

Recupero di marchi di fabbrica mediante
scheletrizzazione e funzioni di taglia ridotte

F.Neri

23 novembre 2005

Indice

1	Computer Vision	5
1.1	image processing	6
1.1.1	Scheletrizzazione	6
1.1.2	Rilevamento dei Contorni	9
2	Funzioni di Taglia: Definizioni e Proprietà	11
2.1	Definizioni delle Funzioni di Taglia Ridotte	12
2.2	Prime proprietà	12
2.3	Rappresentazione algebrica delle Funzioni di Taglia	13
2.4	Distanza tra Funzioni di Taglia	15
2.4.1	Distanza di Hausdorff	16
3	Descrizione dell'esperimento	21
3.1	Scelta delle funzioni misuranti	22
3.2	Dati Statistici	22
3.3	Risultati	23
3.4	Scheletrizzazione vs. Rilevamento dei Bordi	42
3.5	Rumore vs Rimozione del Rumore	43
3.6	conclusioni	44
	Bibliografia	47

Capitolo 1

Computer Vision

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. Il problema che studieremo fa parte di quel campo della Computer Vision che va sotto il nome di Shape Matching, o riconoscimento di forme reali studiando una o più immagini digitali. Un problema tipico di Shape Matching è recuperare immagini contenute in un database di grosse dimensioni, per fare ciò è necessario sviluppare strumenti che analizzano il contenuto delle immagini stesse. Possiamo suddividere l'approccio generale ad un problema di shape-matching in due passi fondamentali:

1. scegliere un descrittore di forma;
2. scegliere una misura di dissimilarità per i descrittori.

Entrambe le due scelte sono di importanza centrale per avere risultati invarianti per un certo gruppo di trasformazioni. Per l'esperimento che vedremo nel terzo capitolo, useremo e studieremo un particolare descrittore che combina informazioni sia geometriche che topologiche e permette di scegliere le

invarianze per i gruppi di trasformazioni richiesti. Tale strumento sono le Funzioni di Taglia che vedremo descritte nel prossimo capitolo.

1.1 Image processing

Computer Vision e (digital)-image processing sono strettamente legate, la distinzione tra le due non è così netta. La computer vision adotta molti metodi che tradizionalmente appartengono all' image processing. Una distinzione formale è che l' image processing si occupa di trasformazioni di immagini, cioè sia gli input che gli output di questi processi sono sempre immagini. Nessuno di questi processi tuttavia si spinge oltre a occuparsi di interpretare le caratteristiche o i contenuti di un'immagine. Esempi di image-processing sono la rimozione del rumore, miglioramento di immagini sfuocate o sbiadite tramite correzioni di colore e luminosità, compressione di immagini, rilevamento di contorni e scheletri. I metodi che interessano a noi fanno parte di quel campo dell' image processing che vanno sotto il nome di operazioni matematiche morfologiche [5], o più semplicemente *operazioni morfologiche*, in particolare parleremo di algoritmi per: la rimozione del rumore, rilevamento dei contorni e scheletrizzazione. L'idea principale delle operazioni morfologiche è analizzare il contenuto di un'immagine confrontando essa con piccoli campioni in tutti i punti dell'immagine. Modificando questo campione, detto *struttura di controllo*, si possono ottenere effetti diversi. In questo lavoro avremo a che fare con immagini bidimensionali binarie in bianco e nero, pertanto quanto detto di seguito vale solo per questo tipo di immagine. Le operazioni morfologiche di cui studieremo più avanti gli effetti sui recuperi dei marchi sono due: la scheletrizzazione e il rilevamento dei contorni.

1.1.1 Scheletrizzazione

Gli scheltri si prestano molto al tipo di esperimento che eseguiremo poiché forniscono un modo per semplificare le forme in un'immagine con curve sottili anche un solo pixel, mantenendo però alcune caratteristiche topologiche e metriche come la connessione e l'estensione.

Definizione 1.1.

Sia \mathbf{R} un sottoinsieme piano, \mathbf{B} il suo bordo, P un punto di \mathbf{R} . Un *vicino prossimo* di P in \mathbf{B} è il punto M in \mathbf{B} tale che non esiste nessun altro punto su \mathbf{B} tale che la sua distanza da P sia minore di PM . Se P ha più di un *vicino prossimo*, allora P si dice *punto scheletrico* di \mathbf{R} ; l'insieme dei *punti scheletrici* è detto lo *scheletro* di \mathbf{R} .

Questa definizione implica che i *punti scheletrici* siano i centri delle circonferenze contenute in \mathbf{R} che sono tangenti al bordo di \mathbf{R} in almeno due punti distinti.

Osservazione 1. Se P è un punto del bordo \mathbf{B} in cui la curvatura di \mathbf{B} è un massimo locale, allora esiste un ramo dello scheletro che termina in P .

Descriviamo ora le operazioni morfologiche necessarie per definire la scheletrizzazione e che ci serviranno anche in seguito per il rilevamento dei contorni.

Un po' di notazioni.
Indichiamo con

$$A = \{a : val(a) = 1\} \quad \bar{A} = \{b : b \notin A\} = \{b : val(b) = 0\}$$

Indichiamo la traslazione di A mediante h :

$$A_h = \{a + h : a \in A\}$$

Nota che trattando con immagini binarie si tratta di traslazione mediante vettori a coordinate intere (in \mathbb{Z}^2). B rappresenta una struttura di controllo, usualmente è una matrice $3 * 3$ con valori logici 1 e 0.

Definizione 1.2. Erosione-Dilatazione

Definiamo la Dilatazione tramite B come:

$$D(A, B) = \bigcup_{\beta \in B} (A_\beta)$$

Definiamo la Erosione tramite B come:

$$E(A, B) = \bigcap_{\beta \in B} (A_{-\beta})$$

Dove A rappresenta la nostra immagine e B è una struttura di controllo. Denotiamo con $nB = D^n(B, B)$ cioè la ripetizione su B di n volte la dilatazione tramite B .

Una proprietà importante è la loro complementarità, cioè:

$$\bar{E}(A, B) = D(\bar{A}, -B)$$

$$\bar{D}(A, B) = E(\bar{A}, -B)$$

1	1	1
1	1	1
1	1	1

-	1	-
1	1	1
-	1	-

Figura 1.1: A sinistra la struttura tipo N_8 a destra quella tipo N_4

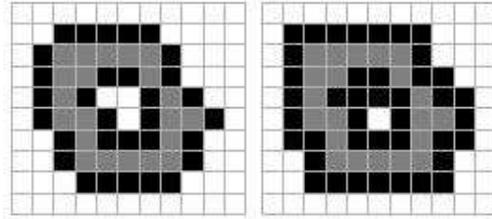


Figura 1.2: A sinistra una dilatazione ottenuta con N_4 , a destra una ottenuta con N_8

La dilatazione in generale aumenta l'estensione mentre l'erosione la diminuisce. La dilatazione ed erosione è ben definita quando si definisce una struttura di controllo B che definisce le traslazioni da compiere. Mostriamo le due strutture di controllo standard utilizzate in tal senso.

E' facile vedere come applicando prima un'erosione e poi una dilatazione si riesca a eliminare il rumore provocato dai punti isolati.

Combinando dilatazione ed erosione si ottengono:

Definizione 1.3. Opening-Closing

Opening:

$$A \circ B = D(E(A, B), B)$$

Closing:

$$A \bullet B = E(D(A, -B), -B)$$

Anche in questo caso valgono le proprietà:

$$\overline{A \bullet B} = \bar{A} \circ B$$

$$\overline{A \circ B} = \bar{A} \bullet B$$

Introduciamo ora due operatori Booleani:

Definizione 1.4.

Siano M, N due immagini binarie.

$M \cap N =$ Prodotto logico elemento per elemento tra M e N .

$M \cup N =$ Somma logica elemento per elemento tra M e N

Ora definiamo l'operazione Hit-Miss.

Definizione 1.5.

Siano A un'immagine binaria, B_1 e B_2 due strutture di controllo tali che $B_1 \cap B_2 = \emptyset$

$$HitMiss(A, B_1, B_2) = E(A, B_1) \cap \bar{E}(\bar{A}, B_1)$$

La definizione sopra data nel caso continuo non è altrettanto semplice da implementare nel caso discreto di immagini binarie, esistono pertanto svariati tipi di scheletri che possono o no soddisfare le proprietà sopra descritte. Un procedimento base di scheletrizzazione $S(A)$ di un'immagine A è il seguente: Si definisce prima

$$S_k(A) = E(A, kB) \cap [E(A, kB) \circ B]$$

Dove K è il più grande valore di k per cui $S_{k+1}(A) = \emptyset$, la struttura di controllo B può essere per esempio la N_4 o la N_8 . Lo scheletro è:

$$S(A) = \bigcup_{k=0}^K S_k(A)$$

Una felice proprietà di questo scheletro è che è possibile risalire all'immagine originaria quando si conoscono B e tutti gli S_k , non mantiene però le proprietà topologiche. Un'alternativa a questo tipo di scheletro che mantenga le proprietà topologiche si ottiene implementando un algoritmo di *thinning*:

$$Thinning(A) = Thin(A, B_1, B_2) = A \cap HitMiss(A, B_1, B_2)$$

che dipende strettamente dalla scelta di B_1 e B_2 . Per questo fatto è possibile ottenere diversi tipi di scheletri tramite thinning. Gli scheletri che useremo noi sono stati creati usando il software MatLab e la funzione di scheletrizzazione in esso implementata; gli scheletri ottenuti mantengono la topologia, l'estensione ma non sono 'sottili' un solo pixel. Osserviamo che la scheletrizzazione è un procedimento molto sensibile al rumore, la presenza di anche un solo pixel fuori posto può generare macroscopiche imperfezioni, per questo motivo potrebbe essere necessario eseguire a monte delle procedure di rimozione del rumore. Vediamo alcuni esempi di instabilità.

1.1.2 Rilevamento dei Contorni

A questo punto, potendo utilizzare tutte le operazioni morfologiche sopra definite è facile descrivere un metodo per rilevare i contorni di un'immagine.

