



UNIVERSITÀ DEGLI STUDI DI PADOVA

DEPARTMENT OF
INFORMATION
ENGINEERING
UNIVERSITY OF PADOVA



**FACOLTÀ DI INGEGNERIA
CORSO DI LAUREA IN INGEGNERIA DELL'INFORMAZIONE**

Tesi di laurea Triennale

Registrazione automatica di immagini dermoscopiche

Automatic registration of
dermoscopic images

Studente:
Marco Dalla Santa

Relatore:
Chia.mo Prof. Enoch Peserico
Stecchini Negri De Salvi

Indice

Indice	1
1 Introduzione medica	3
1.1 Nevi e melanomi	3
1.2 La prevenzione	4
2 Registrazione di Immagini	7
2.1 Introduzione	7
2.2 Descrizione del problema	9
2.3 Strategia di realizzazione	9
Cross-correlazione	10
Approccio multiscala	12
2.4 Struttura dell'algoritmo	13
Conversione in scala di grigi	13
Ricerca del nevo	14
Contrasto localizzato	15
Ritaglio	16
Contrasto automatico	17
Sogliatura e media nulla	18
Zeropadding	18
Registrazione	20
Ultimo livello	20
Livelli intermedi	22
Primo livello	23
3 Aspetti implementativi	25
3.1 Linguaggio e librerie	25
3.2 Organizzazione del progetto	27
4 Prestazioni	29
4.1 Robustezza e precisione	29
4.2 Tempi di esecuzione	30
4.3 Trades off	31
Livelli intermedi	31
Primo livello	33
4.4 Conclusioni	35

Introduzione medica

1.1 Nevi e melanomi

I nevi (o nei) sono macchie cutanee, localizzate e persistenti, più o meno rilevate rispetto alla superficie della pelle. Possono essere dovuti ad un insieme anomalo di vasi sanguigni (nevi vascolari), ad accumulo di sostanza pigmentata (melanina) come nel caso di lentiggini ed efelidi oppure alla proliferazione dei melanociti, ossia le cellule che producono la melanina (nevi melanocitici). La colorazione è variabile: quella prevalente è marrone in diverse tonalità, ma può essere anche nera, blu, rosso-violacea oppure con un alone bianco, mentre le dimensioni variano da pochi millimetri a qualche centimetro. I nevi possono essere presenti in numero variabile già alla nascita, ma continuano a comparire a tutte le età, soprattutto in relazione all'esposizione alla luce solare. Durante la vita possono subire modificazioni: possono aumentare leggermente di dimensioni e diventare nodulari oppure regredire fino alla scomparsa. In genere i nevi sono asintomatici e non vanno trattati, tuttavia l'escissione chirurgica è necessaria quando questi diventano sintomatici (comparsa di prurito, dolore) o cambiano rapidamente le loro caratteristiche morfologiche: si tratta infatti di segni di una possibile degenerazione maligna, in particolare se si presentano nei nevi di recente comparsa. Sono in particolare i nevi melanocitici che possono portare alla formazione di un melanoma, un tumore cutaneo maligno caratterizzato da un'elevatissima propensione alla metastasi, ovvero alla diffusione delle cellule maligne a tutti i tessuti dell'organismo.



Figura 1.1: Nevo melanocitico



Figura 1.2: Melanoma

La terapia prevede l'asportazione completa della lesione e di parte della superficie cutanea circostante. Se il melanoma è rimasto confinato agli strati cutanei superficiali, la prognosi è generalmente buona, con guarigione del paziente, viceversa, se il melanoma ha raggiunto gli strati più profondi

perché ha avuto molto tempo di accrescersi prima della sua identificazione ed asportazione, i rischi di vita per il paziente sono molto elevati.

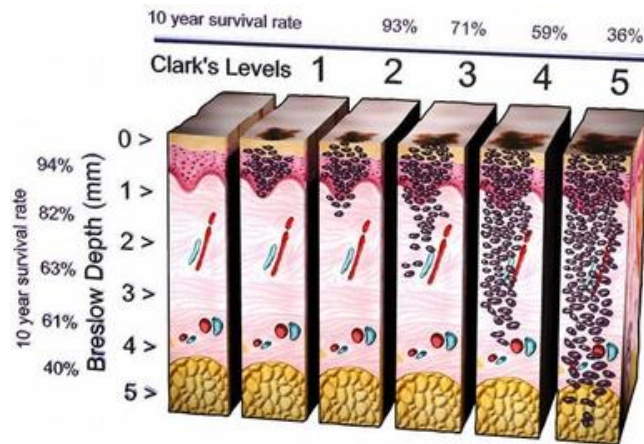


Figura 1.3: Evoluzione e tasso di mortalità del melanoma

1.2 La prevenzione

L'evoluzione del melanoma è strettamente legata al suo spessore, per cui la diagnosi precoce è fondamentale, individuare il melanoma quanto più precocemente possibile rappresenta la principale arma per tentare di ridurre la mortalità. Non esistono tuttavia criteri assoluti per differenziare i nevi nevocellulari dal melanoma: dimensioni inferiori al centimetro, limiti periferici netti, uniformità del colore e assenza di desquamazione, di fenomeni infiammatori e di disturbi soggettivi sono elementi in linea di massima a favore della natura nevica di una lesione mentre sono da considerare sospetti i nevi con margini frastagliati, forma e pigmentazione irregolari, ulcerazioni e segni di infiammazione circostante, quelli che cambiano improvvisamente di dimensioni e consistenza o nei quali la pigmentazione tende a diffondersi alla cute circostante. La prevenzione del melanoma prevede l'utilizzo della cosiddetta regola ABCDE che consiste nel monitorare periodicamente, soprattutto nei nevi a rischio, le seguenti proprietà:

1. Asimmetria: i melanomi sono generalmente asimmetrici;
2. Bordi: i bordi sono irregolari e frastagliati al contrario di quelli dei nevi;

3. Colore: il melanoma è spesso policromo, ossia di colori diversi e disomogenei;
4. Dimensione: una lesione di diametro superiore ai 6mm è da considerarsi sospetta;
5. Evoluzione: cambiamenti di forma, colore e superficie sono sintomi di rischio.

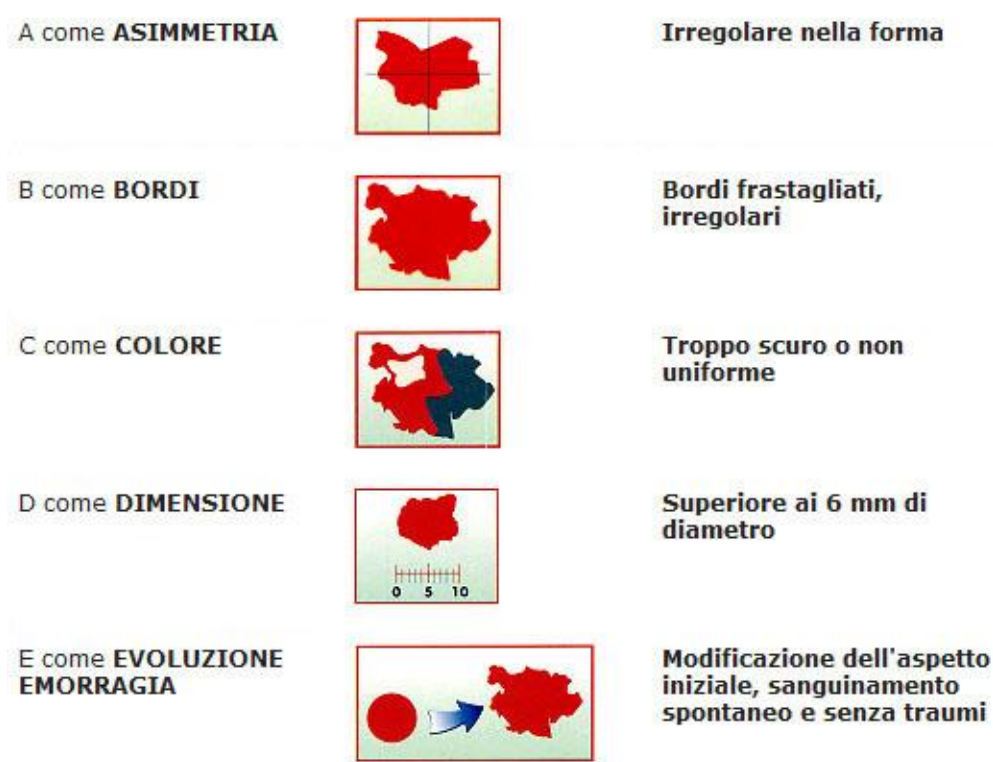


Figura 1.4: Regola ABCDE

Queste caratteristiche devono essere monitorate da un dermatologo per permettere una diagnosi corretta, tuttavia il supporto fornito da un sistema informatizzato che segnala le situazioni a rischio da sottoporre all'attenzione medica e quelle invece stazionarie permette di ridurre il carico di lavoro del medico e quindi di porre un'attenzione maggiore ai casi da monitorare con controlli più frequenti. Un sistema affidabile che affianchi il medico svolgendo la parte di lavoro che non richiede necessariamente l'intervento umano, alleggerendone così il carico lavorativo, viene ad essere particolarmente importante in casi come questo dove è fondamentale la diagnosi precoce.

Registrazione di Immagini

2.1 Introduzione

Negli ultimi anni in ambito biomedico il progressivo aumento nell'uso di sistemi hardware sempre più sofisticati per l'acquisizione di immagini radiologiche a fini diagnostici e la sempre crescente dimensione e risoluzione (sia spaziale sia temporale) delle informazioni acquisite ha creato la necessità di sistemi software in grado di supportare i medici ed i radiologi nell'interpretazione e sfruttamento proficuo della grande quantità di informazione disponibile.

Una procedura di importanza critica e sulla quale sono stati condotti numerosi studi è la registrazione di immagini: il termine si riferisce al processo di stima di una trasformazione spaziale ottimale che permetta di sovrapporre due o più immagini della stessa scena prese in tempi diversi, da diversi punti di vista o con diversi sensori. In altri termini, l'obiettivo di un algoritmo di registrazione è quello di mappare i punti di un'immagine nei corrispondenti punti dell'immagine che vogliamo registrare permettendo di confrontare o integrare i dati ottenuti da diverse misure ed evidenziando così ogni possibile cambiamento in dimensioni, forma o posizione.

Il processo di registrazione di immagini è molto variegato per poter soddisfare al meglio le necessità di ogni problema specifico.

