzy, distinguendo i soliti gruppi di funzioni. Osserviamo come ancora una volta le percentuali siano aumentate in maniera particolare per le linee. La percentuale massima raggiunta con tutte le funzioni ha superato il 90% e altri gruppi di funzioni hanno raggiunto l'80% di successo.

4.8 Conclusioni

Giunti al termine di questo lavoro possiamo trarre le nostre conclusioni distinguendo i diversi strumenti che abbiamo testato e messo a confronto nelle varie fasi. Relativamente alla prima fase, quella di estrapolazione delle informazioni dalle immagini, possiamo dire che le funzioni di taglia e il metodo delle distanze proposte da [4] sono entrambi validi strumenti, anche se in ultima analisi il metodo di [4] si è rivelato più performante, con una differenza di percentuali media del 3.9% e minima (1,5%) nel caso più significativo di tutte le funzioni. D'altra parte, in fase di assegnazione, concludiamo che il metodo fuzzy è chiaramente più performante dei metodi proposti in [4] sia rispetto al caso delle funzioni di taglia sia alle stesse distanze di [4]. L'aumento delle percentuali ottenuto nell'ultimo esperimento in cui si è considerata l'unione delle funzioni misuranti dei due metodi, ha ulteriormente provato che l'assegnazione fuzzy è un ottimo metodo di assegnazione ed è indipendente dalla scelta degli strumenti nella fase iniziale. In conclusione, il nostro problema di Computer Vision sul riconoscimento automatico dello stile fumettistico a partire dai nostri dati iniziali è stato risolto con buone percentuali di successo. La percentuale massima raggiunta del 91% è valida in un qualsiasi contesto applicativo futuro ma il costo computazionale speso per raggiungerla è ancora troppo elevato.

Appendice

Alcuni esempi

Esempio di un'immagine della collezione Scotellaro:



Figura 4.1: Scotellaro30.jpg

Vediamone il suo grafico relativo alla funzione salti:

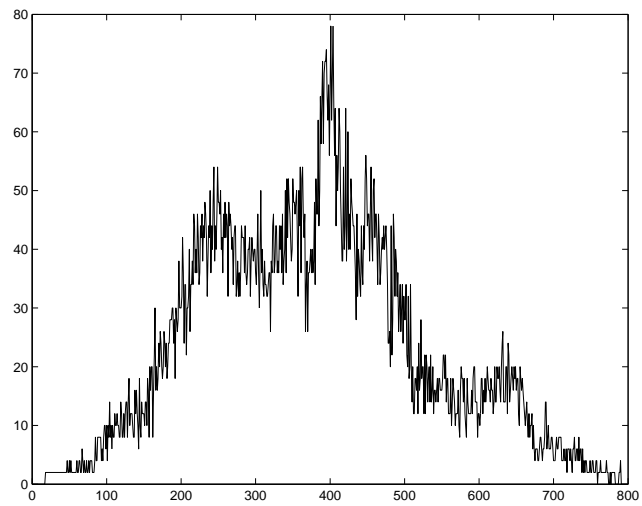


Figura 4.2: Istogramma Salti

Vediamo ora la classifica ottenuta dopo aver calcolato le funzioni di taglia e usato il merge method somma:

Vediamo i risultati dell'assegnazione dell'autore mediante:

```
num funmis 4  num training files 113 distance = haus_0_0
fmis 5  image scotellaro30.jpg.ang  dist = 0.000000
fmis 5  image scotellaro10.jpg.ang  dist = 63.063187
fmis 5  image scotellaro24.jpg.ang  dist = 65.499741
fmis 5  image ang(19).jpg.ang  dist = 67.511749
fmis 5  image ang(21).jpg.ang  dist = 68.083191
fmis 5  image scotellaro19.jpg.ang  dist = 68.491837
fmis 5  image scotellaro2.jpg.ang  dist = 69.735916
fmis 5  image cos16.jpg.ang  dist = 73.012405
fmis 5  image ang(22).jpg.ang  dist = 74.276299
fmis 5  image scotellaro7.jpg.ang  dist = 75.542557
fmis 5  image scotellaro21.jpg.ang  dist = 75.715111
fmis 5  image scotellaro27.jpg.ang  dist = 79.209419
fmis 5  image ang(3).jpg.ang  dist = 81.534172
fmis 5  image ang(1).jpg.ang  dist = 82.569153
fmis 5  image scotellaro6.jpg.ang  dist = 82.879387
fmis 5  image scotellaro22.jpg.ang  dist = 83.362579
fmis 5  image scotellaro3.jpg.ang  dist = 84.063538
fmis 5  image ang(2).jpg.ang  dist = 86.322548
fmis 5  image scotellaro23.jpg.ang  dist = 86.730003
fmis 5  image scotellaro12.jpg.ang  dist = 87.141533
fmis 5  image ang(14).jpg.ang  dist = 88.165527
fmis 5  image scotellaro13.jpg.ang  dist = 89.088005
fmis 5  image cos7.jpg.ang  dist = 92.961784
fmis 5  image cos23.jpg.ang  dist = 95.088005
fmis 5  image cos14.jpg.ang  dist = 95.174461
fmis 5  image cos12.jpg.ang  dist = 95.518684
fmis 5  image scotellaro15.jpg.ang  dist = 96.307663
fmis 5  image cos3.jpg.ang  dist = 96.848869
fmis 5  image scotellaro1.jpg.ang  dist = 98.022308
fmis 5  image ang(26).jpg.ang  dist = 98.649109
fmis 5  image scotellaro20.jpg.ang  dist = 100.000000
fmis 5  image ang(30).jpg.ang  dist = 102.560219
fmis 5  image scotellaro8.jpg.ang  dist = 104.745277
fmis 5  image scotellaro18.jpg.ang  dist = 106.058395
fmis 5  image scotellaro14.jpg.ang  dist = 107.620499
fmis 5  image ang(25).jpg.ang  dist = 108.142136
fmis 5  image scotellaro17.jpg.ang  dist = 108.806252
fmis 5  image ang(24).jpg.ang  dist = 110.873009
fmis 5  image scotellaro26.jpg.ang  dist = 111.861694
```

Figura 4.3: Classifica di Scotellaro30 con Merge Method Somma

PRIMOVICINO a Scotellaro,
PUNTEGGI a Scotellaro.