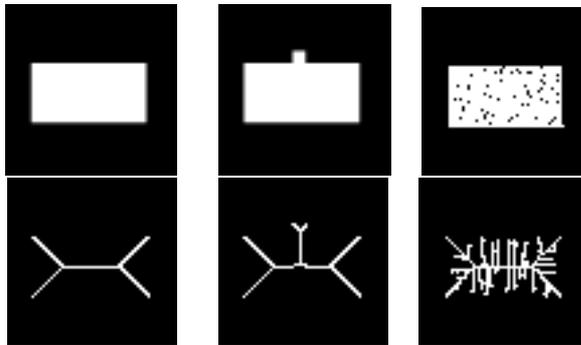


Figura 1.3: sopra le immagini primitive e sotto i rispettivi scheletri

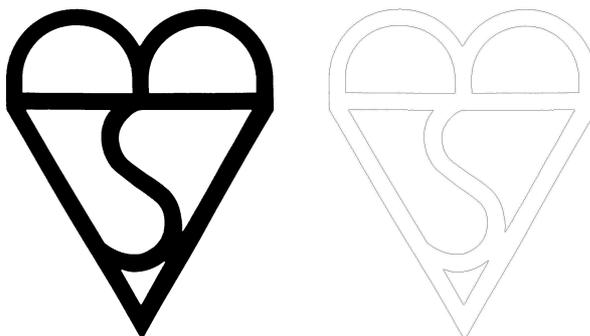
1	1	1	-	-	-	-	1	-
1	1	1	-	1	-	1	1	1
1	1	1	-	-	-	-	1	-

Figura 1.4: A sinistra $B=N_8$, al centro B_1 e a destra B_2

Anche in questo caso esistono in letteratura diversi modi per ricavare i contorni di un'immagine binaria. Prendiamo le strutture di controllo B, B_1, B_2 , simmetriche, dove - sta per 'non faccio confronti su questi pixel'. Un'altro modo per ricavare un contorno che mantenga la connessione-4, rispettivamente a 8 è:

$$\partial A = A \cap E(A, N_8)$$

$$\partial A = A \cap E(A, N_4)$$



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 2. 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 3. 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 $\ell_{(\mathcal{M}, \varphi)}(x, y)$.

2.2 Prime proprietà

Elenchiamo di seguito alcune osservazione 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. $\ell_{(\mathcal{M}, \varphi)}(x, y) = 0 \forall x < \min_{P \in \mathcal{M}} \varphi(P)$.

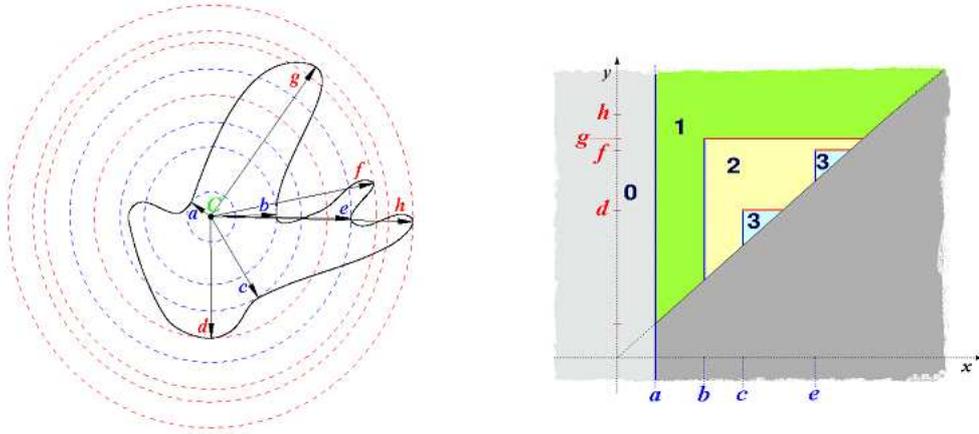


Figura 2.1: Fig 1.1

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 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 tanto più 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 $r = k$ con $k \in \mathbb{R}$, definiamo il numero $\mu(r)$ 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:

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

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'altraparte 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 cardinanià 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 verificata quando data la distanza tra due

Funzioni di Taglia è possibile modificare le relative funzioni misuranti in maniera che le Funzioni di Taglia coincidano.

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

La distanza di Hausdorff è largamente usata in Computer Vision anche al di fuori delle Funzioni di Taglia, per un approfondimento su questo si veda [4]; Noi la useremo 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 2.7 (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 2.8 (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:

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

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 $O(nm)$.

Se \mathcal{M} è un'immagine digitale

Per calcolare con il calcolatore le Funzioni di Taglia occorre prima discretizzare la varietà \mathcal{M} , il modo più semplice e usato per rappresentare la coppia di taglia fa uso dei grafi. Nel nostro caso abbiamo a che fare con delle immagini digitali bidimensionali binarie, risulta semplice perciò associare queste a un grafo; i vertici sono rappresentati dai singoli pixel in primo piano, due vertici sono legati da uno spigolo se i pixel sono connessi rispetto la connessione-4 (cioè se uno è contenuto nell' N_4 centrato sull'altro pixel). A ogni vertice infine si associa il valore della funzione misurante φ .

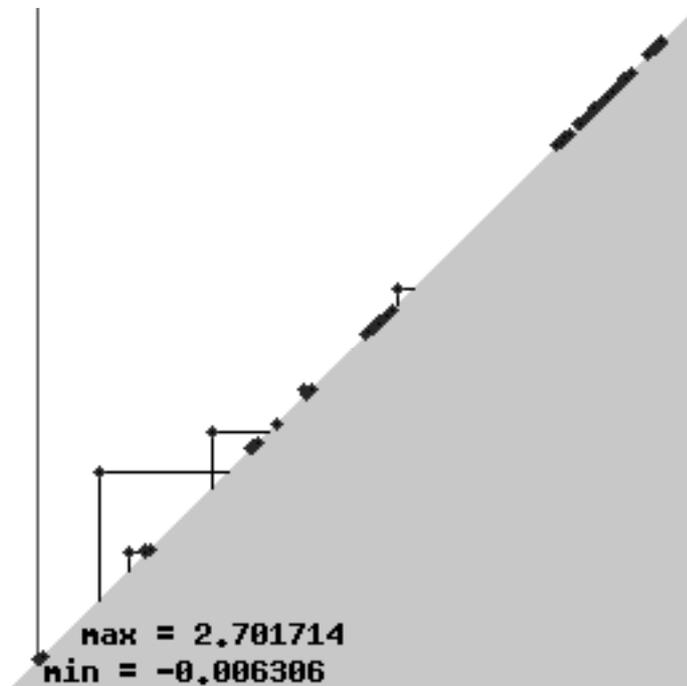
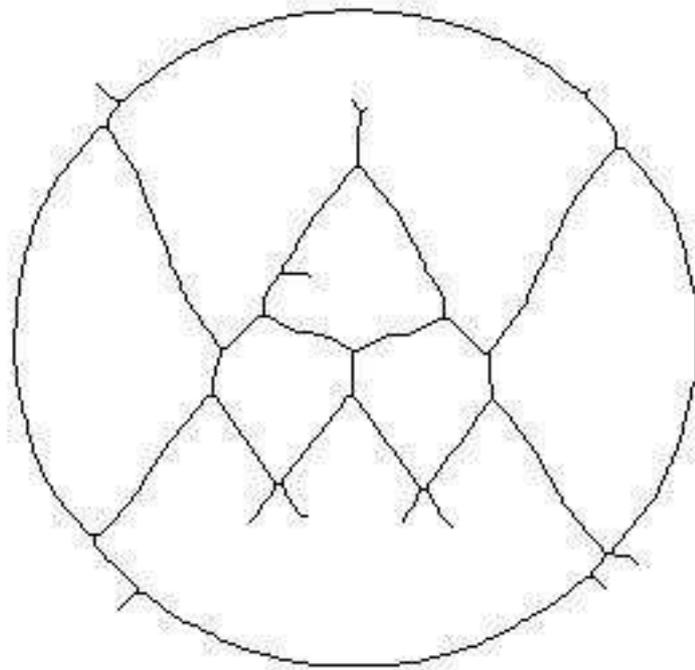


Figura 2.2: A sinistra l'immagine in assenza di rumore e a destra la relativa Funzione di Taglia

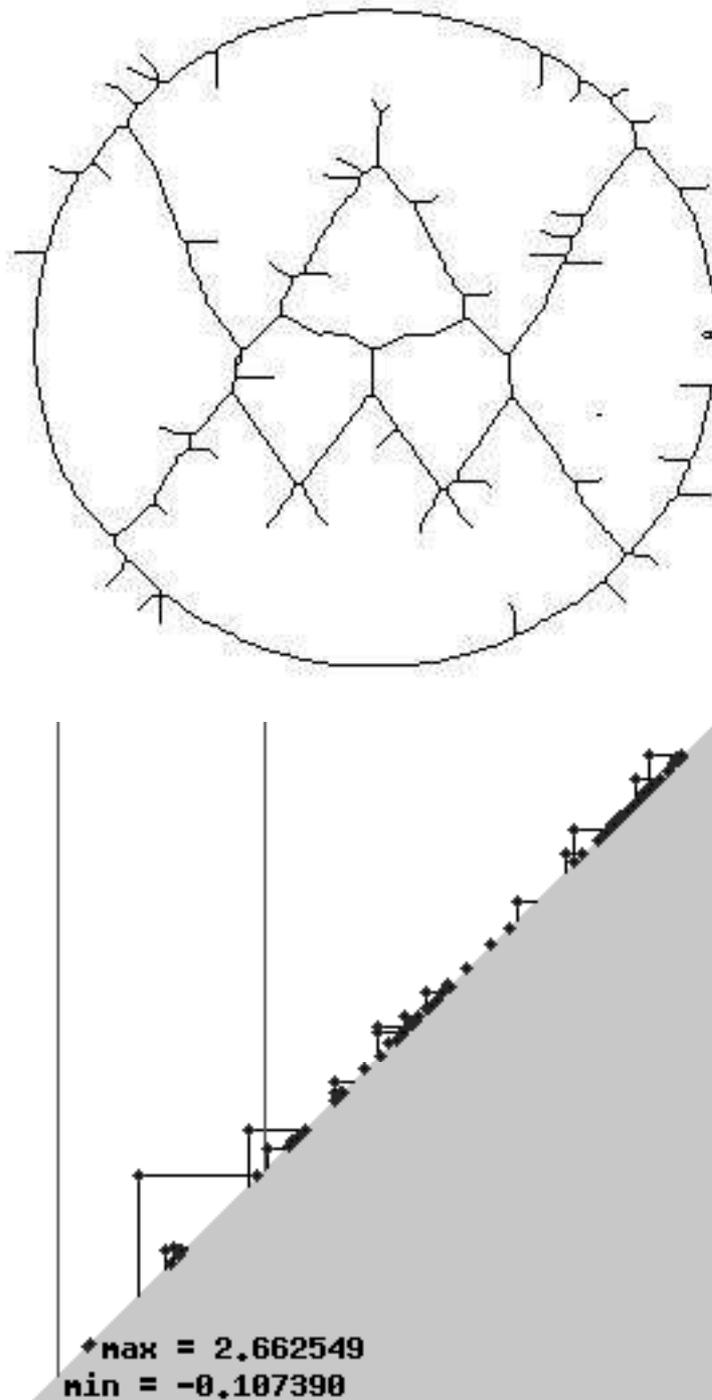


Figura 2.3: A sinistra l'immagine con presenza di rumore e a destra la relativa Funzione di Taglia