Per confrontare tra loro le immagini esistono due diversi procedimenti principali: il primo metodo è basato sul confronto di regolarità nell'intensità dei segnali mentre il secondo sulla ricerca di corrispondenze tra caratteristiche salienti dell'immagine come punti, linee e contorni. Entrambi possono essere applicati all'immagine intera o a sotto immagini.

Si applicano poi vari modelli di trasformazione per correlare lo spazio dell'immagine bersaglio a quello dell'immagine di riferimento: possono essere trasformazioni lineari come traslazione, rotazione, scalatura, che generano registrazioni cosiddette rigide oppure trasformazioni non lineari come funzioni a base radiale o modelli di deformazione dette elastiche.

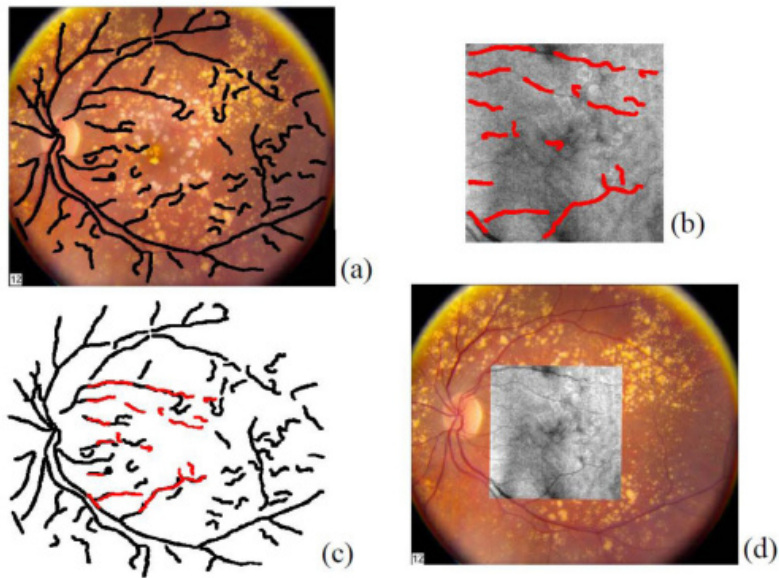


Figura 2.1: Esempio di registrazione tramite elementi grafici di riferimento (retina umana)

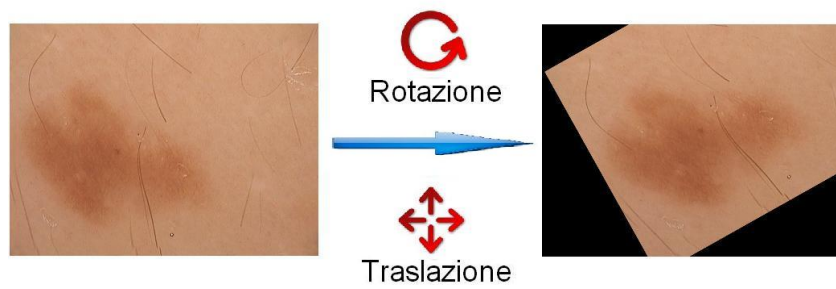


Figura 2.2: Esempio di trasformazione rigida



Figura 2.3: Esempio di trasformazione elastica

2.2 Descrizione del problema

Nel caso in esame si vuole produrre un sistema informatizzato che rappresenti uno storico dei nevi a scopo preventivo evidenziando i possibili rischi di loro degenerazioni maligne. Quindi il tipo di registrazione che si vuole realizzare è pensato per immagini dermoscopiche (bidimensionali) raffiguranti il medesimo nevo a distanza temporale, con le inevitabili differenze di posizionamento e rotazione che ne derivano. Le foto sono ottenute da dermatoscopi o smartphone dotati di un'apposita ottica, quindi hanno dimensione costante. La procedura prevederà perciò un confronto basato sull'intensità vista la natura dei nevi, spesso privi di veri e propri punti di riferimento grafici, ed una trasformazione di tipo rigido che calcoli rotazione e traslazione ottimali di una seconda immagine (nuova) rispetto ad una di riferimento (precedente) senza operazioni di scalatura, vista anche la necessità di notare un'eventuale crescita avvenuta nel tempo.

2.3 Strategia di realizzazione

Per ottenere una registrazione rigida dal confronto dell'intensità tra due immagini simili l'idea è quella di definire una funzione di similitudine tra le due immagini e confrontarne i risultati ottenuti per ogni possibile combinazione di rotazione e traslazione. Le possibilità più valide spesso utilizzate sono le funzioni sommatoria dei quadrati, o dei moduli, delle differenze tra intensità dei corrispondenti pixel ad esempio:

$$S = \frac{1}{A} \sum_{x=1}^{L_x} \sum_{y=1}^{L_y} |I_1(x, y) - I_2(x, y)|$$

dove A rappresenta l'area, I_1 e I_2 le immagini, L_x e L_y le dimensioni laterali. Questa mostrerà un minimo in corrispondenza della massima somiglianza tra i singoli punti delle due figure, quindi della registrazione ottimale. Tuttavia soluzioni simili che operano un confronto direttamente sulle immagini, pur portando generalmente a risultati qualitativamente soddisfacenti, necessitano di particolare attenzione poiché sono particolarmente sensibili al rumore, tipico nelle immagini mediche, ed inoltre risultano molto dispendiose in termini computazionali. Infatti eseguire una simile operazione su immagini di un MegaPixel, dove si compiono spostamenti di un pixel per volta e la rotazione minima è calcolata come lo spostamento di un solo pixel all'estremità della foto, comporterebbe l'analisi di circa quattro milioni

di traslazioni per oltre 6250 rotazioni diverse, per un totale di confronti superiore ai 25 miliardi.

Cross-correlazione

La soluzione a tali inconvenienti viene dalle trasformate integrali, spostando lo studio dal dominio spaziale a quello delle frequenze. L'analisi precedente di tutte le traslazioni tra immagini può così essere drasticamente semplificata e ridotta ad un'unica operazione: la funzione di correlazione incrociata o cross-correlazione. Essa rappresenta in generale la misura di similitudine tra due segnali come funzione di uno spostamento applicata ad uno di essi. Date due immagini la correlazione incrociata in due dimensioni esamina tutte le possibili traslazioni relative ed assegna a ciascuna un valore proporzionale alla somiglianza delle immagini correnti, mostrando in corrispondenza della massima similitudine un picco di intensità proporzionale al loro grado di affinità.

Nel dettaglio la correlazione incrociata si ottiene con il seguente procedimento: date due immagini puramente reali i e j per prima cosa si calcola la trasformata di Fourier discreta bidimensionale di entrambe utilizzando la FFT (Fast Fourier Transform).

$$I = \mathcal{F}\{i\}, \quad J = \mathcal{F}\{j\} \quad I, J \in \mathbb{C}$$

Dalle due trasformate tramite la loro moltiplicazione complessa normalizzata si ottiene una densità spettrale di potenza incrociata.

$$H = \frac{I \cdot J}{|I \cdot J|} \quad H \in \mathbb{C}$$

Si determina la figura di cross-correlazione normalizzata applicando la trasformata di Fourier inversa.

$$h = \mathcal{F}^{-1}\{H\} \quad h \in \mathbb{R}$$

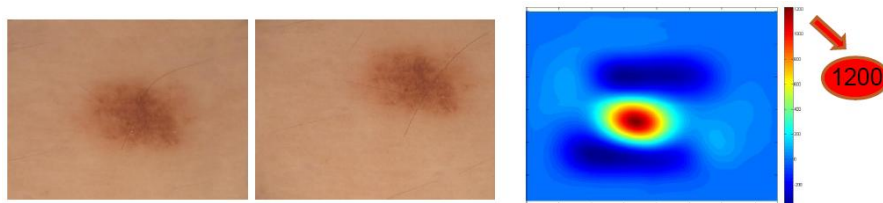
Questa presenta un picco chiaramente visibile che indica somiglianza relativa (intensità) e traslazione della seconda immagine rispetto alla prima (posizione). Localizzando la posizione del picco si ottiene la traslazione ottimale per sovrapporre la seconda immagine alla prima:

$$(\Delta x, \Delta y) = \underset{(x,y)}{\operatorname{argmax}}\{h\}$$

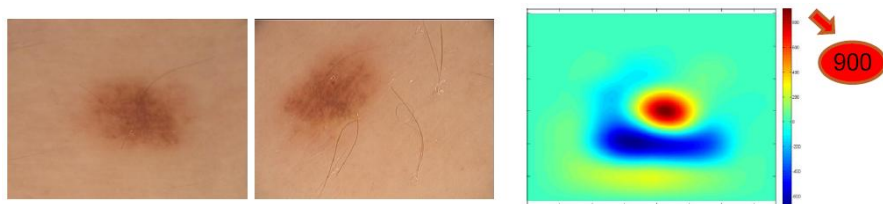
Per immagini con la stessa angolazione la collocazione del massimo della

funzione permette perciò di risalire direttamente alla traslazione ottimale in un unico passaggio. Se invece si vuole individuare prima la rotazione si esegue più volte la cross-correlazione tra l'immagine di riferimento e quella da registrare ruotata in tutte le sue possibili angolazioni; si procede poi al confronto tra i valori delle intensità dei picchi ottenuti per ogni angolo e si prende quello corrispondente al valore massimo. Poiché la correlazione incrociata fornisce valori di massimo proporzionali alla similitudine tra i segnali si vanno in questo caso a confrontare le intensità di questi senza prendere in considerazione la loro collocazione spaziale.

$$\theta_m = \underset{(\theta)}{\operatorname{argmax}} \left\{ \underset{(x,y)}{\operatorname{argmax}} \{ \operatorname{xcorr}[I_1, I_2.\operatorname{rotate}(\theta)] \} \right\} \quad \theta \in [0, 360]$$



(a) Cross-correlazione eseguita su immagini con uguale angolazione



(b) Cross-correlazione eseguita su immagini con diversa angolazione

Figura 2.4: Esempi grafici di risultati della correlazione incrociata, è evidente la differente intensità dei picchi

L'operazione descritta rassomiglia la convoluzione nel dominio spaziale, in questo modo però l'utilizzo della trasformata di Fourier veloce (FFT) porta un grande guadagno nella performance. Per trasformazioni semplici quali traslazione e rotazione, uno studio condotto nel dominio delle frequenze, risulta quindi in generale più performante ed accurato del corrispettivo nel dominio spaziale poiché resiste maggiormente al rumore e necessita di un numero molto minore di operazioni.

