

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Corso di Laurea in Matematica

**Spazi di configurazioni e trasformata
generalizzata di Hough**

Tesi di Laurea in Geometria

Relatore:
Chiar.mo Prof.
Massimo Ferri

Presentata da:
Giulia Pasqualini

Sessione II
Anno Accademico 2015-2016

Alle mie famiglie.

Indice

Introduzione	2
1 Introduzione alla visione artificiale	5
1.1 Concetto generale di visione	5
1.2 Problema generale relativo alla visione artificiale	6
2 Preliminari Matematici	9
2.1 Spazi Proiettivi	9
2.1.1 Dualità Proiettiva	9
2.2 Spazi di configurazioni	12
2.2.1 Coordinate generalizzate	12
2.3 Spazi di moduli	13
2.4 Trasformata di Hough	14
2.4.1 Trasformata semplice di Hough	15
2.4.2 Trasformata generalizzata di Hough	16
2.5 Confronto tra i metodi	19
3 Applicazioni del metodo della trasformata di Hough	21
3.1 Applicazione alla geologia: allineamento dei vulcani	21
3.1.1 Introduzione al problema generale	21
3.1.2 Applicazione della trasformata di Hough	23
3.2 Applicazione alla medicina: identificazione dell'aorta	27
3.2.1 Introduzione al problema generale	27
3.2.2 Applicazione della trasformata di Hough	27
3.3 Applicazione alla medicina: analisi delle vescicole di membrana	31
3.3.1 Introduzione al problema generale	31
3.3.2 Applicazione della trasformata di Hough	31
4 Conclusione	34
Bibliografia	35

Introduzione

L'intelligenza artificiale (I.A.), nata circa nel 1950, è l'abilità di un computer di svolgere funzioni e ragionamenti tipici della mente umana; secondo le parole di Marvin Minsky, uno dei "pionieri" della I.A., lo scopo di questa nuova disciplina sarebbe quello di *"far fare alle macchine delle cose che richiederebbero l'intelligenza se fossero fatte dagli uomini"*.

Molteplici sono i campi di ricerca a cui si rivolge questa disciplina, ma uno dei maggiori ostacoli che essa incontra riguarda la visione artificiale. Quest'ultima, infatti, non presuppone solamente un insieme di compiti specifici puramente meccanici per riprodurre un'immagine (compiti che un computer sarebbe capace di compiere molto meglio di un essere umano) ma racchiude in sé anche l'interpretazione e la successiva ri-rappresentazione della stessa. Da circa sessant'anni, i più grandi studiosi di questa disciplina hanno cercato di trovare metodi e modelli fisico-matematici, che potessero rispondere a tutte le necessità poste dal problema della visione. Ciò che ne risulta, però, è un insieme di metodi creati specificatamente per ogni singolo problema, che cozzano col presupposto iniziale di ricerca di un metodo universale che possa essere flessibile quanto il sistema di visione umano.

Questo elaborato si propone di descrivere una delle tecniche maggiormente usate dalla visione artificiale per la rilevazione di forme specifiche in un'immagine, attraverso il confronto con diversi metodi ad essa simili: la trasformata di Hough. Quest'ultima è un modello matematico dal panorama assai ampio, dal momento che utilizza strumenti propri della geometria, della fisica-matematica e della geometria algebrica. Inoltre, ne esistono diversi tipi: dalla semplice, preposta alla ricerca di rette, alla generalizzata per la ricerca di forme circolari, fino alla ricerca di forme arbitrarie.

Ciò su cui ci concentreremo sarà la ricerca di rette e cerchi in un'immagine, per poi approfondire entrambi i modelli, enfatizzandone l'importanza, con l'enunciazione di tre applicazioni recenti, riguardanti campi di studio indispensabili e sempre attuali come la medicina o la geologia.

In particolare vedremo che, seppur il problema della visione è ben lontano dall'essere risolto, alcuni quesiti sono stati ampiamente portati a termine

grazie all'uso di tecniche come quella esaminata in questo elaborato.

Più nello specifico, la tesi si compone di tre capitoli:

Primo capitolo: presenta una panoramica generale sul concetto di visione e successivamente dà un'idea dei problemi che devono affrontare i ricercatori nel campo della visione artificiale.