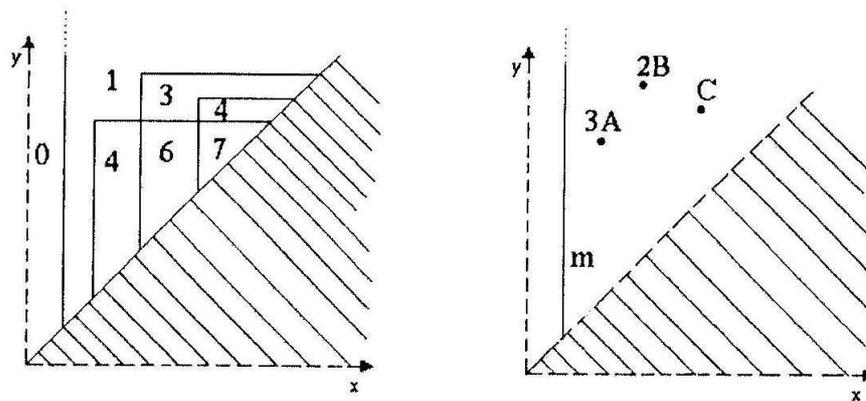


Figura 2.4: Punti e rette angolari di una funzione di taglia ridotta: A,B,C sono punti angolari, rispettivamente di molteplicità 3, 2, 1. m è l'unica retta angolare.

Capitolo 3

Descrizione dell'esperimento

L'esperimento consiste nel testare le Funzioni di Taglia su un dataset contenente 308 marchi di fabbrica e simulare prove di recupero. Le immagini sono state ricavate da una digitalizzazione tramite scanner e si presentano in formato standard '.jpg' in bianco e nero. I marchi sono generalmente diversi tra loro e alcuni appartengono a industrie famose; in tal senso la curiosità del lettore può essere soddisfatta in ?? dove è riportata una porzione del dataset. Il dataset contiene quattro classi di marchi che possono considerarsi commercialmente equivalenti, vale a dire che hanno presentano molto somiglianti. Gli elementi appartenenti ad una stessa classe verranno di seguito chiamati *rilevanti*. Le immagini evidenziano alcune imperfezioni, rumore ai bordi, pixel mancanti o pixel isolati. Partendo da siffatto dataset ne sono stati creati altri quattro: i primi due contenenti gli scheletri e i contorni ottenuti applicando le relative operazioni morfologiche descritte nel primo capitolo, gli altri due ottenuti allo stesso modo ma successivamente a operazioni di image processing per la rimozione del rumore. Si noti che l'immagine processing è stato fatto prima della scheletrizzazione e non sugli scheletri (analogamente per i bordi). In seguito abbiamo calcolato le Funzioni di Taglia e la (pseudo)distanza di Hausdorff, su tutti e cinque i dataset: immagini normali, scheletri, bordi, scheletri dopo image processing e contorni dopo image processing. Infine calcoleremo la precisione del recupero dei rilevanti e cercheremo di capire come influiscono le operazioni morfologiche su di essa. Valuteremo inoltre quanto sia influente l'operazione di image processing di pulizia dell'immagine. Ricordiamo che le operazioni di pulizia dell'immagine, almeno in questo caso, sono molto costose in termini di tempo, poichè non possono essere automatizzate e richiedono un controllo da parte di un operatore.

3.1 Scelta delle funzioni misuranti

Il primo passo da fare per utilizzare le Funzioni di Taglia è quello di scegliere una funzione misurante adeguata al problema. Ci chiediamo in primo luogo se abbiamo bisogno di conservare un' invarianza per qualche tipo di trasformazione. Per loro natura i marchi di fabbrica sono abbastanza rigidi e hanno quasi sempre un verso di lettura, perciò non è necessario avere un' invarianza per rotazioni; infatti se un marchio ruotato di un certo angolo è tale da far pensare che sia girato 'male', allora molto difficilmente può essere confuso con l'originale. Un' invarianza di cui abbiamo senz'altro bisogno è per le traslazioni e per scala. Un'altra invarianza che potrebbe tornarci utile è quella per inversione di colori, ma può essere ottenuta molto semplicemente studiando il contorno piuttosto che l'intera immagine. Ciò di cui abbiamo bisogno è una funzione che ci descriva da più punti di vista la disposizione dei pixel che formano l'immagine. La funzione misurante che adotteremo consiste in 16 distanze da determinati punti fissi. Fissiamo un riferimento cartesiano (O, \vec{x}, \vec{y}) nel piano affinché possiamo indicare con $p = (x_p, y_p)$ un suo generico punto. Definiamo una funzione misurante $\varphi_p(x, y) = d(p, (x, y))$ con d distanza euclidea; Per ottenere le invarianze volute, ogni immagine binaria viene per prima cosa normalizzata (ma senza abbassare la risoluzione) e traslata affinché il centro di massa cada sull'origine O del sistema di riferimento. Definiamo ora formalmente l'insieme delle 16 funzioni misuranti.

$$\Phi = \left\{ \varphi_p \mid p = \frac{\bar{r}}{2} \left(\cos \left(\bar{\alpha} + i \frac{\pi}{2} \right), \sin \left(\bar{\alpha} + i \frac{\pi}{2} \right) \right) i = 1, \dots, 4 \right\} \cup \\ \left\{ \varphi_p \mid p = \bar{r} \left(\cos \left(\bar{\alpha} + i \frac{\pi}{4} \right), \sin \left(\bar{\alpha} + i \frac{\pi}{4} \right) \right) i = 1, \dots, 7 \right\} \cup \\ \left\{ \varphi_p \mid p = \frac{3\bar{r}}{2} \left(\cos \left(\bar{\alpha} + i \frac{\pi}{2} \right), \sin \left(\bar{\alpha} + i \frac{\pi}{2} \right) \right) i = 1, \dots, 4 \right\},$$

dove le costanti \bar{r} e $\bar{\alpha}$ assumono rispettivamente i valori 0.8 (tutte le immagini sono scalate rispetto al raggio medio) e 0.349 (che corrispondono approssimativamente a una rotazione di 20°). Avendo sedici funzioni misuranti si ottengono sedici Funzioni di Taglia differenti, ognuna relativa ad una φ_p . La distanza tra le immagini è calcolata come la media delle distanze tra ogni Funzione di Taglia.

3.2 Dati Statistici

Come strumento di valutazione dei risultati abbiamo usato un metodo che ci permetta di tenere sufficientemente in considerazione la posizione dei rilevanti recuperati. Per fare ciò calcoleremo *precisione*, *recall* e *Rank-medio normalizzato* di cui diamo le definizioni formali.

Definizione 3.1. Definiamo $P(k)$ e $R(k)$, rispettivamente la *precisione* e il *recall* sulle prime k immagini recuperate:

$$P(k) = \frac{NR(k)}{k} \quad R(k) = \frac{NR(k)}{N_{rel}}$$

dove $NR(k)$ è il numero dei rilevanti recuperati nelle prime k posizioni, N_{rel} è il numero dei rilevanti.

Osservazione 4. Per $k = N_{rel}$ si ha $P(N_{rel}) = R(N_{rel})$.

Osservazione 5. $0 < P(k), R(k) \leq 1$

Definizione 3.2. Definiamo \widetilde{Rank} medio normalizzato il seguente numero:

$$\widetilde{Rank} = \frac{1}{NN_{rel}} \left(\sum_{i=1}^{N_{rel}} R_i - \frac{N_{rel}(N_{rel} + 1)}{2} \right)$$

dove R_i è la posizione dell' i -esimo rilevante recuperato, N è il numero di elementi del dataset.

Osservazione 6. Il \widetilde{Rank} è 0 per un recupero ottimale, 1 per un recupero totalmente fallito.

3.3 Risultati

Riportiamo qui un esempio di recupero del marchio 1138103 da tutti e cinque i dataset, riportando i primi 10 recuperati e le posizioni dei rimanenti rilevanti. Nelle figure successive visualizzeremo i marchi in forma standard, non mostreremo infatti gli scheltri o i contorni poichè preferiamo evidenziare non tanto le forme quanto gli effetti di queste trasformazioni.



1138103.jpg

Immagini normali:

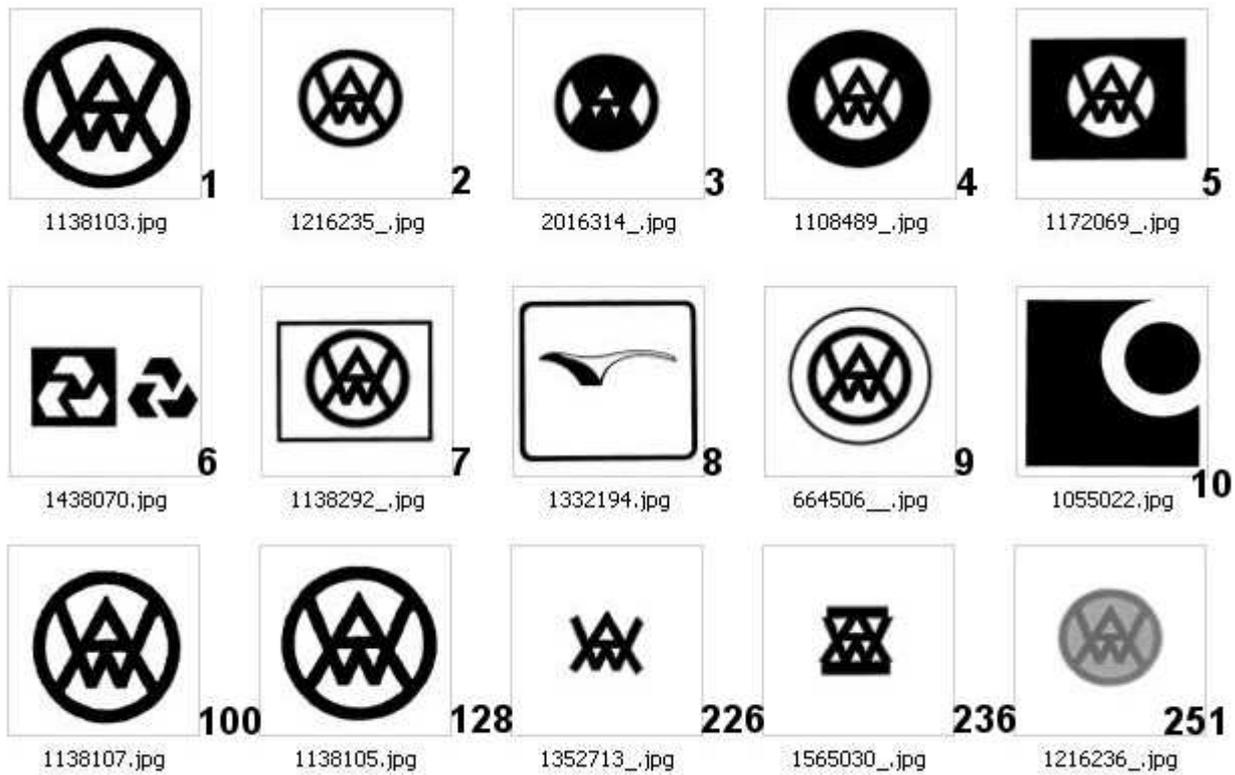


Figura 3.1:

Mediante scheletri:

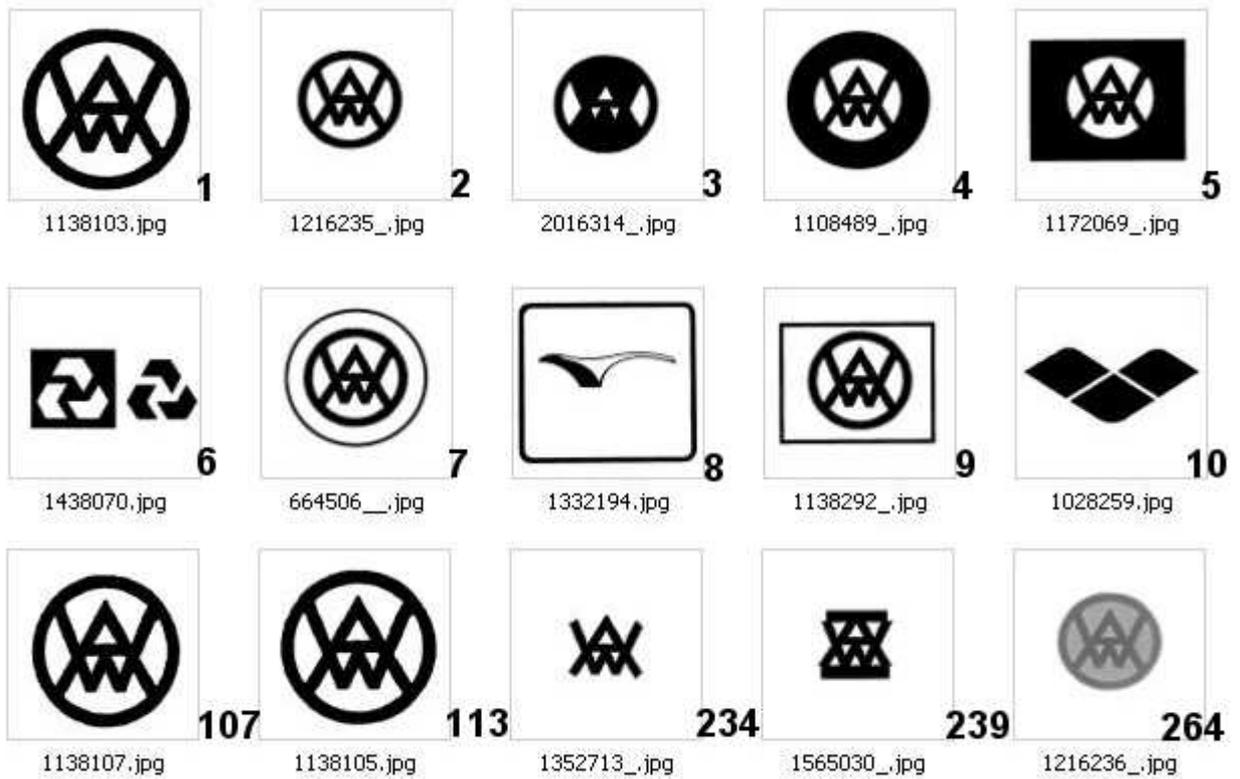


Figura 3.2:

Mediante contorni: In questo caso non è stato rilevato nessun rilevante (diverso da se stesso) nelle prime 10 posizioni, il più vicino si trova in 26° posizione.

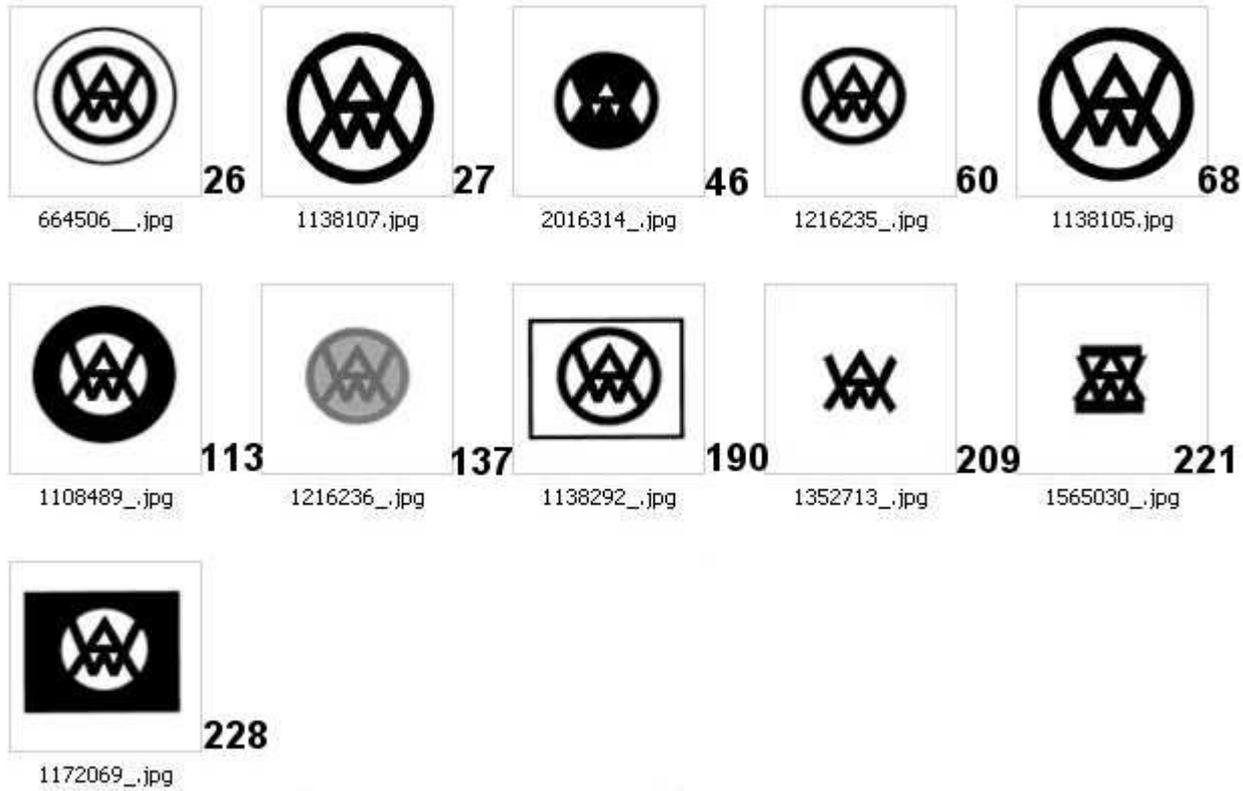


Figura 3.3:

Mediante scheletri dopo image-processing:

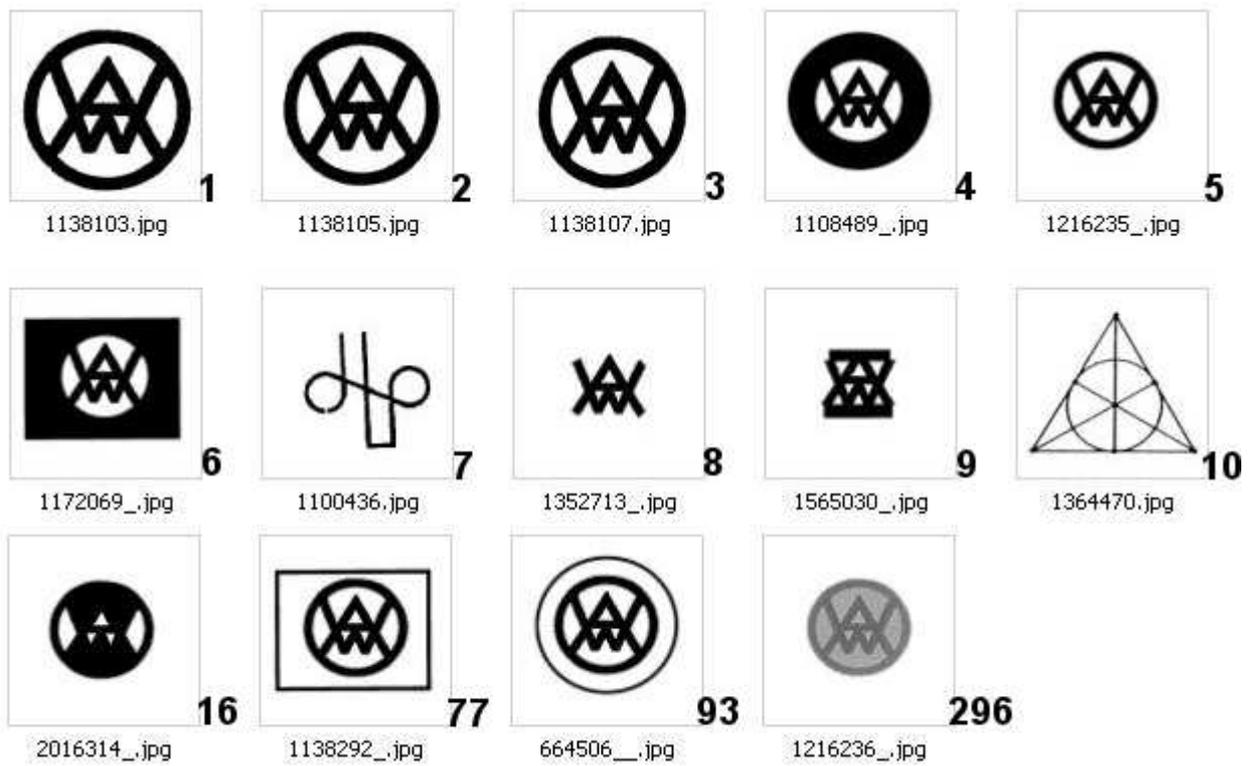


Figura 3.4:

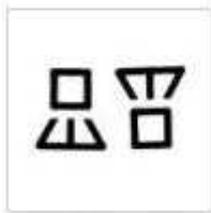
Mediante contorni dopo image-processing:



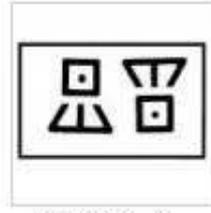
Figura 3.5:

L'esito del recupero in fig 3.1 è soddisfacente, la maggior parte dei rilevanti (7 su 12) è infatti recuperata nelle prime 10 posizioni. Ai fini pratici ciò che a noi interessa maggiormente è il numero di rilevanti recuperati nelle prime dieci o quindici posizioni. L'esito del recupero in fig 3.3 mediante gli scheletri è sostanzialmente lo stesso del precedente; osserviamo inoltre che anche due dei tre marchi estranei recuperati compaiono in entrambi i casi. Il recupero in fig 3.3 mediante contorni è invece alquanto fallimentare. Tuttavia il recupero in fig 3.5 mediante contorni dopo image-processing ha esiti molto buoni con precisione persino superiore a quella delle immagini normali. Questo può far pensare che l'algoritmo di rilevamento dei bordi è molto sensibile al rumore che poi viene riflesso sul calcolo delle funzioni di taglia. Osservando il recupero in fig 3.3 notiamo che i risultati sono ancora molto buoni e che la differenza tra il prima e dopo image-processing non è così marcata come si nota per i contorni.