Approccio multiscala

Il secondo accorgimento per ridurre la quantità di calcoli derivanti dallo studio di tutte le possibili trasformazioni spaziali dell'immagine da registrare è un approccio a risoluzione multiscalata. Ciò significa che al primo stadio dell'analisi si utilizzano le immagini con risoluzione dimezzata N volte, riducendo sensibilmente il numero di rotazioni e traslazioni che vi si possono compiere. Quindi si passa progressivamente a risoluzioni crescenti, generalmente raddoppiandola, operando però una selezione tra i risultati ottenuti in precedenza e andando a studiare solo quelli più promettenti, perfezionandoli. Mantenendo di volta in volta solo i risultati migliori per le fasi successive si riduce allo stretto necessario il numero di analisi alle risoluzioni più elevate. Con questa tecnica si vanno a diminuire sia il numero di confronti sia la dimensione delle immagini analizzate nel corso della registrazione; ciò permette di abbattere drasticamente i tempi di esecuzione considerato l'aumento esponenziale del tempo di calcolo della FFT per immagini di dimensione crescente, senza tuttavia intaccarne la precisione grazie all'affinamento progressivo nella ricerca del risultato.

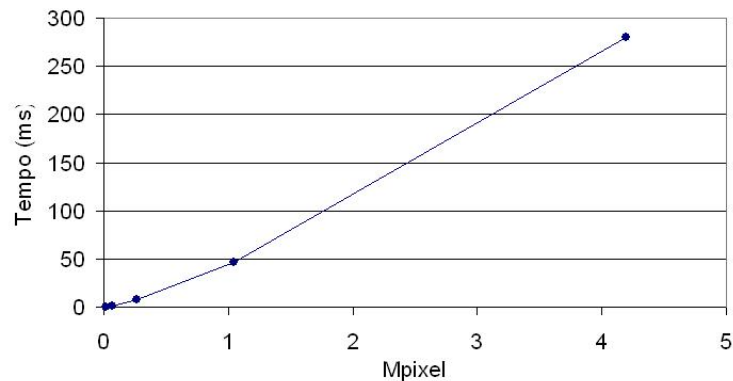


Figura 2.5: Tempi di esecuzione della FFT in funzione della risoluzione

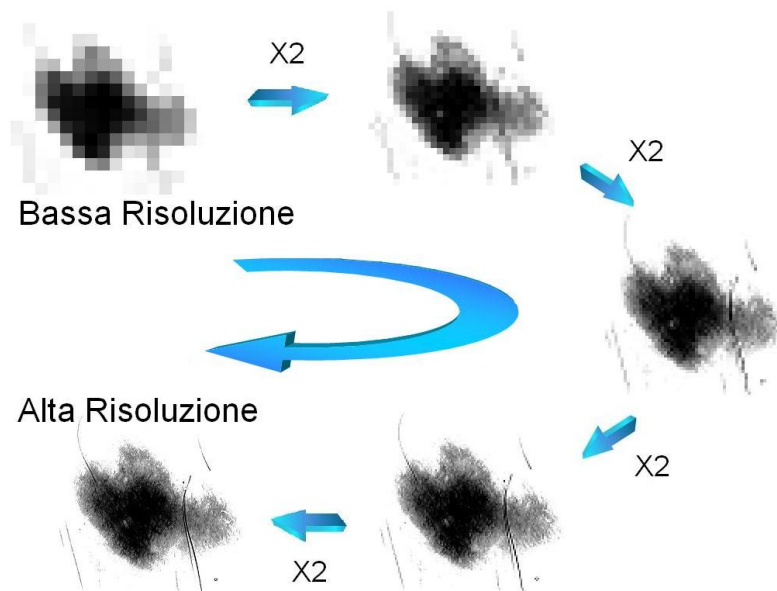


Figura 2.6: Logica a risoluzione multiscalata

2.4 Struttura dell'algoritmo

Prima di poter svolgere il processo vero e proprio di registrazione vengono effettuate sulle immagini diverse operazioni preliminari per migliorare la precisione e la robustezza dell'algoritmo. Inizialmente si lavora sulle immagini originali con operazioni valide per l'intero seguito del procedimento e che saranno quindi eseguite un'unica volta su ogni foto: queste vengono per prima cosa convertite in scala di grigi, dopodichè si localizza il nevo, se ne accentua il contrasto in un intorno e si crea un'immagine circoscritta allo stesso.

Conversione in scala di grigi

Innanzitutto le due immagini da registrare vengono convertite in scala di grigi poiché le informazioni sul colore non vengono considerate nel confronto tra le diverse foto non essendo necessarie nè affidabili nella localizzazione e registrazione del nevo. Il colore è infatti particolarmente soggetto a rumore e cambiamento tra un'immagine e l'altra e rischierebbe di essere più dannoso che utile in foto non scattate con strumenti di precisione e da addetti esperti. Saranno, se necessario, analisi mediche successive a valutare eventuali cambiamenti di pigmentazione. La conversione è fatta tramite il metodo di analisi delle componenti principali sulle componenti RGB la

quale fornisce un'immagine più marcata di quella ottenibile con una media pesata, ed una parziale eliminazione del rumore. Dopo la conversione in scala di grigi l'immagine ha una descrizione unidimensionale del colore ed il valore dei pixel varia tra [0] (nero) e [255] (bianco), con precisione a 32 bit. La seconda immagine viene poi capovolta orizzontalmente e verticalmente, o con risultato analogo ruotata di 180° , per compensare la rotazione che sarà introdotta in seguito nella convoluzione dovuta alla simmetria intrinseca della Trasformata di Fourier.

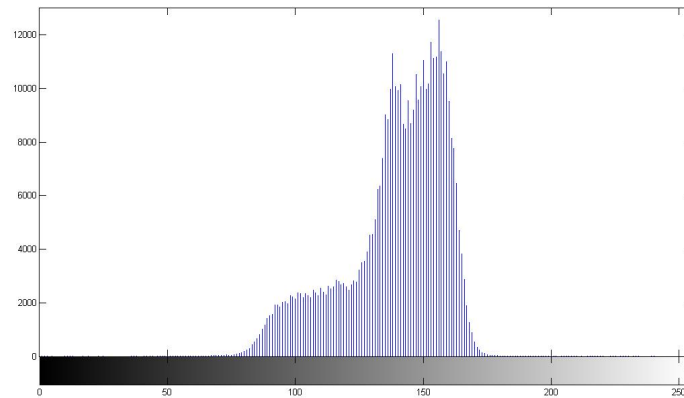


Figura 2.7: Istogramma dell'immagine convertita in scala di grigi, sono visibili due picchi corrispondenti alle diverse tonalità dello sfondo soggetto a rumore ed uno minore corrispondente al nevo

Ricerca del nevo

Viene quindi fatta una localizzazione del nevo cercando dimensione e posizione di un cerchio che meglio lo approssimi. A tal fine si studiano nell'immagine tutti i cerchi dati dalle possibili combinazioni di posizioni e dimensioni; ad ognuno si assegna un punteggio calcolato come differenza pesata tra la media del colore dei pixel all'interno della circonferenza e di quelli all'esterno. Dopo aver dimensionando opportunamente i fattori di peso nella differenza dei valori medi, effettuando prove su varie foto, si può isolare la regione con maggiore concentrazione di colore scegliendo la circonferenza con punteggio minore. Al raggio del cerchio trovato viene infine aggiunto un margine di sicurezza ottenendo così quella che sarà una buona approssimazione della collocazione del nevo.

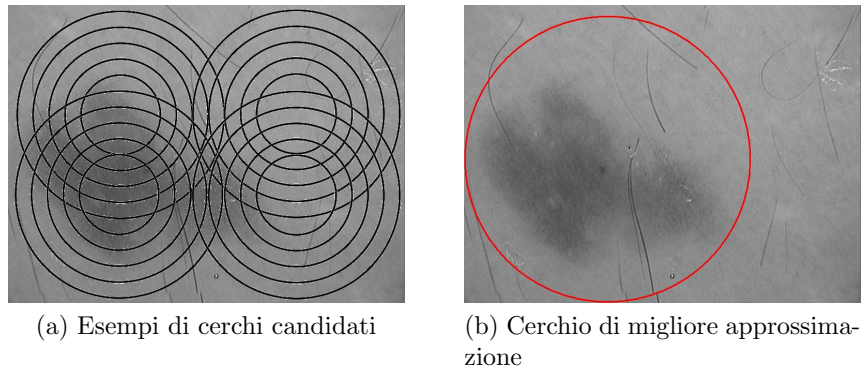


Figura 2.8: Procedura di localizzazione del nevo

Contrasto localizzato

Una volta nota la posizione di quest'ultimo, per farlo in risalto rispetto al resto dell'immagine ed appianare lo sfondo la si modifica moltiplicandone i singoli punti per una funzione di tipo gaussiano bidimensionale con origine nel centro del nevo e varianza data dal raggio del cerchio che ne descrive la dimensione. In questo modo si aumenta il contrasto tra la zona significativa dell'immagine, enfatizzata, e quella periferica, resa più omogenea dai bassi valori assunti dalla gaussiana fuori dal cerchio del nevo.

$$\left[A \cdot e^{-\left(\frac{(x-x_c)^2}{r_c^2} + \frac{(y-y_c)^2}{r_c^2}\right)} \right] \times (I(x,y) - \text{media}(I)) + \text{media}(I)$$

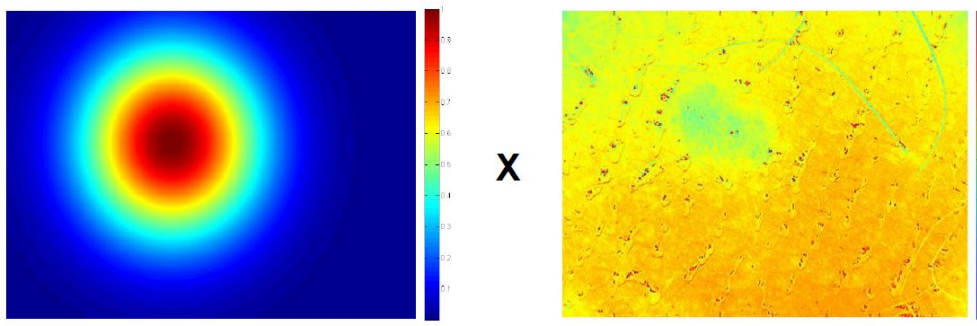


Figura 2.9: Funzione gaussiana di contrasto locale

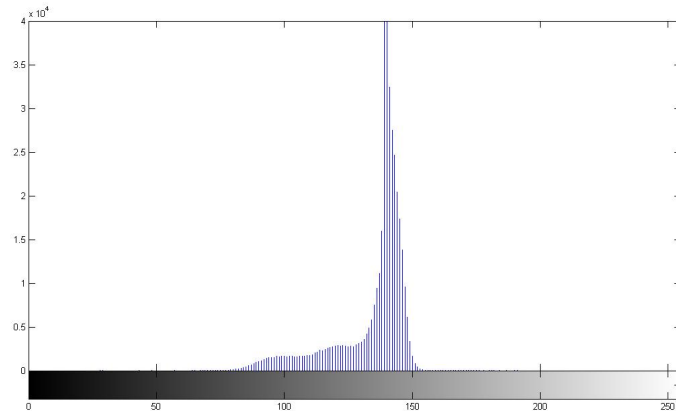


Figura 2.10: Istogramma dell'immagine sulla quale è stato enfatizzato il contrasto nella zona del nevo, è ben visibile un picco localizzato che corrisponde alla tonalità dello sfondo reso omogeneo

Ritaglio

Infine si ritaglia una sottoimmagine circoscritta al cerchio del nevo che sarà perciò di forma quadrata, o quasi a seconda della posizione di quest'ultimo nella foto, in modo da eliminare la parte eccedente dello sfondo circostante. Questa immagine sarà utilizzata nella ricerca dello spostamento angolare delle foto, dove lo sfondo in esubero è ininfluenza ai fini del calcolo della rotazione e anzi potrebbe solo aumentarne la complessità date le dimensioni maggiori.