Secondo capitolo: riguarda i preliminari matematici, utilizzati per comprendere il concetto che sostiene la trasformata di Hough. In particolare presenta quindi il funzionamento di quest'ultima, sia nel caso semplice che generalizzato, per poi passare ad un confronto con i preliminari suddetti.

Terzo capitolo: è diviso in tre paragrafi. Il primo descrive un'applicazione della trasformata semplice di Hough alla geologia per lo studio della distribuzione dei vulcani. Il secondo e il terzo presentano due applicazioni della trasformata generalizzata di Hough alla medicina, rispettivamente per l'identificazione dell'aorta e per l'analisi automatizzata delle vescicole di membrana.

Capitolo 1

Introduzione alla visione artificiale

1.1 Concetto generale di visione

La visione è la modalità sensoriale dominante nell'uomo. Essa ci fornisce, infatti, una notevole quantità di informazioni riguardo ciò che ci circonda e ci rende capaci di interagire intelligentemente con l'ambiente, il tutto senza il contatto fisico diretto.

Ma cosa significa vedere? In prima istanza la visione si può schematizzare come un processo di informazione attraverso il quale si scopre cosa c'è nel mondo e dove si trova, quindi le posizioni e l'identità degli oggetti e le relazioni che intercorrono tra di essi.[1, 2]

Questo processo è il risultato di ciò che viene denominato ottica, la quale si occupa di studiare la fisiologia della visione umana in modo puramente tecnico, sottolineandone quindi i passaggi fisici che la caratterizzano. Ma attribuire alla visione il mero significato di ottica, significherebbe banalizzare il problema. Se siamo capaci di sapere cosa c'è nel mondo e dove si trova, significa che il nostro cervello è capace, in qualche modo, di rappresentare questa informazione. La visione, infatti, non mira alla semplice acquisizione di informazioni dal mondo esterno attraverso l'immagine, ma coinvolge anche il successivo processo di riconoscimento e rielaborazione dell'oggetto rappresentato, ed è ciò che permette all'uomo di poter ri-rappresentare ciò che ha già interpretato. [3]

La visione è, perciò, qualcosa di più di una capacità sensoriale, è un'intelligenza, ed è proprio questa dualità (informazione-elaborazione) che dà forma al problema della visione.

Se ci si concentra sul funzionamento del cervello umano, si è più inclini a lo-

dare i poteri deduttivi di un logico, che la capacità di una persona qualunque di riconoscere un viso o un oggetto. [4] Comunemente infatti, dal momento che noi vediamo senza alcuno sforzo e in modo del tutto naturale, siamo portati a pensare che il processo che permette di vedere sia relativamente semplice, tanto da non costituire nemmeno un vero e proprio problema.

Ma con l'avvento della tecnologia e in particolar modo dei computer, nel momento in cui si è cercato di fornire a quest'ultimo la capacità di visione e non di semplice ottica, ci si è accorti della grande complessità di tale compito.

Se si considera il computer come un sistema di elaborazione di informazioni, proprio come il cervello umano, dato che ha memoria e capacità di calcolo a sufficienza, si potrebbe, infatti, supporre di applicargli una telecamera e scrivere un algoritmo che fornisca in output una descrizione di ciò che la telecamera vede.

Ma ciò si è rivelato un problema di elevata complessità, dal momento che il computer ha risultati di gran lunga migliori rispetto ad un essere umano quando si richiede un'ampia capacità di memoria; tuttavia qualsiasi compito che presupponga una elaborazione e quindi un'interpretazione, risulta quasi impossibile per esso. [3]

I progressi della visione artificiale negli ultimi cinquant'anni hanno rivelato un dato che suona ironico: la visione è non solo intelligente ma è addirittura più difficile da capire o da riprodurre che il più sofisticato dei ragionamenti matematici. Basti pensare che l'intelligenza artificiale oggi è molto più avanti in ricerche e sviluppi nella creazione di sistemi che possano svolgere alcuni compiti di medici e avvocati, piuttosto che mansioni apparentemente meno complesse, come giardinieri o cuochi.