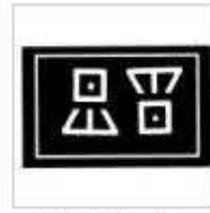
Ora studiamo il comportamento dell'immagine 2014722. Prestiamo attenzione alla relazione con il marchio 2014724 e 2040904, osservando che il secondo è sostanzialmente il negativo del primo.



2014722_.jpg



2014724_.jpg



2040904_.jpg

Riportiamo per questo caso i primi 10 marchi recuperati, e le posizioni dei rimanenti rilevanti recuperati.

Immagini normali:

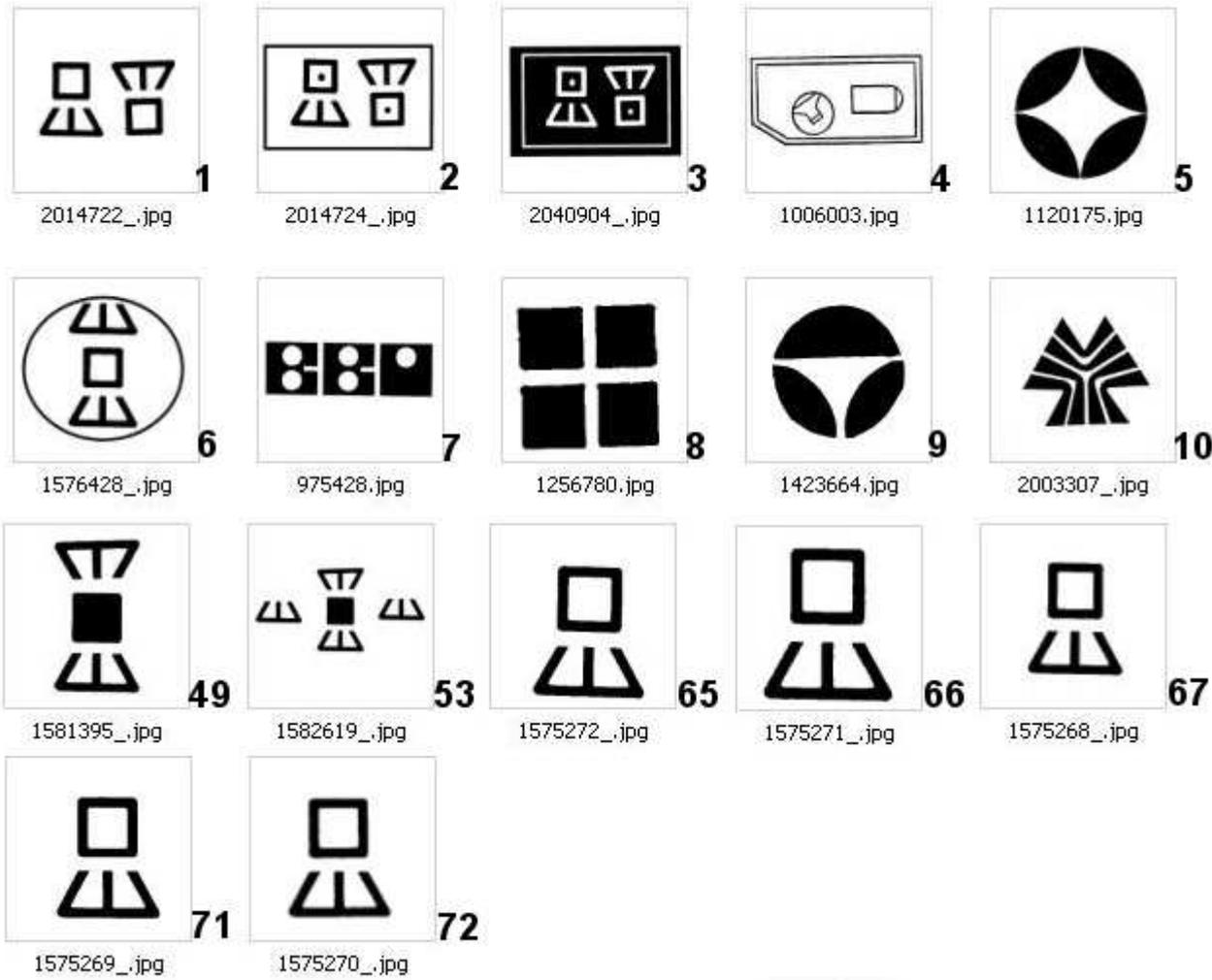


Figura 3.6:

Riportiamo ora solo le posizioni in cui sono stati rilevati i rilevanti.
Mediante contorni:

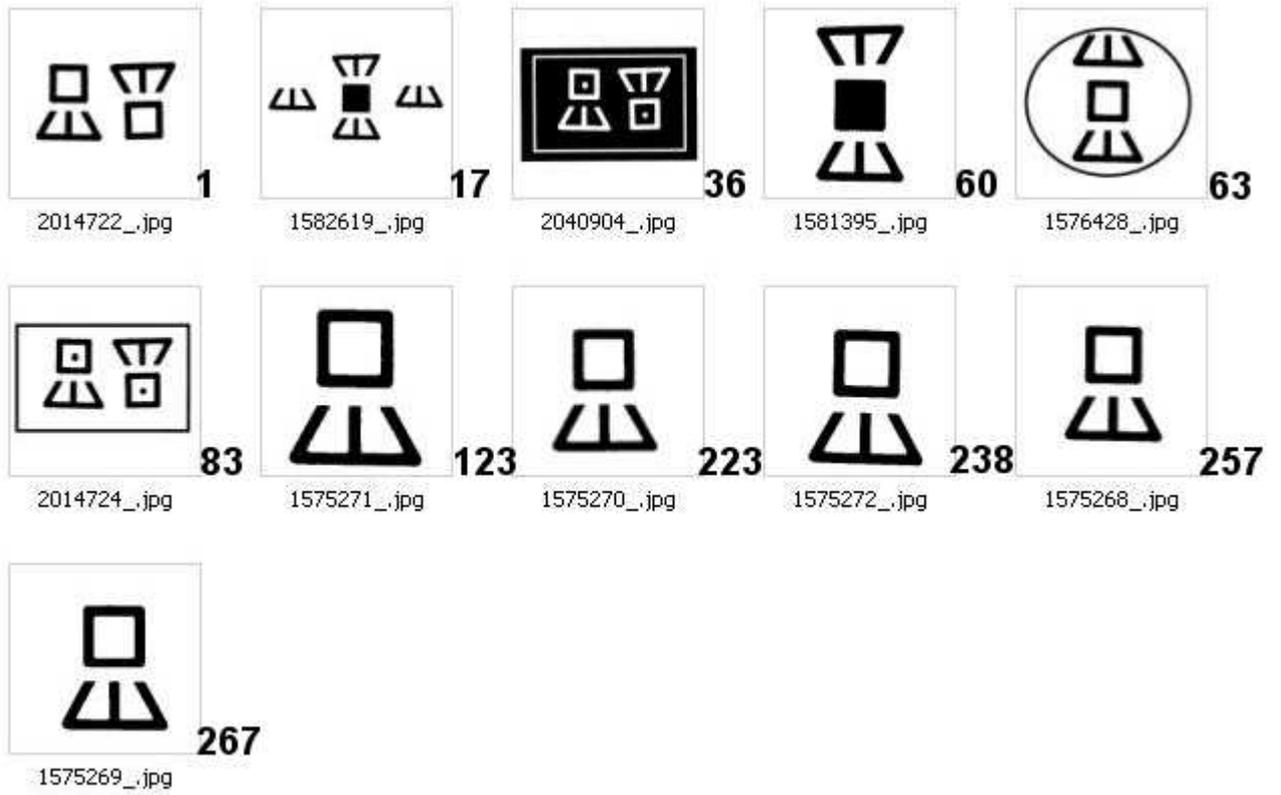


Figura 3.7:

Mediante contorni dopo image-processing:

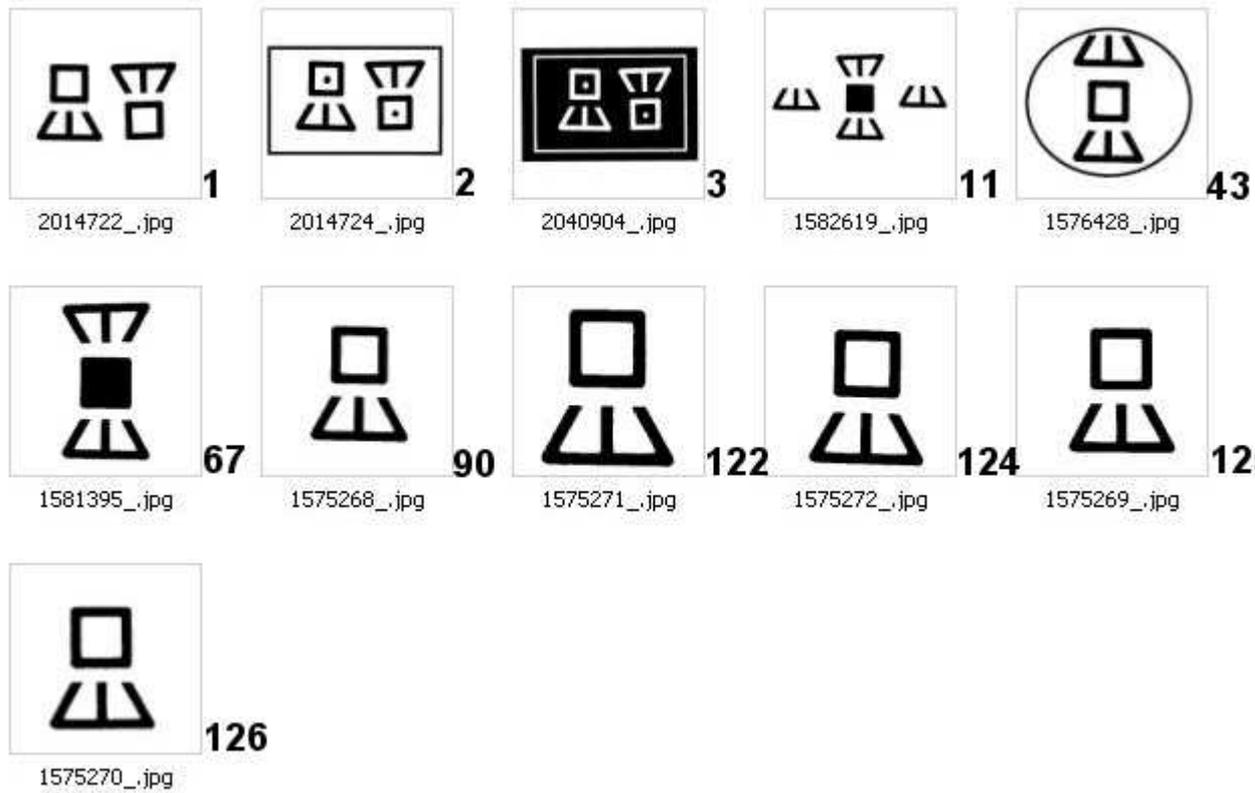


Figura 3.8:

Mediante scheletri dopo image-processing:

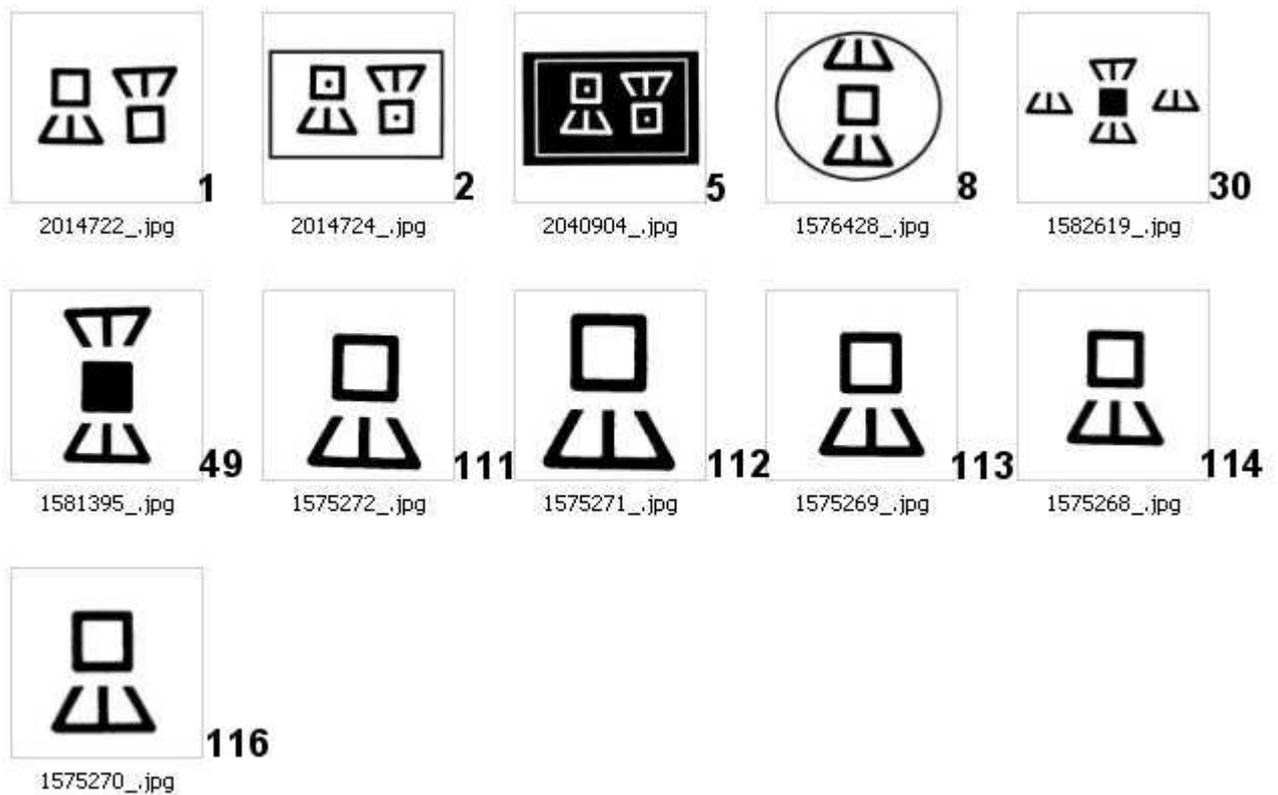


Figura 3.9:

Ricordiamo che il programma che calcola le funzioni di taglia legge come sfondo i pixel bianchi (valore 0) e come primo piano i pixel neri (valore 1), pertanto è concesso aspettarsi che l'immagine negativa venga vista molto diversa dall'immagine primitiva. In tal senso anche lo scheletro di un'immagine negativa è profondamente diverso dallo scheletro dell'immagine positiva. Tuttavia i risultati del recupero di 2014724 in fig. 3.3 si discostano dalle aspettative. Si tratta di un caso oppure c'è da aspettarsi che esista qualche relazione tra un'immagine e la propria inversa per colore? Visto il ripetersi del recupero della 2040904 nelle prime posizioni (5° posizione) anche per gli scheletri dopo image-processing (vedi fig. 3.3) è improbabile che si tratti di un caso, quanto meno dovrà trattarsi di una particolarità intrinseca a quell'immagine. D'altra parte il rilevamento dei contorni, a priori, può rilevarsi un metodo per ottenere un'invarianza per inversione di colore, tuttavia i contorni delle due immagini (positiva e negativa) differiscono sempre almeno

per la cornice. A causa di questo difetto aggiunto alla presenza di rumore il recupero dei contorni in fig. 3.3 ha portato risultati peggiori rispetto alle immagini normali; riducendo il rumore (vedi fig. 3.3) i risultati ritornano ad essere soddisfacenti.

Studiamo ora il marchio 2010916.
Per quanto riguarda le immagini normali osserviamo che i rilevanti ven-



Figura 3.10:

gono recuperati tutti (esclusi il 1516142, 8869761, 2043261) nelle prime 20 posizioni. La 1516142 è una tessitura (vedremo più avanti i problemi che riguardano questo genere di immagini) ed è plausibile che venga recuperata così lontana; la 8869761 come vedremo in tabella 3.3 è causa di alcuni dei peggiori recuperi effettuati. La 2043261 invece è sostanzialmente uguale alla 2010916, pertanto la causa della sua distanza dalla 2010916 dipenderà completamente dalla presenza di rumore, infatti dopo image-processing (vedi fig.3.3) questa viene rilevata nelle prime posizioni. Osserviamo inoltre come nel caso di immagini normali ed entrambi gli scheletri le immagini 1565030 e 1352713 appaiono nelle prime posizioni; d'altra parte se torniamo a vedere i recuperi in fig.3.1 e 3.3 vediamo che i precedenti due marchi sono rilevati molto lontani. Sembra che gli scheletri di queste immagini vengano viste più vicino a 2010916 che a 1138103, alla cui classe di rilevanti effettivamente appartengono. Osserviamo come dopo image-processing i marchi 1516142, 2043261 vengano subito riconosciuti.

Immagini normali:

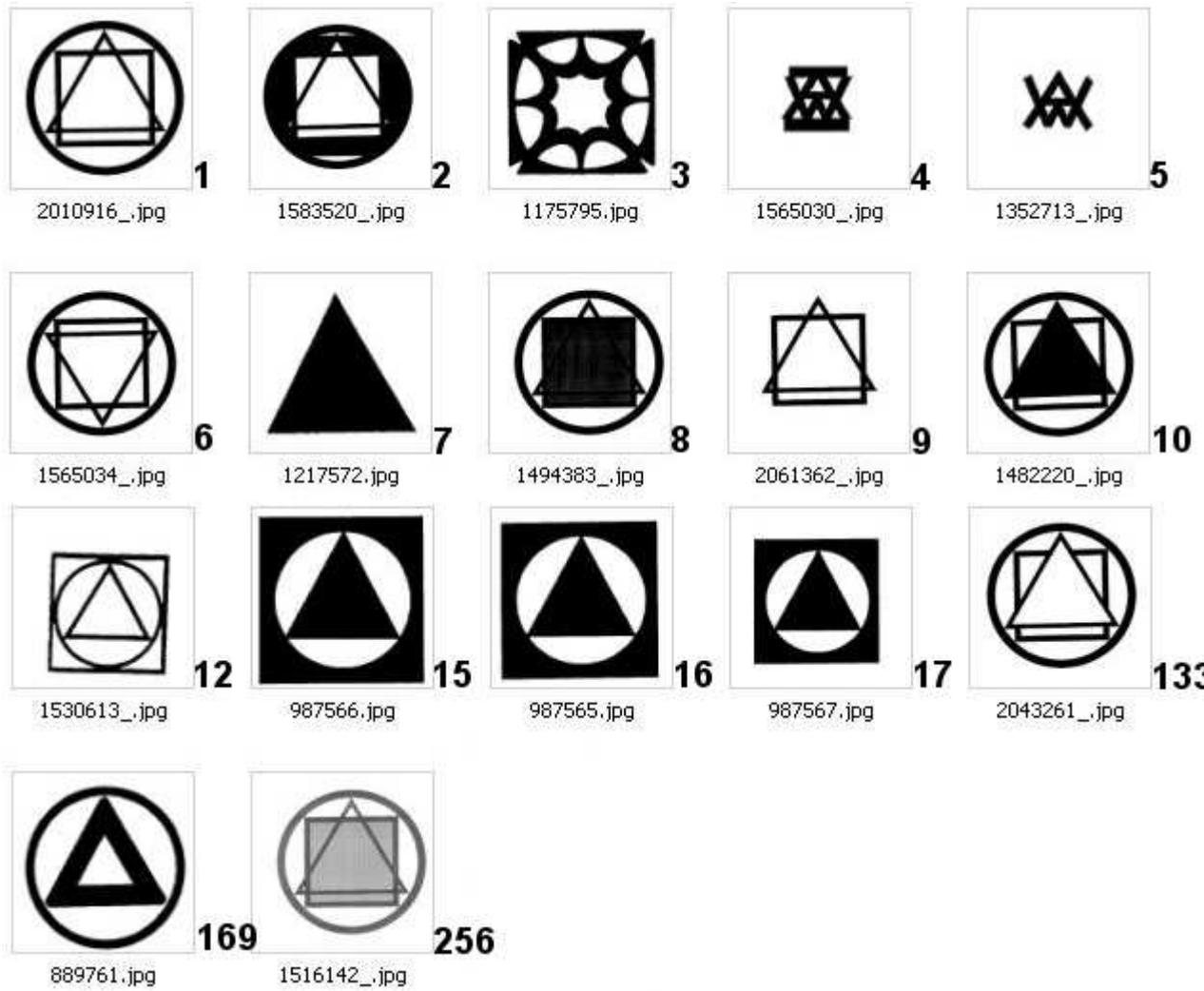


Figura 3.11:

Mediante scheletri:

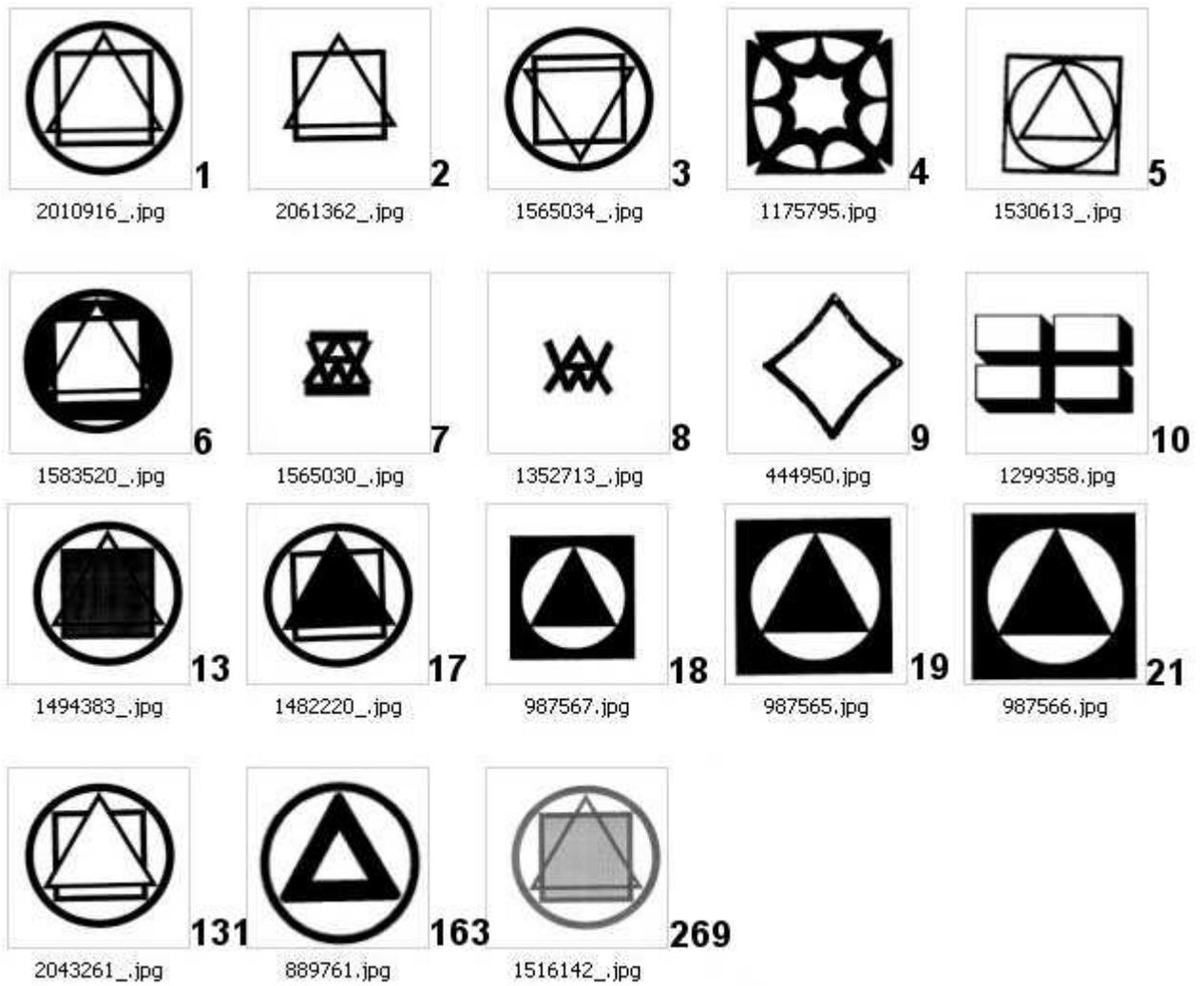


Figura 3.12:

Mediante contorni:



Figura 3.13:

Mediante scheletro dopo image-processing:

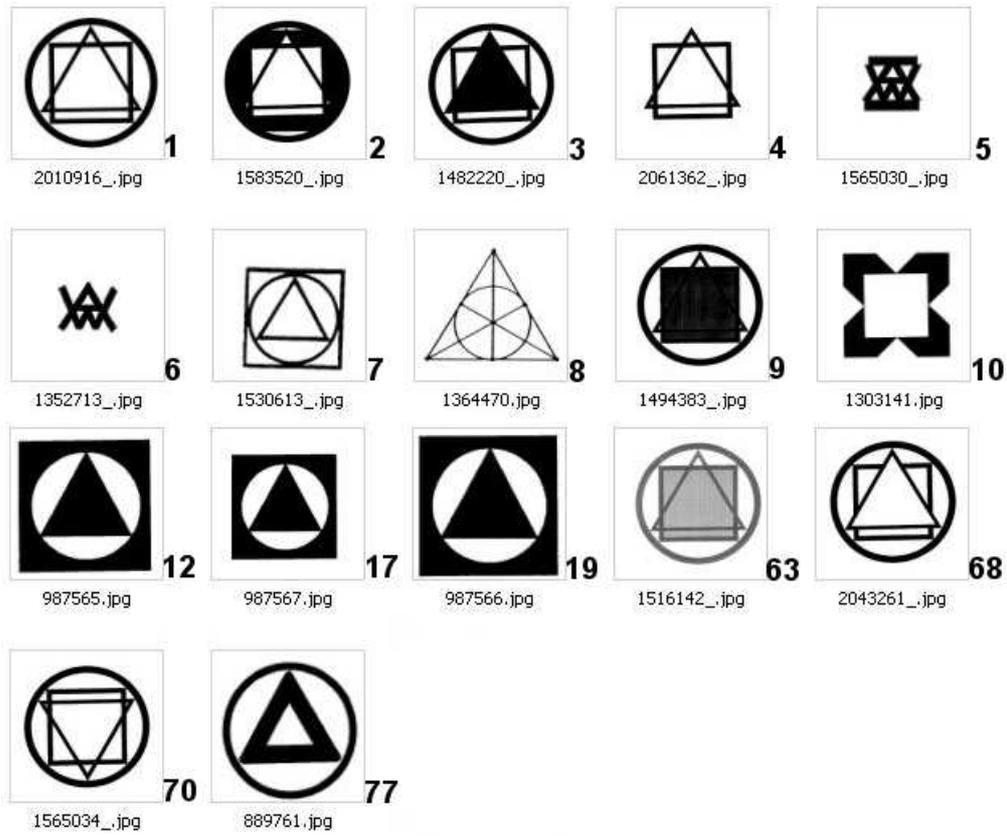


Figura 3.14:

Mediante contorni dopo image-processing:

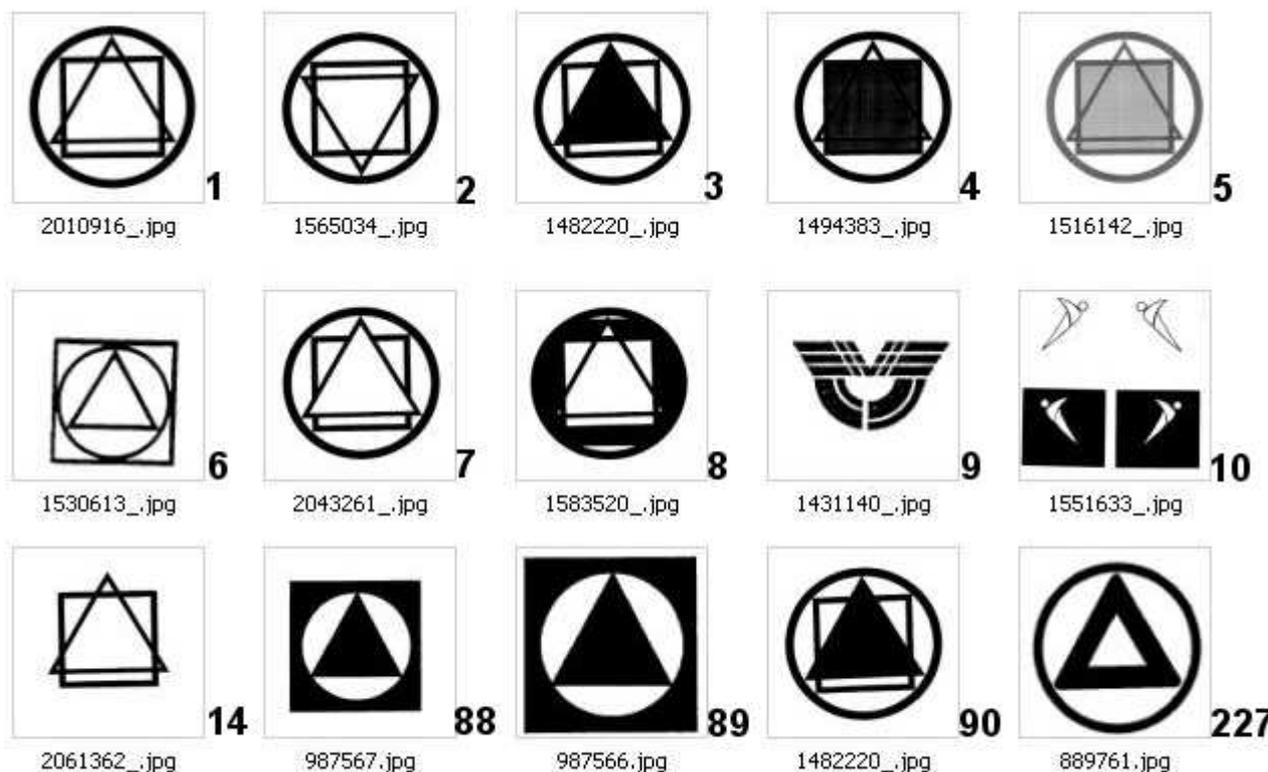


Figura 3.15:

Casi particolari

Ci sono stati alcuni marchi per cui non è stato possibile eseguire il recupero, in particolare il 1216236 e il 1516142. Questi due marchi sono delle tessiture, cioè nuvole di pixel più o meno dense che conferiscono all'immagine tonalità di grigio pur trattandosi sempre di un'immagine binaria. La conseguenza di questo è la presenza di un numero altissimo di componenti connesse che si riflettono sulle funzioni di taglia sotto forma di un insieme molto numeroso di rette angolari (cioè punti di discontinuità); per questo motivo si sono verificati overflow di memoria ed errori nel calcolare le distanze con le altre size function. Su questi due immagini più che mai è necessario un'operazione di image-processing che va oltre la riduzione del rumore.