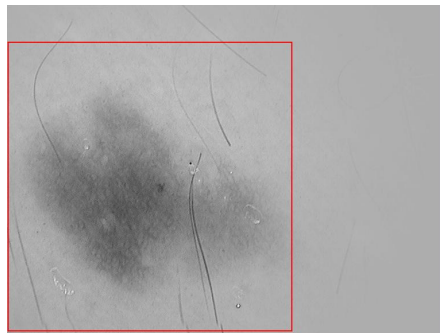


Figura 2.11: Esempio di ritaglio attorno al nevo

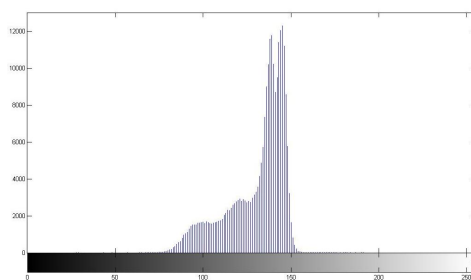
L'immagine ottenuta con questi primi passaggi è ottimizzata per essere sottoposta a registrazione. Si divide qui l'analisi in varie fasi durante le quali si risalirà dapprima alla rotazione utilizzando la logica multiscala esposta

in precedenza e quindi alla traslazione in un unico ultimo passaggio. In ogni fase prima di ricavarne le FFT le immagini necessitano di ulteriori necessari passaggi di elaborazione che vanno eseguiti dopo che queste sono state opportunamente ridimensionate a seconda del livello. Questi ultimi interventi servono per migliorare il contrasto globale, eliminare parte del rumore dall'immagine e renderle a media nulla per poter poi eseguire la correlazione incrociata su dati normalizzati.

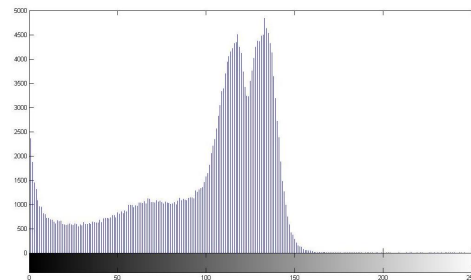
Contrasto automatico

Si modifica il contrasto della figura nel seguente modo: in principio si ricava un valore di soglia in base ad una data saturazione. Da questa con lo studio dell'istogramma dei livelli di grigio della figura si ricavano i valori di soglia minima e massima entro cui racchiudere l'immagine. A questo punto ogni pixel dell'immagine viene normalizzato entro un intervallo da 0 a 1 dividendolo per il contrasto massimo (la differenza tra la soglia superiore e quella inferiore) oppure viene fatto corrispondere ad uno degli estremi (0, 1) se fuori dai valori di soglia, in modo da eliminare tratti dovuti al rumore e livellare le regioni esterne di intensità simili.

$$I(x, y) = \begin{cases} 0, & \text{se } I(x, y) < hmin \\ 1, & \text{se } I(x, y) > hmax \\ \frac{I(x, y)}{hmax - hmin}, & \text{altrimenti} \end{cases}$$



(a) Istogramma dell'immagine ritagliata



(b) Istogramma dell'immagine dopo l'ottimizzazione del contrasto

Figura 2.12: Confronto fra gli istogrammi prima e dopo l'operazione di ottimizzazione del contrasto, si nota come l'istogramma ottenuto occupi una più ampia gamma di valori grazie alla riscalatura in funzione dei due valori di soglia

Sogliatura e media nulla

L'immagine viene poi resa a media nulla, per poter ottenere direttamente la cross-correlazione normalizzata, sottraendo ad ogni punto il valore medio della figura; anche in questo passaggio si opera una sogliatura azzerando i valori sopra ad una soglia massima. In questo modo si ottiene l'eliminazione del rumore chiaro, come possibili bolle presenti tra la lente ed il derma, ed il livellamento della regione di sfondo dell'immagine.

$$I(x, y) = \begin{cases} I(x, y) - \text{media}(I), & \text{se } I(x, y) \leq k \\ 0, & \text{se } I(x, y) > k \end{cases}$$

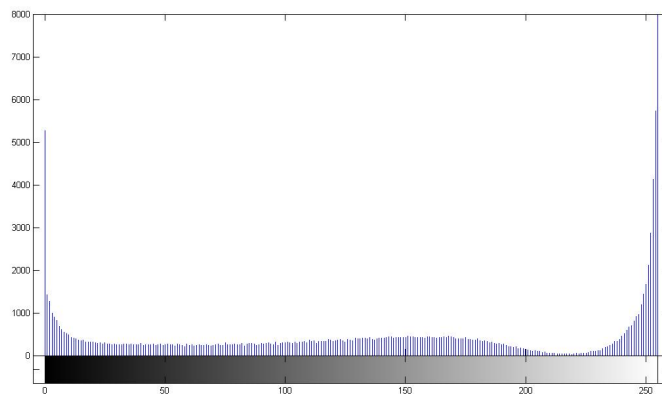


Figura 2.13: Istogramma dell'immagine dopo la resa a media nulla e sogliatura, si notano i due picchi relativi al nevo ed allo sfondo collocati alle estremità dei valori a causa della sottrazione della media, ed in particolare l'elevato valore del picco corrispondente alla tonalità bianca per opera della sogliatura

Zeropadding

Per ultima si applica una funzione di zero-padding, ossia l'aggiunta di uno sfondo neutro all'immagine per aumentarne la dimensione. Questa operazione serve per fare in modo che durante la correlazione incrociata, dove le immagini vengono considerate come segnali periodici, il prodotto convoluzionale non risenta di sovrapposizioni non volute tra i due, rendendola cioè un'operazione aciclica. Si deve perciò raggiungere una dimensione tale da permettere alla seconda immagine di traslare sulla prima in tutte le possibili

posizioni senza che la sua porzione significativa (non nulla) ne oltrepassi i lati, che vorrebbe dire sovrapporsi dal momento che verrebbe riproposta ciclicamente al lato opposto a quello eccedente. Si aggiunge perciò a entrambi i lati uno sfondo neutro di dimensione pari alla diagonale della seconda immagine, cioè la massima larghezza che assume questa durante le rotazioni, per scongiurare qualsiasi interferenza.

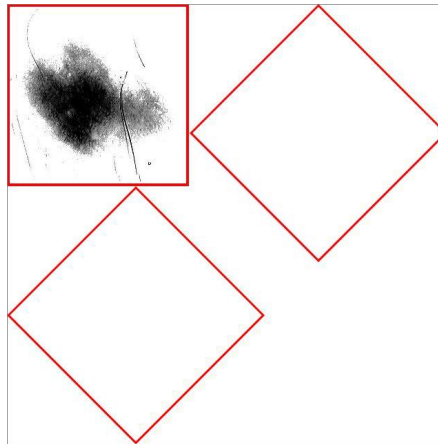


Figura 2.14: Procedura di zeropadding

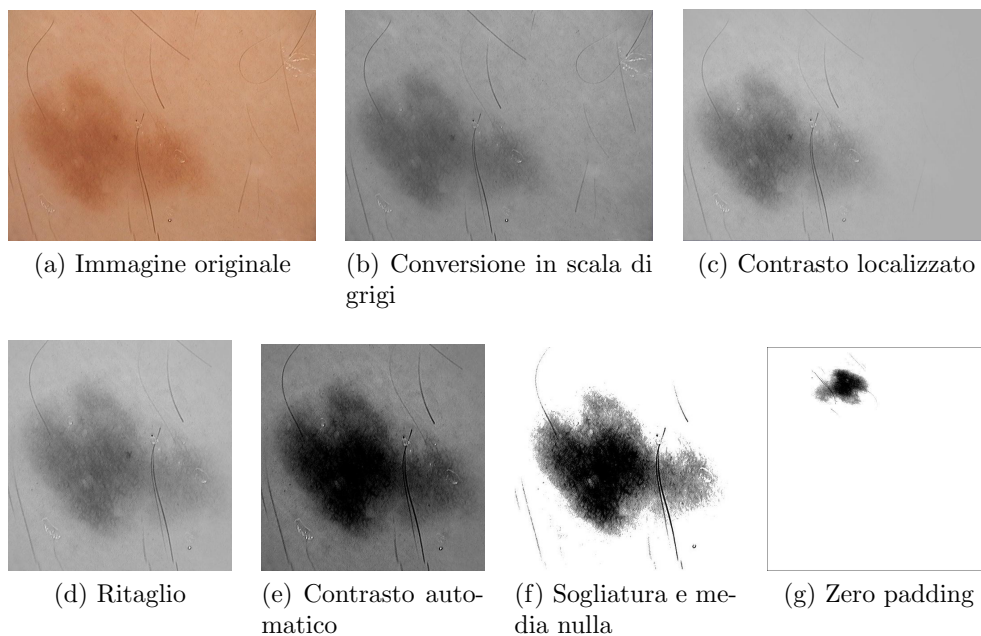


Figura 2.15: Fasi di pre elaborazione

Registrazione

A questo punto operando sulle immagini ottenute si può procedere alla loro registrazione. Questa seguirà la logica multiscalata esposta in precedenza ed eseguirà ad ogni livello di risoluzione, dalla minore alla massima, la correlazione incrociata selezionando i risultati migliori fino a trovarne uno definitivo.