Vediamo ora un output del calcolo della funzione di assegnazione fuzzy:

	Vert	Orizz	Dx	Sx
Distanza da n				
FalT	0.17316	0.37113	0.30631	0.50955
FFiE	0.11111			
FFlT	0.04834	0.30266	0.06169	0.27513
FFnD	0.28571	0.25000	0.25000	0.28571
FFnP	0.42947	0.23223	0.29649	0.18310
FliE	0.14286			
FneD	0.25000	0.25000	0.25000	0.25000
FneP	0.20000	0.29412	0.02669	0.12470
LinE	0.22807			
QneD	0.09292	0.30149	0.18317	0.23090
QneP	0.14611	0.45210	0.19334	0.19100
SaIT	0.44370	0.33362	0.34759	0.20213

Distanza da o				
FalT	0.34428	0.09278	0.44108	0.18344
FFiE	0.22222			
FFlT	0.07734	0.30266	0.75954	0.27513
FFnD	0.28571	0.25000	0.25000	0.14286
FFnP	0.17919	0.23223	0.14917	0.19804
FliE	0.14286			
FneD	0.25000	0.25000	0.25000	0.25000
FneP	0.20000	0.29412	0.19274	0.09592
LinE	0.22807			
QneD	0.11480	0.15252	0.11550	0.20795
QneP	0.12086	0.14058	0.14204	0.16454
SaIT	0.30172	0.23928	0.34759	0.50183

```

Distanza da a
FalT      0.02531 0.16495 0.02757 0.02038
FFiE      0.22222
FFlT      0.05259 0.09202 0.04731 0.17460
FFnD      0.14286 0.25000 0.25000 0.28571
FFnP      0.16039 0.23223 0.41475 0.30943
FliE      0.35714
FneD      0.25000 0.25000 0.25000 0.25000
FneP      0.40000 0.11765 0.43365 0.62350
LinE      0.31579
QneD      0.09323 0.31210 0.15476 0.16701
QneP      0.05625 0.22424 0.22961 0.12967
SalT      0.01886 0.12006 0.02674 0.01617

Distanza da c
FalT      0.45725 0.37113 0.22504 0.28662
FFiE      0.44444
FFlT      0.82173 0.30266 0.13146 0.27513
FFnD      0.28571 0.25000 0.25000 0.28571
FFnP      0.23096 0.30332 0.13959 0.30943
FliE      0.35714
FneD      0.25000 0.25000 0.25000 0.25000
FneP      0.20000 0.29412 0.34692 0.15588
LinE      0.22807
QneD      0.69905 0.23389 0.54657 0.39414
QneP      0.67677 0.18308 0.43501 0.51479
SalT      0.23572 0.30704 0.27807 0.27987

CALCOLA MEDIE
dist da n: 0.24375 dist da o: 0.23304 dist da a:
0.19663 dist da c: 0.32657
Traslazione
dist da n: 0.22073 dist da o: 0.17055 dist da a:
0.00000 dist da c: 0.60872
scotellaro30.jpg.res viene attribuita a c      OK

```

Figura 4.4: Esempio di procedimento fuzzy sul gruppo Tutte le funzioni.

Riportiamo ora l'esito di un test con assegnazione fuzzy per l'immagine cos22.jpg (appartenente a Costantini) protagonista di numerosi insuccessi:

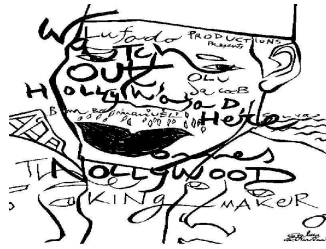


Figura 4.5: cos22.jpg

Base	Costantini OK
Frequenze	Angelo NO
Fascie	Scotellaro NO
Quantonero	Sacco NO
Quantonero2	Costantini OK
Salti	Costantini OK
Linee	Scotellaro NO
TUTTE	Scotellaro NO

Bibliografia

- [1] M.d'Amico, P.Frosini, C.Landi: *Using matching distance in Size Theory: a survey.*
- [2] A.Cerri, M.Ferri, D.Giorgi: *A complete Keypics experiment with Size Function* 12 Settembre 2005.
In: Proc. CIVR 2005, Singapore, Springer LNCS 3568, W.-K. Leow, M.S. Lew, T.-S. Chua, et al. Eds. (2005), 357-366.
- [3] A.Verri, C. Uras, P.Frosini, M.Ferri: *On the use of size function for shape analysis.*
Biol. Cybern. 70 (1993), 99-107.
- [4] N.Cavazza: *Titolo* Tesi di Laurea Specialistica 12 ottobre 2007.
- [5] A.Kaufmann *Theory of Fuzzy subsets Vol 1.*
Academic Press Inc 1975.
- [6] D. Ballard and C. Brown *Computer Vision*, Prentice-Hall,
1982, Chap. 4.
- [7] R. Boyle and R. Thomas *Computer Vision:A First Course*, Blackwell
Scientific Publications, 1988, Chap. 5.