Questo è anche dovuto al fatto che il "potere" della visione umana è proprio del cervello, il quale è un organo complesso e ancora misterioso, il cui funzionamento anche dal punto di vista della visione è sconosciuto. Proprio lì sta la difficoltà, nel dover dare una capacità al computer, di cui non sappiamo ancora nulla. [4]

1.2 Problema generale relativo alla visione artificiale

Lo scopo della visione robotica è quello di costruire macchine che analizzino immagini e producano descrizioni di ciò che è l'immagine. Queste descrizioni dovrebbero catturare gli aspetti degli oggetti che poi verranno ripresi perchè utili per svolgere dei compiti.

La visione può essere considerata come un elemento di un ciclo di feedback

che coinvolge sensi e altri elementi che sono preposti alle decisioni e le implementazioni di queste decisioni. L'input delle macchine è un'immagine, o più di una, mentre il suo output è una descrizione che deve soddisfare due criteri:

- Deve recare qualche relazione riguardo ciò che viene ripreso
- Deve contenere tutte le informazioni di cui abbiamo bisogno per certi compiti

Per capire l'importanza e la forza dell'approccio della visione artificiale, occorre essere a conoscenza delle difficoltà e dei limiti che essa affronta. Un aneddoto divertente che ci permette di capire il contesto nel quale ha origine la visione artificiale, appartiene alla biografia di Marvin Minsky, uno dei padri fondatori dell'intelligenza artificiale.

Lui, circa cinquant'anni fa, aveva assegnato ad un suo studente un compito: collegare una telecamera ad un computer e far in modo che quest'ultimo descrivesse ciò che la telecamera riprendeva. Tale progetto si è esteso tanto da diventare una vera e propria materia di ricerca, e nonostante i molti progressi fatti, il problema della visione artificiale non è stato ancora risolto. [4]

L'approccio globale a tale problema consiste dapprima nello sviluppare teorie di tipo fisico-matematico che descrivano alcune parti del processo della visione, poi nel trasformare tali teorie in algoritmi che possano funzionare in situazioni reali. [5]

Inizialmente, i primi scienziati che si occupavano di visione tentavano di risolvere il problema con ogni mezzo. A fronte di un singolo compito, essi utilizzavano qualsiasi strategia che potesse togliere ambiguità all'immagine in modo da renderla ri-rappresentabile dalla macchina, senza tener conto di quanto fossero limitate e circoscritte le informazioni che occorrevano al fine di risolvere il suddetto compito.

Eppure, la visione artificiale ha fatto grandi progressi da quando lo studente di Minsky ha dovuto risolvere quel problema. Ora infatti, i metodi prevedono regole ad hoc studiate per catturare specifiche caratteristiche presenti nelle immagini.

Questi metodi si basano su una suddivisione in parti del problema, i quali sono trattabili indipendentemente tra loro. Il processo complessivo di lavoro può essere suddiviso, infatti, in moduli concettualmente indipendenti come ad esempio la visione stereoscopica, la determinazione della forma degli oggetti per mezzo delle ombre, il rilevamento dei profili e così via ed ha condotto alla messa appunto di metodi matematici (come le trasformate di Hough) e non preposti alla risoluzione di questi singoli problemi.

Questo approccio è quello che viene utilizzato tutt'ora e, per la maggior parte

dei casi, ha portato alla costruzione di macchine per la visione con risultati apprezzabili, ma che presentano notevoli limiti.

Infatti i pionieri di questo cambiamento, si sono resi conto che: *senza una teoria sulla descrizione dell'immagine abbastanza generale da guidare l'interpretazione di un'immagine qualsiasi, la visione artificiale sarebbe condannata ad un'infinita ripetizione di espedienti, dove ciascuno è concepito più intelligentemente del precedente, ma nessuno di essi sarebbe capace di raggiungere la flessibilità del sistema visivo umano.* [4]

Quindi, il limite principale della visione artificiale è la mancanza di un metodo universale che permetta alla macchina di dare una descrizione completa di un'immagine qualsiasi; inoltre i metodi fino ad ora trovati sono perfetti per la risoluzione di problemi specifici, ma non hanno la flessibilità giusta per portare ad un progresso in direzione del suddetto metodo universale.