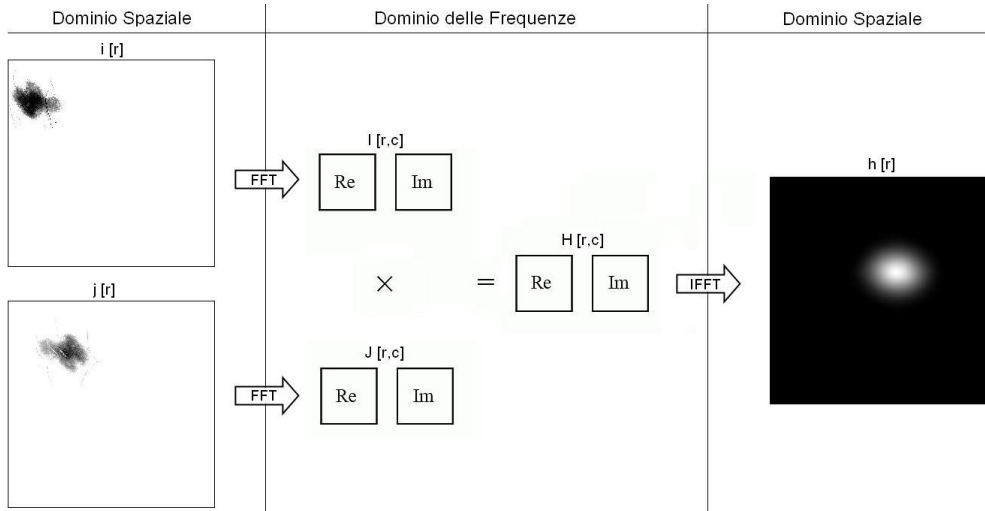


Figura 2.16: Schema della correlazione incrociata: le due immagini da registrare sono trasformate nel dominio delle frequenze tramite FFT. I rispettivi spettri sono moltiplicati e la FFT inversa di questo risultato è usata per tornare nel dominio spaziale. La posizione del picco nell'immagine risultante indica la traslazione ottimale.

Ultimo livello

Nel primo passaggio della registrazione si lavora con le immagini all'ultimo livello di risoluzione, cioè ridotte al minimo stabilito da un parametro determinato empiricamente, e tramite la correlazione incrociata si cerca tra tutte le possibili rotazioni della seconda immagine rispetto alla prima quella che massimizza la somiglianza tra le due. Le rotazioni a questo livello vanno a coprire i 360° con intervalli di angolo minimo che corrispondono allo spostamento di un pixel agli estremi dell'immagine. Questa rotazione minima apprezzabile sarà calcolata come:

$$\Delta = \arcsin \frac{1}{\max(L_x, L_y)}$$

Quindi isolando tra i valori di intensità i più elevati si ottengono gli angoli candidati per il risultato finale, in numero definito da un parametro che permetta un certo margine di sicurezza, determinato con una serie di prove su varie immagini, tipicamente otto valori per garantire la massima certezza del risultato.

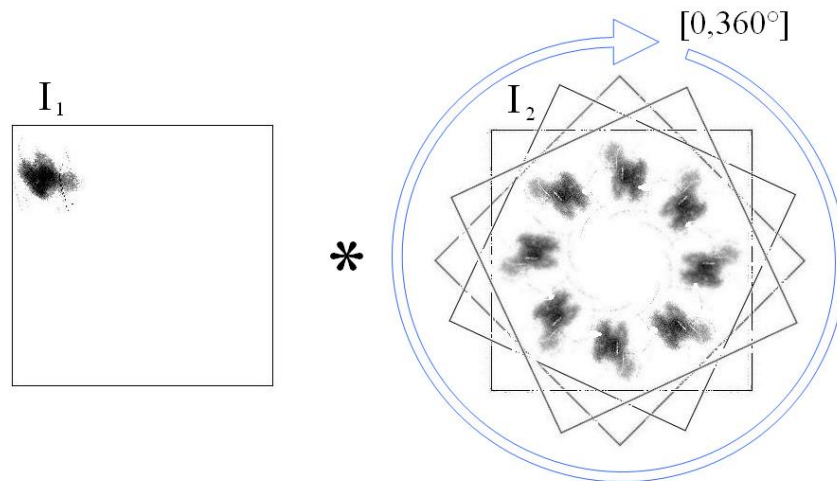


Figura 2.17: Esecuzione della cross-correlazione all'ultimo livello, in cui viene esaminato l'intero spettro angolare.

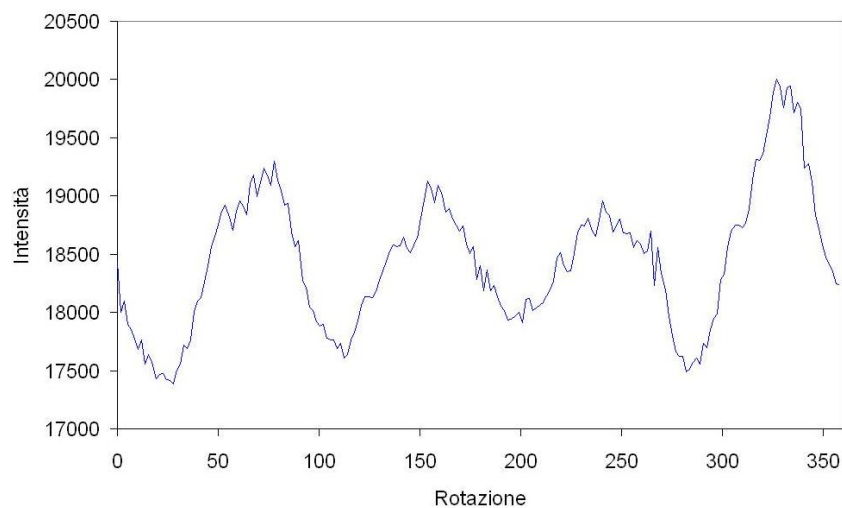


Figura 2.18: Risultato della correlazione incrociata in funzione dell'angolo in una rotazione completa ottenuta all'ultimo livello; dal grafico si notano gli angoli che massimizzano la similitudine ed una certa simmetria sui 180° .

Livelli intermedi

Nel passo successivo si riprendono i risultati selezionati al livello minimo e si ripete lo studio sulla corrispondenza delle due immagini con la cross-correlazione. Lo studio viene fatto negli angoli candidati e nei loro immediatamente adiacenti, poiché raddoppiando la risoluzione dell'immagine diminuisce anche l'angolo minimo, calcolato come in precedenza, e quindi si vanno ad esplorare rotazioni non ancora prese in esame. In modo analogo a quanto fatto in precedenza si selezionano gli angoli migliori dimezzandone progressivamente il numero data la maggiore affidabilità dell'operazione a risoluzione elevata; si itera tale il procedimento per risoluzioni crescenti fino ad arrivare all'immagine originale, o ad una dimensione cui corrisponde il grado di accuratezza desiderato. Alla conclusione di questa fase intermedia si è individuato un singolo angolo migliore degli altri che è perciò il risultato definitivo.

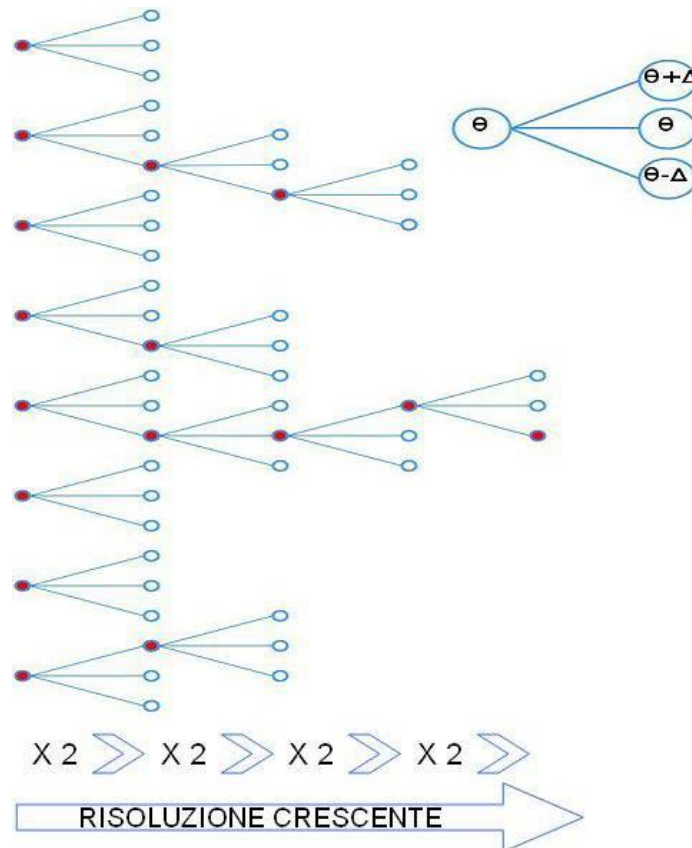
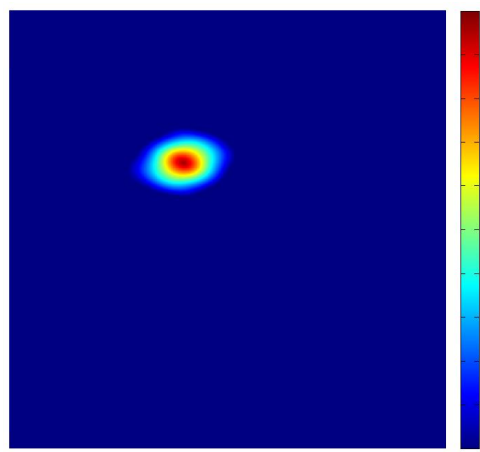


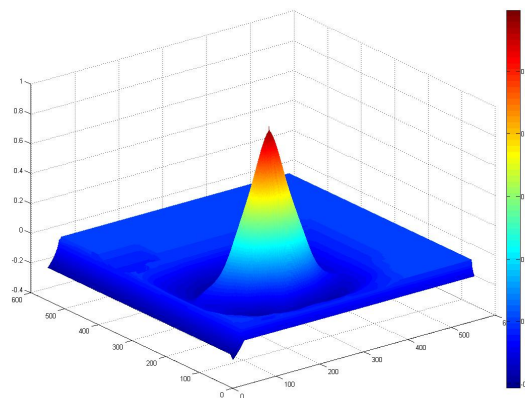
Figura 2.19: Selezione del risultato in logica a multiscala

Primo livello

A questo punto l'ultimo passo, per trovare la traslazione della seconda immagine, si svolge sempre con la correlazione incrociata fatta questa volta con le immagini nel formato originale, cioè non tagliate attorno al nevo, e ruotando la seconda dell'angolo trovato. Le coordinate del picco nell'immagine della cross-correlazione indicano la traslazione che massimizza la somiglianza reciproca.