Riportiamo ora nel grafico 3.3 la relazione tra le grandezze di precisione e recall medie calcolate su un campione di 40 marchi presi da tutte le quattro classi. Inoltre riportiamo in tabella 3.3 i valori medi, massimi e minimi del \widetilde{Rank} e il numero del marchio per cui si ottengono i valori di massimo e minimo.

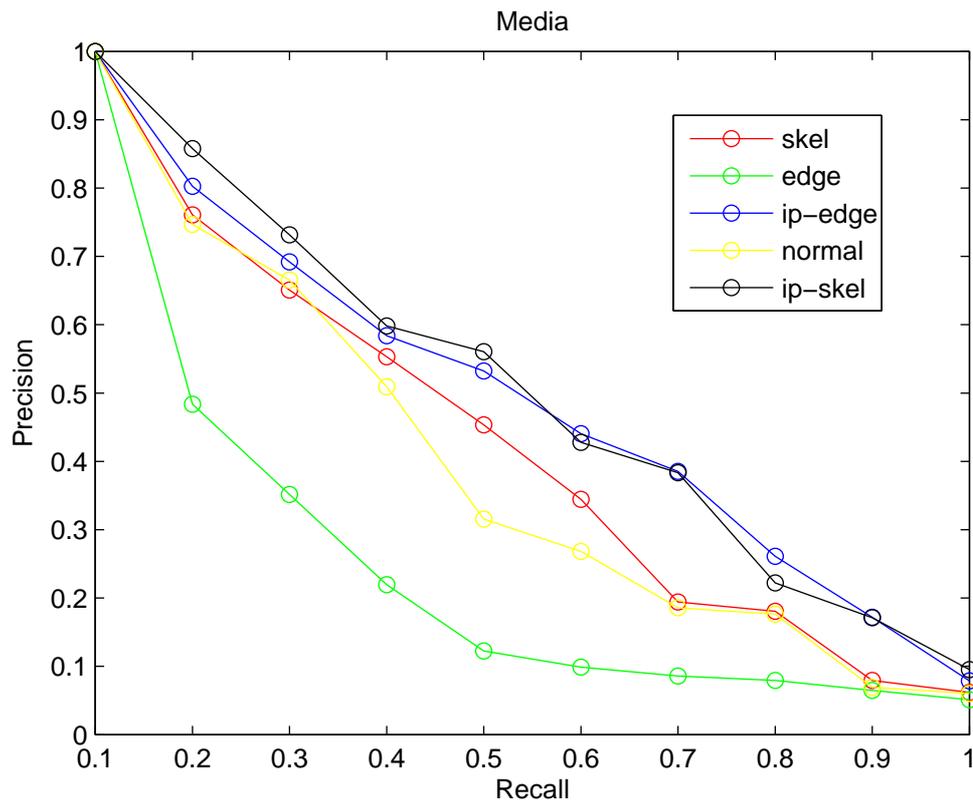


Figura 3.16:

<i>Rank</i>	medio	min	max	nome-min	nome-max
normali	0.2284	0.0871	0.7445	1576428	889761
scheltri	0.2286	0.0682	0.7610	1581395	889761
contorni	0.2616	0.1019	0.6537	1575271	1576428
scheletri-ip	0.1295	0.0482	0.4257	2061362	1164811
contorni-ip	0.1779	0.0244	0.5467	1138107	889761

Tabella 3.1:

3.4 Scheletrizzazione vs. Rilevamento dei Bordi

Abbiamo già visto nel capitolo I come la scheletrizzazione sia un'operazione morfologica suscettibile al rumore, in particolare come un singolo pixel fuori posto, possa generare nello scheletro un grave imperfezione. Per questo motivo ci aspetteremmo che i recuperi mediante scheltri siano poco precisi. Tuttavia, come mostra il grafico in fig 3.3, non sembra essere così. Osserviamo infatti come la precisione relativa agli scheletri sia uguale e in certi punti maggiore di quella relativa alle immagini normali. D'altra parte se osserviamo la curva relativa ai contorni, notiamo che è drasticamente peggiore di entrambi. Questa tendenza è evidente anche nella tabella 3.3 dove il valori di Rank medio per scheltri e immagini normali è sostanzialmente uguale; decisamente maggiore (e quindi peggiore) è quello relativo ai contorni. Si direbbe che l'operazione di rilevamento dei contorni non sia un buon metodo per fare dei recuperi e che in tal senso la scheletrizzazione è un metodo molto più affidabile. Se consideriamo le curve relative ai contorni e scheletri dopo image-processing, notiamo come queste siano piuttosto simili e decisamente migliori delle immagini normali, in particolare sottolineiamo la distanza tra contorni prima e dopo image-processing. In base a questo, possiamo pensare che rispetto alla funzione di taglia da noi usata, il rilevamento dei contorni sia più sensibile al rumore di quanto non sia la scheletrizzazione. Analizziamo ora un po' più approfonditamente gli effetti delle operazioni morfologiche a prescindere dalla distinzione dei rilevanti. Abbiamo preso in considerazione i primo 20 recuperati di una query, dopodichè abbiamo contato quanti di questi vengono recuperati a loro volta nelle prime 20 posizione sui rimanenti dataset. Il confronto è stato svolto su un campione di 20 marchi appartenenti al dataset delle immagini normali, non necessariamente contenuti in una delle quattro classi di rilevanti. In particolare abbiamo notato che i recuperati in comune con gli scheletri sono più del 60% , e meno del 20% per i contorni.

Questo può significare che la scheletrizzazione mantiene quasi invariate alcune proprietà intrinseche che vengono lette dalle funzioni di taglia.

3.5 Rumore vs Rimozione del Rumore

L'immagine processing è un'operazione che ha un costo elevato, riuscire a svincolare qualsiasi metodo di confronto tra immagini da questo sarebbe una vittoria. Se da una parte la scheletrizzazione e il rilevamento dei contorni sono procedure automatiche, così non è per la rimozione di imperfezioni poichè necessita della presenza di un utente che valuti quali operazione eseguire. Ristrettamente al nostro caso possiamo fare alcune considerazioni. Nel caso si voglia fare recuperi mediante scheletri, allora la rimozione del rumore porta un discreto miglioramento dei risultati, tuttavia abbiamo osservato che la scheletrizzazione è già un metodo di recupero abbastanza stabile. Nel caso si voglia fare recuperi mediante contorni, allora la rimozione del rumore diventa una procedura assolutamente necessaria. Anche qui abbiamo fatto un confronto, come spiegato nel paragrafo precedente, tra recuperati nei dataset ottenuti con le stesse operazioni morfologiche. I marchi comuni nei recuperi mediante le scheletrizzazioni sono circa il 35%, mentre per i contorni sono la metà (17%), questo può significare che l'azione di rimozione del rumore sui contorni è effettivamente molto influente poichè non solo modifica sensibilmente i rilevanti ma in generale tutte le immagini.

3.6 Conclusioni

Le funzioni di taglia forniscono un metodo valido per il riconoscimento delle immagini. La potenza di questo metodo rispetto ad altri sta nella discreta resistenza al rumore. Questa caratteristica può essere impiegata nella costruzione di un programma efficiente per la ricerca di marchi di fabbrica in database di grandi dimensioni. Non può considerarsi un metodo risolutivo, ma certamente utilizzandolo a livello di test può far risparmiare tempo, denaro e complicazioni a chi vuole registrare un proprio marchio. Le operazioni morfologiche eseguite a monte del calcolo delle funzioni di taglia hanno dato risultati differenti. L'utilizzo dei bordi è risultato alquanto instabile e sensibile alle imperfezioni, contrariamente a ciò che ci saremo aspettati; L'utilizzo degli scheletri si è rilevato stabile e resistente al rumore; Affiancare il test mediante scheletri a quello sulle immagini normali porterebbe un grado di affidabilità in più alla ricerca su database. Infine possiamo dire che la pulizia delle immagini aumenta sensibilmente la precisione di recupero; tuttavia ai fini di un utilizzo pratico ha un costo, poichè necessita che le immagini vengano controllate almeno una volta una ad una per fare le opportune correzioni.

Elenco delle figure

1.1	A sinistra la struttura tipo N_8 a destra quella tipo N_4	8
1.2	A sinistra una dilatazione ottenuta con N_4 , a destra una ottenuta con N_8	8
1.3	sopra le immagini primitive e sotto i rispettivi scheletri	10
1.4	A sinistra $B=N_8$, al centro B_1 e a destra B_2	10
2.1	Fig 1.1	13
2.2	A sinistra l'immagine in assenza di rumore e a destra la relativa Funzione di Taglia	18
2.3	A sinistra l'immagine con presenza di rumore e a destra la relativa Funzione di Taglia	19
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.	20
3.1	24
3.2	25
3.3	26
3.4	27
3.5	28
3.6	30
3.7	31
3.8	32
3.9	33
3.10	35
3.11	36
3.12	37
3.13	38
3.14	39
3.15	40
3.16	41

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.
- [3] A.Verri, C. Uras, P.Frosini, M.Ferri: *On the use of size function for shape analysis.*
- [4] W.Rucklidge: *efficient visual recognition using the Hausdorff distance* volume 1173 of *Lecture notes in computer science* Springer.
- [5] Henk J.A.M. Heijmans: *Matematical Morphology as a Tool for Shape Description* Pag 147. 'Shape in Picture', NATO ASI Series, Computer and System Science Vol.126.