Capitolo 2

Preliminari Matematici

2.1 Spazi Proiettivi

Definizione 2.1 (Spazio Proiettivo). *Sia K un campo, sia V un K -spazio vettoriale di dimensione finita. Lo spazio proiettivo associato a V è l'insieme $P(V)$ i cui elementi, chiamati punti di $P(V)$, sono i sottospazi vettoriali $L \subset V$ uno-dimensionali.*

La dimensione di $P(V)$ è definita come $\dim(V) - 1$ e si denota con $\dim(P(V))$.

Convenzionalmente un punto di $P(V)$ è denotato con il simbolo $[v]$, dove $v \in V \setminus \{0\}$ è un generatore del sottospazio uno-dimensionale di V :

$$\langle v \rangle = \{uv : u \in K\}$$

Es: $V = K^n + 1$, allora $P(V)$ si indica con $P^n(K)$ ed è uno spazio n -dimensionale chiamato n -esimo spazio proiettivo numerico.

Definizione 2.2 (Sottospazio Proiettivo). *Sia V un K -spazio vettoriale di dim n e $P(V)$ lo spazio proiettivo associato, si definisce sottospazio proiettivo, o sottovarietà proiettiva, di $P(V)$ di dim k il proiettivo di un sottospazio vettoriale $W \subset V$ di dim $k+1$.*

2.1.1 Dualità Proiettiva

Definizione 2.3 (Spazio Proiettivo Duale). *Sia V un K -spazio vettoriale di dimensione finita. Lo spazio proiettivo $P^* = P(V^*)$, dove $V^* = \text{Hom}(V, K)$ si dice spazio proiettivo duale di $P = P(V)$.*

P e P^* hanno la stessa dimensione poiché $\dim(V) = \dim(V^*)$.

Per definizione, due funzionali lineari $F, F' : V \rightarrow K$, entrambi non nulli, definiscono lo stesso elemento di P^* , in simboli $[F] = [F'] \iff F' = \lambda F$ per qualche $\lambda \neq 0$.

In tal caso $\ker(F) = \ker(F')$, quindi l'iperpiano $\ker(F)$ di V dipende solamente dal punto $[F] \in P^*$. Definiamo quindi:

Definizione 2.4 (Applicazione di dualità). $\delta : P^* \rightarrow \{ \text{iperpiani di } P \}$, è detta applicazione di dualità ed ha questa forma:

$$\delta([F]) = P(\ker(F))$$

Definizione 2.5 (Sistemi lineari di iperpiani). Sia S un sottospazio di P di dimensione $k \leq n - 1$. L'insieme $\Lambda(S)$ i cui elementi sono gli iperpiani di P che contengono S si dice sistema lineare di iperpiani di centro S .

Proposizione 2.1. Supponiamo che il sottospazio S di P abbia dimensione k , ed equazioni cartesiane:

$$\begin{cases} F_1(X_0, \dots, X_n) = 0 \\ \vdots \\ F_{n-k}(X_0, \dots, X_n) = 0 \end{cases} \quad (2.1)$$

allora il sistema lineare $\Lambda(S)$ consiste degli iperpiani di equazione

$$\lambda_1 F_{n-k}(X_0, \dots, X_n) + \dots + \lambda_{n-k} F_{n-k}(X_0, \dots, X_n) = 0 \quad (2.2)$$

dove $\lambda_1 \cdots \lambda_{n-k} \in K$ sono scalari non tutti uguali a zero.

Dimostrazione. Ogni iperpiano della forma (2.2) appartiene a $\Lambda(S)$, cioè contiene S , perchè le coordinate di ogni punto di S annullano i polinomi F_1, \dots, F_{n-k} e quindi annullano il primo membro della (2.2). Viceversa, sia H un iperpiano contenente S , di equazione:

$$a_0 X_0 + \dots + a_n X_n = 0 \quad (2.3)$$

Il sistema (2.1) e quello costituito da (2.1) e (2.3) sono equivalenti. Quindi (2.3) è combinazione lineare di (2.1), cioè è della forma (2.2), come si voleva. \square

A partire da queste definizioni e dalla proposizione enunciamo il teorema, detto teorema di dualità, del quale daremo due formulazioni:

Teorema 2.2. *Sia $\dim(P) = n$, l'applicazione di dualità (2.4) è una biiezione che fa corrispondere ad ogni sottospazio proiettivo di dimensione $n-k-1$ di P^* un sistema lineare di iperpiani di P di centro un sottospazio proiettivo di dimensione k . Si ottiene in questo modo una corrispondenza biunivoca tra sottospazi di P^* e sottospazi di P che rovescia le inclusioni.*

Dimostrazione. Per prima cosa dimostriamo che l'applicazione δ è biunivoca, quindi che è iniettiva e suriettiva.