(a)



(b)

Figura 2.20: Risultato della cross-correlazione eseguita sulle immagini con la stessa angolazione, mostra un picco chiaramente visibile

Aspetti implementativi

3.1 Linguaggio e librerie

Il linguaggio scelto per la realizzazione del progetto è il C++, poiché permette una certa versatilità nell'utilizzo grazie alla programmazione orientata agli oggetti, unita ad elevate prestazioni e compatibilità con sistemi Windows, Unix e mobile.

Per la gestione ed elaborazione delle immagini si è utilizzata la libreria open-source CImg che è di semplice utilizzo, ha una struttura dati delle immagini facilmente accessibile, fornisce un'esauriente insieme di funzioni di image-processing ed è di immediata compatibilità col progetto essendo scritta in codice C++. Questa libreria tuttavia ha presentato dei grossi limiti per quanto riguarda il calcolo della trasformata di Fourier discreta veloce (FFT) sulle immagini bidimensionali. Innanzitutto tale calcolo è applicabile solo ad immagini quadrate il cui lato sia una misura completamente fattorizzabile per due. La forma quadrata richiesta non rappresenta un problema visto che l'operazione di zeropadding viene fatta in due dimensioni restituendo un'immagine con lati uguali, al contrario il dover raggiungere il quadrato perfetto di due più prossimo obbliga a ridimensionamenti talvolta notevoli. A tali aspetti si aggiunge la lentezza del calcolo, privo di ottimizzazioni e che arriva ad essere la parte fortemente preponderante in termini temporali nell'esecuzione del programma. La ricerca di migliori performance passa perciò per un calcolo in tempi rapidi della FFT; dal punto di vista puramente prestazionale la scelta migliore sarebbe rappresentata dalla libreria FFTW, punto di riferimento in quanto a velocità di esecuzione, tuttavia la volontà di realizzare un programma il più possibile leggero e multi piattaforma e la necessità di processare solo dati reali (immagini) ha fatto propendere per la libreria kissfft, anch'essa distribuita con licenza libera. Questa è infatti molto ridotta in dimensioni, formata solo da codice C++ da aggiungere al progetto senza richiedere installazioni aggiuntive e per quanto riguarda la FFT reale, che è la sola utilizzata per la registrazione, fornisce prestazioni di gran lunga migliori rispetto a CImg senza imporre i limiti alla forma dell'input, e molto vicine a quelle ottenute con la FFTW. In tale libreria l'aumento del tempo di esecuzione per dimensioni maggiori non è strettamente monotono, infatti tra dimensioni molto simili viene ad essere significativa la natura numerica della dimensione laterale dell'immagine: per la struttura del calcolo dell'FFT i numeri completamen-

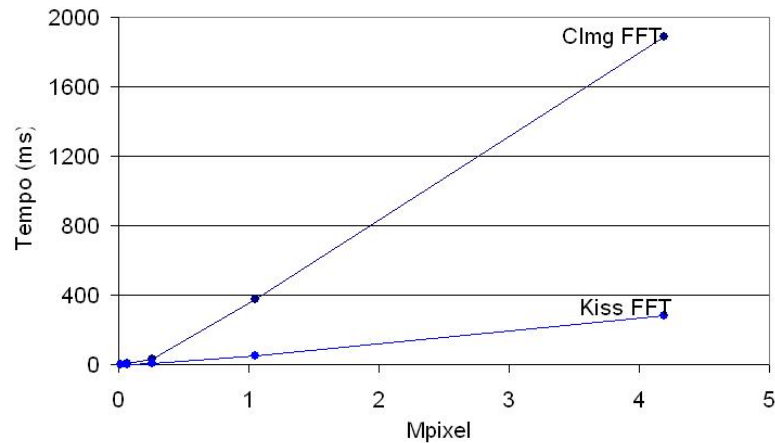


Figura 3.1: Confronto tra tempi di esecuzione di CImg FFT e KissFFT

te fattorizzabili per due, tre e cinque sono più veloci da processare dei valori che non riportano tale caratteristica, talvolta con differenze di due ordini di grandezza. Quindi per velocizzare l'operazione si pospone al passaggio di zeropadding un controllo della dimensione dell'immagine che viene aumentata del minimo necessario per soddisfare i suddetti requisiti. In questo modo pur aumentando la dimensione dell'input si ottengono prestazioni migliori, e l'aumento risulta comunque più contenuto vista la maggiore densità dei multipli ricercati rispetto ai quadrati di due.

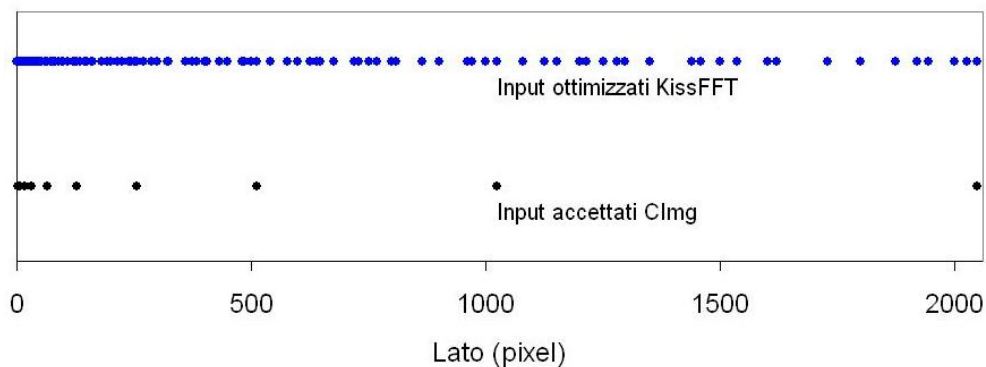


Figura 3.2: Confronto tra le densità di input utilizzabili con CImg e KissFFT

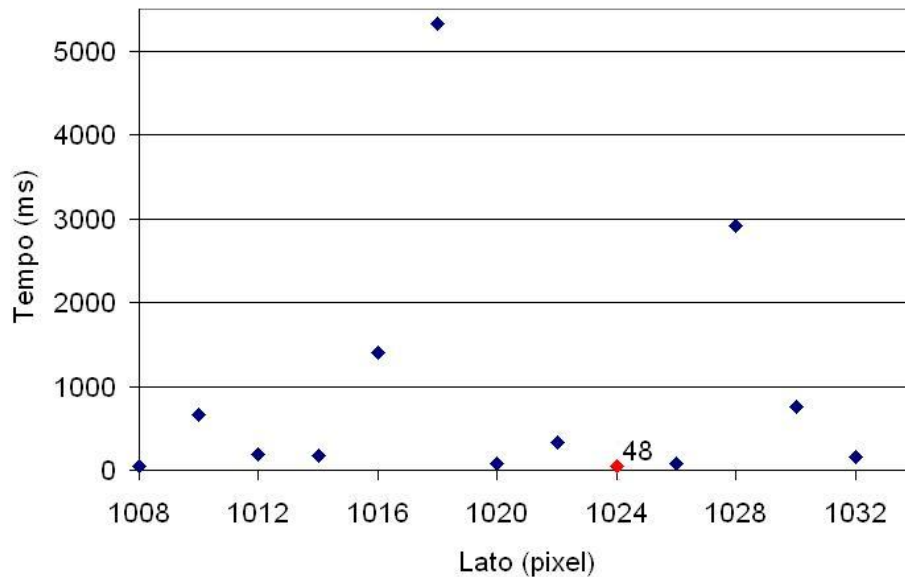


Figura 3.3: Tempi di esecuzione per immagini con dimensione laterale crescente per intervalli minimi

3.2 Organizzazione del progetto

La realizzazione del progetto è stata suddivisa in classi distinte che si occupano dello svolgimento delle varie parti del processo di registrazione, coordinate da una procedura principale. La conversione in scala di grigi dell'immagine e la successiva localizzazione spaziale del nevo sono procedimenti cui fanno capo due classi distinte. Una terza classe fornisce una serie di strumenti per l'elaborazione di immagini utilizzati nelle fasi di ottimizzazione del contrasto automatica e localizzata, sogliatura e annullamento della media. L'esecuzione in logica multiscalata della registrazione è delegata ad un'apposita classe che studia separatamente il livello minimo, quelli medi che sono in numero variabile a seconda della risoluzione dell'immagine ed il primo, utilizzando un buffer di memoria per salvare gli angoli con i risultati migliori da processare nelle fasi successive. Questa si occupa di tutta la fase che prevede l'impiego della correlazione incrociata. Un'ulteriore classe raccoglie le variabili globali del programma quali i valori di soglia e saturazione, la risoluzione minima del livello al quale iniziare la registrazione, il numero di angoli da tenere in considerazione ad ogni passaggio, i fattori di peso delle aree usati nella ricerca del nevo, il fattore di sicurezza per il raggio

trovato; questi non sono infatti valori determinati esclusivamente a priori ma vengono affinati empiricamente, ed in tal modo li si rende facilmente accessibili e modificabili per migliorare le prestazioni. Una classe principale si occupa di coordinare tutti i passaggi necessari, gestendo le immagini in input, eseguendo in ordine il pre-processing sulle foto, la registrazione per livelli successivi e fornendo in output il risultato nella forma desiderata (immagine registrata o i soli valori di rotazione e traslazione).

Per poter realizzare un'unica libreria che racchiuda l'intero codice della registrazione è stata creata anche una classe di interfaccia provvista di metodi con chiamata JNI di modo da permetterne l'utilizzo in ambiente java; tali metodi accettano in ingresso le due immagini codificate in forma vettoriale e forniscono come uscita l'immagine registrata oppure i soli valori di traslazione e rotazione, a seconda delle richieste dell'utente. I metodi di utilizzo per l'utente prevedono la possibilità di portare a termine la registrazione fino al risultato completo oppure di arrestarla ad un livello più basso considerato di precisione accettabile per ridurre il tempo di esecuzione, con la possibilità di completare poi in seguito il procedimento.