Iniettività: Siano F e F' due funzionali lineari non nulli, e supponiamo che $\ker(F) = \ker(F')$, allora F e F' sono proporzionali, quindi definiscono lo stesso punto di P^* .

Suriettività: Ogni iperpiano di V è un nucleo di un funzionale lineare, quindi è anche suriettiva.

Ora, sia $k = \dim(S)$, allora $n-k-1$ è la dimensione del sistema lineare $\Lambda(S)$ (segue dalla proposizione). Da qui, dato che la mappa è biunivoca, il teorema è verificato. □

Essendo univocamente definita da P , δ identifica in modo intrinseco, cioè dipendente solo da P , l'insieme degli iperpiani di P con lo spazio proiettivo P^* . In particolar modo δ permette di considerare gli iperpiani di P come elementi di uno spazio proiettivo.

Teorema 2.3. *Sia P una proposizione che riguarda i sottospazi di uno spazio proiettivo n -dimensionale $P(V)$, le loro intersezioni, i sottospazi da essi generati e le loro dimensioni. Allora P è vera se e solo se è vera la proposizione “duale” P^* ottenuta sostituendo i termini “intersezione, sottospazio generato, contenuto, contenente, dimensione”, rispettivamente con “sottospazio generato, intersezione, contenente, contenuto, dimensione duale”, dove quest'ultima vale $n-k-1$ se la dimensione vale k .*

Dimostrazione. Segue dal teorema 2.2 □

Esempio: Supponiamo di avere $P^3(K)$, per tre punti non collineari in tale spazio proiettivo passa uno ed un solo piano. Nel duale, i punti sono gli iperpiani, in questo caso saranno tre piani, la cui intersezione è un punto. Alla fine un piano in $P^3(K)$ sarà un punto nel duale.

In questo paragrafo ho utilizzato [6, 7].

2.2 Spazi di configurazioni

I corpi estesi, intesi come sistema meccanico, possono essere pensati come un insieme di particelle puntiformi, dotate di massa ma prive di conformazione interna, legate da una specifica relazione spaziale: descrivere la configurazione del sistema significa quindi specificare la posizione di ogni particella costituente.

L'insieme di tutte le configurazioni del sistema è ciò che viene chiamato spazio delle configurazioni del sistema e la sua dimensione è il più piccolo numero di parametri che servono per descrivere completamente la configurazione del sistema.

La dimensione dello spazio delle configurazioni viene anche chiamato il numero dei gradi di libertà del sistema. Ad esempio, per specificare la configurazione di una singola particella puntiforme non vincolata occorrono tre parametri, quindi il suo spazio delle configurazioni è tre-dimensionale.

Nel caso più generale di un sistema composto da k particelle separate occorrono $3k$ parametri per descriverne le possibili configurazioni.

Se ci sono dei vincoli tra le particelle del sistema, la configurazione è ridotta ad uno spazio di dimensione più piccolo: nel caso di un corpo rigido, nel quale vale quindi il vincolo di rigidità (le posizioni reciproche dei punti rimangono costanti nel tempo), si hanno sei gradi di libertà: il minimo numero di parametri per specificare la posizione nello spazio è tre e il minimo numero di parametri per specificare l'orientazione è tre.