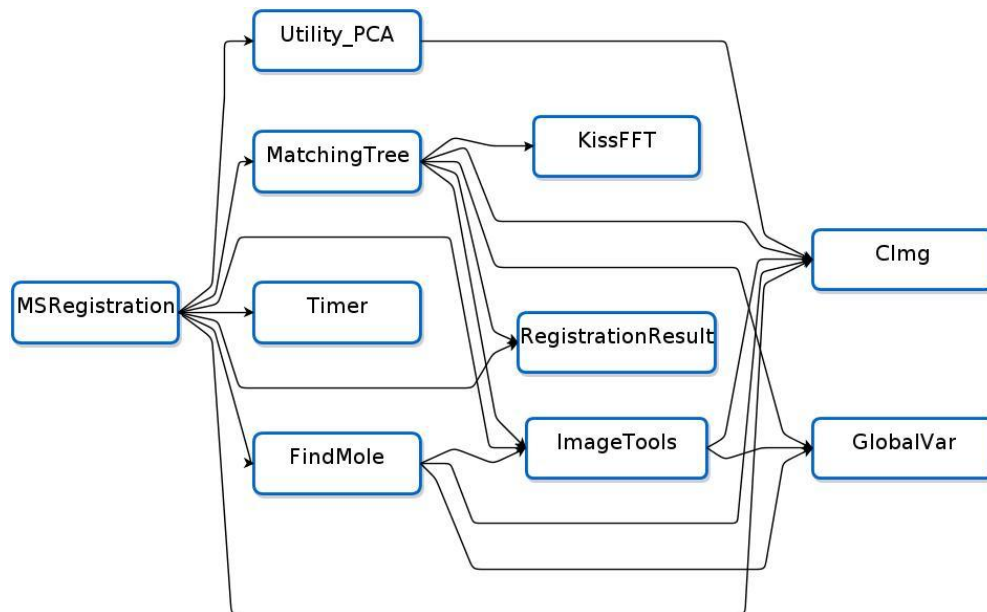


Figura 3.4: Organizzazione gerarchica delle classi del progetto

Prestazioni

4.1 Robustezza e precisione

Il metodo adottato dell'analisi in frequenza, grazie alle sue caratteristiche di bassa sensibilità al rumore si è rivelato molto robusto, registrando correttamente tutte le coppie di immagini sottoposte alla prova. L'algoritmo ha prodotto risultati affidabili sia con coppie di immagini di nevi con forme quasi invariate, sia con nevi che mostrano un'evoluzione avvenuta nel tempo e quindi aumento delle dimensioni e cambiamenti nella forma.

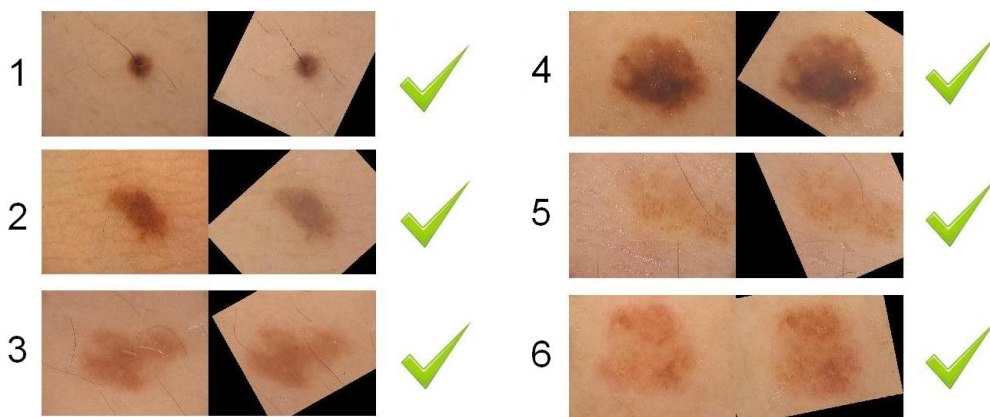


Figura 4.1: Esempi di registrazioni di coppie di immagini con differenze di forma minime



Figura 4.2: Esempi di registrazioni di coppie di immagini con differenze marcate dovute all'evoluzione

Data la natura del confronto basato sull'intensità, nelle foto dai colori poco accentuati è stato fondamentale aver aggiunto i passaggi preliminari di elaborazione dell'immagine per accentuare il contrasto senza i quali la localizzazione del nevo o la registrazione non sarebbero sempre soddisfacenti. Altro accorgimento rivelatosi necessario per incrementare la robustezza del

processo è stato prendere in considerazione ad ogni passaggio della cross-correlazione una selezione dei migliori risultati invece del solo primo, che all'analisi a risoluzioni più elevate avrebbe potuto rivelarsi sbagliato. La precisione è strettamente legata alla risoluzione cui viene analizzata l'immagine. Per quanto riguarda l'angolo la scelta di muoversi per intervalli rappresentati dalla minima rotazione ottenibile assicura l'ispezione di tutte le possibilità, la precisione è quindi dovuta alla dimensione utilizzata nel livello più alto cui si arriva nell'affinamento della ricerca. L'accuratezza della traslazione dipende invece solo dalla dimensione dell'immagine utilizzata per l'ultimo passaggio, effettuato dopo aver individuato l'angolo di rotazione.

4.2 Tempi di esecuzione

In termini temporali gran parte del contributo è dovuto al calcolo della FFT nella correlazione incrociata, mentre le fasi di preparazione e pre-processing ricoprono un ruolo marginale. Come si vede dai risultati la sola operazione dispendiosa nella pre-elaborazione delle immagini è la conversione in scala di grigi a causa dell'utilizzo dell'analisi delle componenti principali. Sono invece la ricerca della rotazione nei livelli intermedi e soprattutto della traslazione nel primo a portare il maggior dispendio temporale.

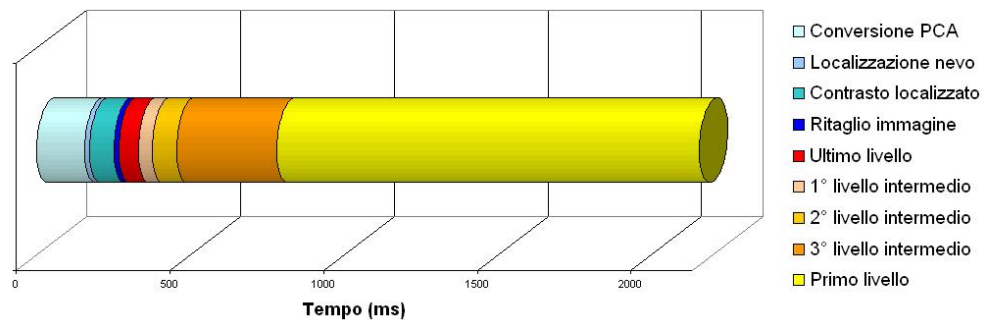


Figura 4.3: Distribuzione dei tempi della registrazione completa

Il tempo nelle varie prove è risultato variare anche in funzione della dimensione del nevo poiché questa va a modificare il tempo di calcolo della sua FFT, minore per aree minori. Per velocizzare l'algoritmo sono già stati adottati accorgimenti quali l'uso della FFT reale, più performante della complessa, e la riduzione del numero di angoli ottenuti ad ogni passaggio allo stretto necessario per garantirne la correttezza. Nonostante ciò la re-

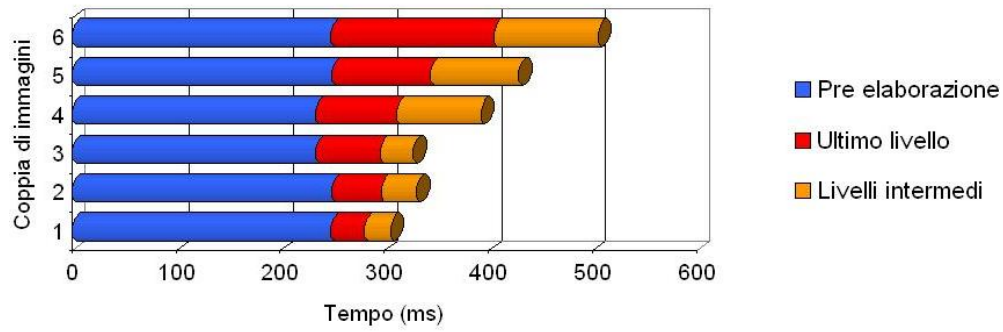
gistrazione con ricerca della massima precisione resta insoddisfacente per l'esecuzione in tempi molto rapidi o l'utilizzo su sistemi mobile con ridotta capacità di calcolo. Per permettere un uso personalizzato della libreria è quindi previsto un metodo di interfaccia con possibilità di selezionare il livello di precisione desiderato il quale va ad influenzare direttamente il tempo di esecuzione. Sta così all'utente la possibilità di scegliere il giusto compromesso tra rapidità e precisione, con la possibilità di avere un risultato parziale in tempi rapidi ed in seguito completare la registrazione.

4.3 Trades off

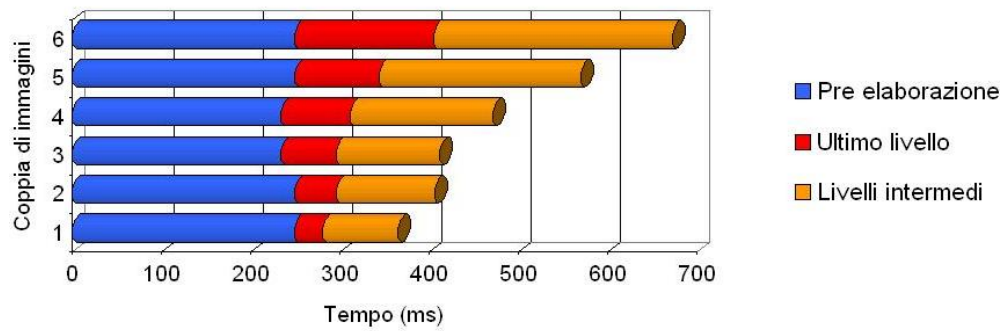
Per venire incontro alla richiesta di brevi tempi esecutivi si deve perciò intervenire riducendo la dimensione delle immagini utilizzate nelle fasi di registrazione, senza compromettere la correttezza della registrazione ma intaccandone la precisione.

Livelli intermedi

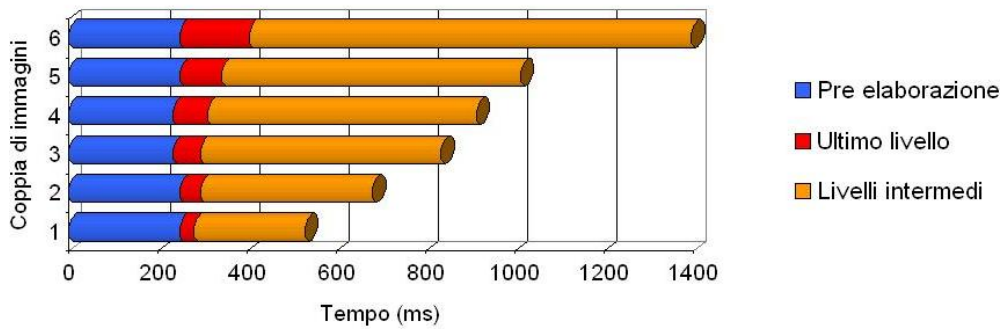
In un primo caso si può arrestare la ricerca dell'angolo ad una risoluzione più bassa senza compiere tutti i livelli ed avere così un risultato meno preciso ma comunque corretto. Eliminando uno o due passaggi i tempi totali della fase intermedia diminuiscono notevolmente e gli errori restano confinati entro i quattro gradi angolari con l'eliminazione di due livelli, entro i due gradi con l'eliminazione di un livello. Tuttavia ciò non ha grande effetto nell'esecuzione globale del programma dato che riveste comunque una porzione marginale del tempo. La registrazione in seguito può essere portata a termine proseguendo dai risultati parziali già ottenuti



(a) Esecuzione di un livello intermedio



(b) Esecuzione di due livelli intermedi



(c) Esecuzione di tre livelli intermedi

Figura 4.4: Tempi di esecuzione al variare della risoluzione raggiunta nei livelli intermedi

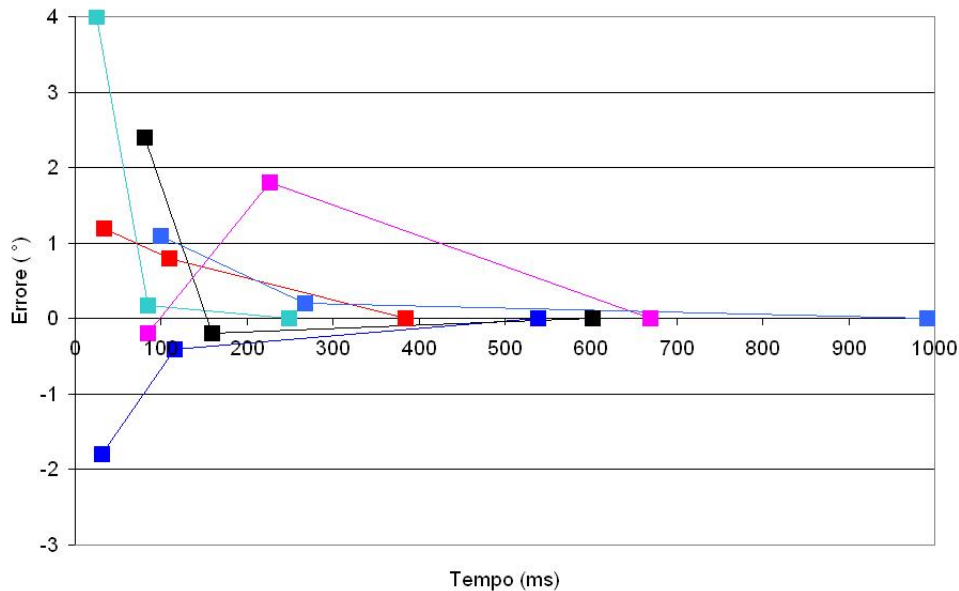
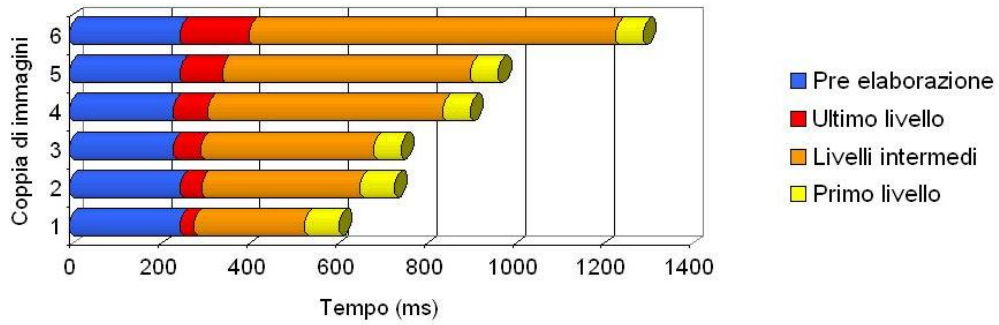


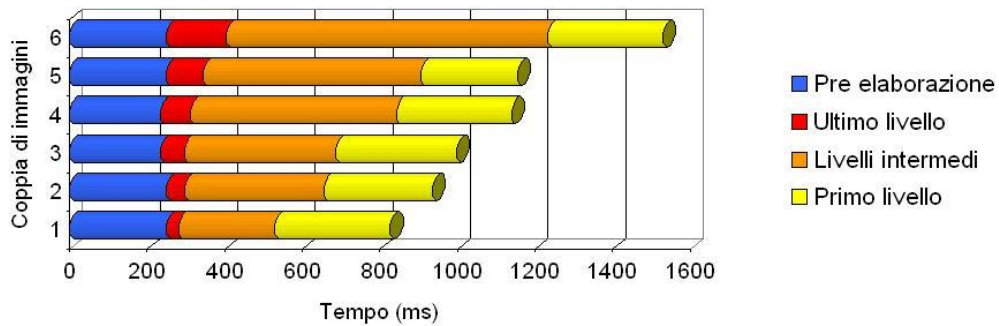
Figura 4.5: Errori angolari in funzione del tempo di esecuzione

Primo livello

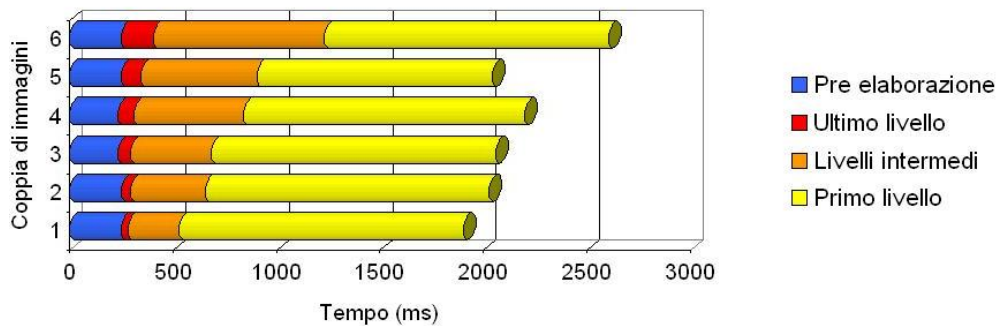
Ben più importante nel totale dei tempi è invece la correlazione incrociata eseguita sul primo livello per cercare la traslazione con l'ultima correlazione incrociata; se eseguita a risoluzione piena essa viene infatti ad impiegare oltre metà del tempo totale del programma. Anche in questo caso la si esegue a risoluzioni ridotte per cercare di ottenere i risultati più velocemente per quanto più imprecisi. Riducendo a metà la dimensione si ottengono tempi ridotti di un fattore circa due con un sacrificio minimo nell'accuratezza della traslazione che risulta differente di non più di tre punti percentuali. Ricercando prestazioni ancora migliori si esegue il passaggio finale con la foto a risoluzione ridotta ad un quarto ottenendo tempi sensibilmente minori ed errori comunque compresi entro il 5%.



(a) Risoluzione un quarto



(b) Risoluzione dimezzata



(c) Risoluzione piena

Figura 4.6: Tempi di esecuzione al variare della risoluzione del primo livello

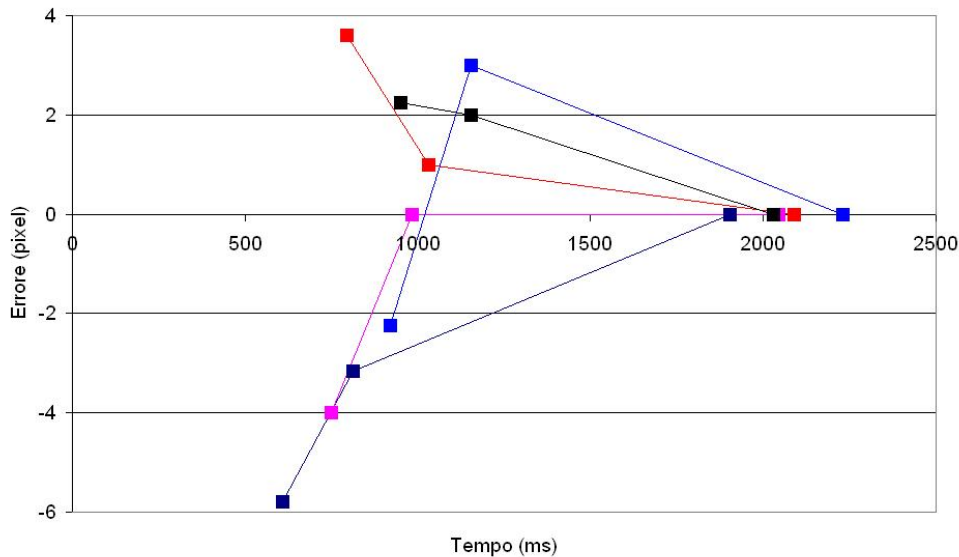


Figura 4.7: Errori di traslazione in funzione del tempo di esecuzione

4.4 Conclusioni

La libreria creata risulta un valido strumento da integrare in sistemi informatici di monitoraggio nevi. Risulta infatti affidabile, personalizzabile per quanto riguarda utilizzo e interfacciamento, compatibile con programmi C++ e Java ed utilizzabile in ambienti Windows, Unix e Android. Per permettere un utilizzo sostenibile su sistemi mobile sarà necessario ridurre il tempo di esecuzione. Un tentativo effettuato durante le fasi di sviluppo per ridurre il numero di FFT calcolate nella registrazione è stato concentrare lo studio nel dominio delle frequenze, calcolando inizialmente una sola trasformata dell'immagine e ricavando le FFT delle successive rotazioni con operazioni su quella nota invece che direttamente dalle immagini; tuttavia tale metodo si è rivelato simile per quanto riguarda i tempi di calcolo, e molto più soggetto a errori a causa delle imprecisioni introdotte nei calcoli sulle trasformate.

Resta quindi aperta la ricerca di algoritmi di registrazione più performanti basati su nuove e diverse tecniche realizzative.