2.2.1 Coordinate generalizzate

Per poter parlare di configurazione del sistema abbiamo bisogno di un insieme di parametri che lo descrivano: le coordinate generalizzate. Nel caso di una particella libera, la sua configurazione è specificata dandone la posizione e ciò, abbiamo visto, richiede tre parametri in quanto ha tre gradi di libertà. Un esempio di coordinate generalizzate per questo sistema sono le coordinate cartesiane relative ad un sistema di assi fissato. In generale, le configurazioni formano uno spazio M di una certa dimensione n , il quale può essere parametrizzato scegliendo una funzione coordinante A che fa corrispondere elementi di spazi di configurazioni (m) a n -uple di numeri reali, $A_i(m)$, i quali rappresentano le coordinate generalizzate.

Queste coordinate generalizzate ci permettono di identificare i punti dello spazio delle configurazioni n -dimensionale con n -uple di numeri reali. C'è una grande varietà di modi per scegliere le coordinate generalizzate, ad esempio per un singolo punto che si muove senza vincolo, possiamo scegliere le coordinate cartesiane, le coordinate polari o altri sistemi di coordinate.

In questo paragrafo ho utilizzato [8].

2.3 Spazi di moduli

Dal punto di vista matematico (come abbiamo visto con lo spazio proiettivo duale e con gli spazi di configurazioni), è sempre possibile prendere in considerazione uno spazio astratto, in cui un punto rappresenta un oggetto con una struttura geometrica ben precisa (semiretta, iperpiano ecc) di un altro spazio. In geometria algebrica tale oggetto matematico è ciò che viene chiamato spazio dei moduli, o spazio dei parametri. Lo spazio dei moduli è, infatti, uno spazio geometrico i cui punti rappresentano oggetti algebrico-geometrici di qualche tipo fissato (per esempio rette) o classi di isomorfismo di tali oggetti. In questo contesto il termine “modulo” è usato come sinonimo di “parametro” e gli spazi di moduli sono intesi come spazi di parametri e non come spazi di oggetti.

Tali spazi si presentano come una soluzione ai problemi di classificazione: se è possibile dimostrare che un insieme di oggetti può avere una struttura di spazio geometrico (spazi di moduli), allora si parametrizzano tali oggetti introducendo le coordinate. Per esempio si consideri il problema di trovare tutti i cerchi nel piano euclideo, a meno di congruenza.

Qualsiasi cerchio può essere descritto dando tre punti, ma molti insiemi differenti di tre punti possono identificare lo stesso cerchio: la corrispondenza è tanti-uno. Tuttavia, i cerchi sono univocamente determinati dando il loro centro e il raggio: due parametri reali e uno reale positivo e dal momento che siamo interessati solo ai cerchi a meno di congruenza, identifichiamo i cerchi aventi centri diversi ma stesso raggio il quale da solo è sufficiente per parametrizzare l'insieme di interesse.

In questo caso lo spazio dei moduli è quindi la semiretta dei numeri reali positivi. Un altro problema risolvibile attraverso lo spazio dei moduli è quello di descrivere tutte le rette passanti per l'origine nel piano reale. Questo problema di classificazione è facilmente risolvibile assegnando ad ogni retta r un parametro, o modulo, che ci aiuterà a distinguerle.

Tutto ciò che dobbiamo fare è di fissare un sistema di riferimento cartesiano x e y nel piano e misurare l'angolo $\theta(L)$ tra la retta L e l'asse x , preso in senso antiorario. Troveremo che i possibili valori che può assumere θ sono tra 0 e π , e per ogni θ così trovato, c'è esattamente una retta L : l'insieme di tutte le rette passanti per l'origine, conosciuto come la retta proiettiva, è in corrispondenza biunivoca con l'intervallo $[0, \pi]$.

In questo paragrafo ho utilizzato [9, 10].

2.4 Trasformata di Hough

La forma è una proprietà importante nelle figure bidimensionali e nelle semplici immagini binarie figura-sfondo è spesso l'unica caratteristica interessante.

Un'evidenza della forma di una figura si trova al confine tra la figura e lo sfondo e descriverla significa mettersi alla ricerca di curve particolari nell'immagine.

La trasformata di Hough è un metodo utile per l'individuazione di tali curve, sia analitiche che non, nel caso di immagini dal bordo binario.

È un metodo che viene detto globale, in quanto si considerano relazioni globali tra i punti dell'immagine, in modo da determinare l'appartenenza di tali punti a linee di forma specificata; è inoltre efficiente in quanto non si prefigge il difficile compito di ricercare una forma, ma di trovare dei picchi nello spazio immagine della curva cercata.

Il metodo si basa sul concetto generale che una curva piana in un sistema di riferimento è definita in forma analitica tramite un insieme di parametri e da un'equazione che lega i parametri alle coordinate $f((x, y), (a_1, a_2, \dots, a_n)) = 0$. In tal caso (x, y) è un punto della curva nello spazio immagine (SI) e (a_1, a_2, \dots, a_n) è una n -upla di valori che individuano un punto nello spazio dei parametri (SP). In particolare:

- ogni punto nello SI corrisponde ad un'ipersuperficie nello SP
- n punti nello SI appartenenti ad una stessa curva generano n superfici che si intersecano in uno stesso punto in SP

Dualmente:

- ogni punto nello SP corrisponde ad una singola istanza della curva nello SI
- n punti nello SP appartenenti ad una stessa ipersuperficie corrispondono ad n curve che si intersecano in uno stesso punto in SI

A partire da questo principio, si possono descrivere le trasformate di Hough.

2.4.1 Trasformata semplice di Hough

La trasformata di Hough semplice si interessa alla rilevazione di allineamenti nell'immagine, quindi di rette.

Per quanto precedentemente detto, data una retta $y = mx + n$, essa è completamente identificata dalla coppia di parametri (m, n) , e lo SP è perciò un piano in cui ogni punto rappresenta una retta.

Dualmente, un punto (x, y) nello SI rappresenta una linea retta $n = x(-m) + y$ nello SP ed ogni punto di questa retta corrisponde ad una retta nello spazio di partenza che passa per il punto (x, y) . Sulla base di ciò, una retta nello SI definita da N punti P_1, \dots, P_N viene identificata come intersezione, nello SP, di N rette ognuna corrispondente ad un P_i .

Dal momento che ogni punto dello SI vota una retta nello SP, quindi in particolare i punti che appartengono a tale retta, il problema si traduce in un problema di ricerca di picchi nel numero di voti che riceve ogni punto dello SP.

Utilizzando questa parametrizzazione per la retta, però, lo SP risulta infinito in quanto m e n possono assumere qualsiasi valore.

Per risolvere questo problema Duda e Hart hanno proposto una modifica, proponendo una rappresentazione alternativa delle rette:

$$x \cos(\theta) + y \sin(\theta) = \rho \quad (2.4)$$

dove ρ rappresenta la distanza della retta dall'origine e $\theta \in [0, \pi]$ il suo orientamento. In tal caso un punto nello SI (x, y) diventa una senoide nello SP. Gli intervalli di variazione di ρ e θ sono in genere fissati sulla base delle caratteristiche dell'immagine originale, tipicamente $-\rho_{max} \leq \rho \leq \rho_{max}$, $-\pi/2 \leq \theta \leq \pi/2$, dove $\rho_{max} = 0.5(NR^2 + NC^2)$ e (NR, NC) sono le dimensioni dell'immagine.

Assumendo per semplicità che un'immagine contenga solo una retta, di parametri (ρ_0, θ_0) costituita dai punti P_1, \dots, P_n l'algoritmo consiste nel:

- quantizzare e rappresentare in un array lo SP (ρ, θ) , definendo $C(\rho, \theta)$ l'accumulatore.
- fissare una soglia o valore minimo di precisione dell'accumulatore
- per ciascun punto $P_i = (x_i, y_i)$ calcolare la senoide r_i corrispondente nello SP
- incrementare tutti i contatori relativi alla senoide r_i nello SP
- selezionare i massimi locali dell'accumulatore (che nell'SI sono le rette), cioè solo le celle associate a valori che superino la soglia fissata:

meccanismo di voting (ciascuna cella dello SP fornisce un contributo o voto alla ricerca della retta)

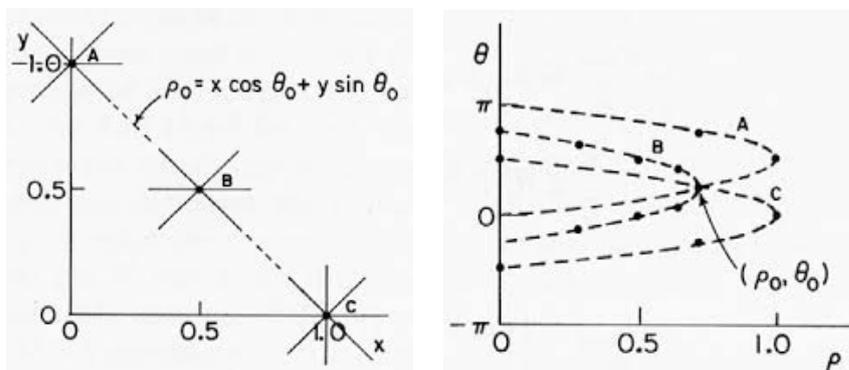


Figura 2.1: Trasformata semplice di Hough

In questo caso tutte le r_i passano per la cella (ρ_0, θ_0) quindi $C(\rho_0, \theta_0) = N$ mentre tutte le altre celle hanno valori più piccoli: selezionando questo massimo si ritrova la retta cercata.

2.4.2 Trasformata generalizzata di Hough

Come abbiamo già visto nell'introduzione al paragrafo, la trasformata di Hough si può generalizzare per la ricerca di altre curve analitiche in un'immagine, ad esempio le forme circolari. Il principio base per la CHT (Circular Hough Transform) è analogo, con qualche modifica, a quello visto per il riconoscimento delle rette: si trasforma un insieme di punti nello SI, in un insieme di voti accumulati nello SP.

Nello specifico, la descrizione delle figure circolari in un'immagine, richiede tre parametri, (a, b, r) :

$$(x - a)^2 + (y - b)^2 = r^2$$

In forma parametrica:

$$\begin{cases} x = a + r \cos(\theta) \\ y = b + r \sin(\theta) \end{cases} \quad (2.5)$$

Dove (a, b) sono due numeri reali che rappresentano le coordinate del centro, r è un numero reale positivo che denota il raggio del cerchio e θ è un angolo che assumerà tutti i valori compresi tra 0 e 2π . Ad un aumento del numero dei parametri, però, corrisponde un aumento nella dimensione dello SP e quindi della complessità dell'algoritmo.

Al fine di semplificare, si utilizzano perciò due approcci al problema: ricerca di cerchi dal raggio fissato o ricerca di cerchi con raggio variabile entro un certo intervallo.

Nel primo caso, supponendo di dover localizzare una circonferenza dal raggio noto r_0 , lo SP si riduce a due dimensioni (come per la retta), essendo necessario stimare solo le coordinate del centro per definire univocamente il cerchio. L'idea di base rimane la stessa: per ogni punto (x, y) dello SI, corrisponde nello SP, un cerchio di raggio r_0 avente centro proprio in (x, y) .

Dalla (2.5) si ottiene:

$$\begin{cases} a = x - \cos(\theta) \\ b = y - \sin(\theta) \end{cases} \quad (2.6)$$

Nello SP si ottiene quindi un insieme di cerchi con stesso raggio r_0 , che si intersecano nel punto (a, b) , il quale corrisponde al centro del cerchio nello SI. Il centro del cerchio da trovare è quindi il punto in comune di tutte quelle circonferenze centrate in un punto qualsiasi (x, y) del suo perimetro.

L'algoritmo che ne segue è il medesimo che viene utilizzato per le rette.

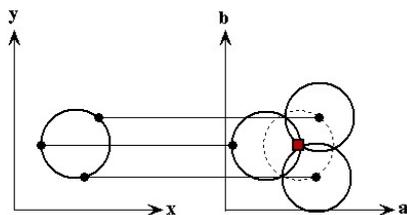


Figura 2.2: CHT

Inizialmente si crea un array bidimensionale, in modo da discretizzare lo SP, ed un accumulatore. Dato che ogni punto del cerchio nello SI contribuisce ad un cerchio nello SP, incrementiamo l'accumulatore in corrispondenza delle celle dell'array che contengono le coordinate appartenenti alla circonferenza di ogni cerchio nello SP. Alla fine del processo, l'accumulatore avrà un picco dove i contributi dei cerchi si sovrappongono, e a tali picco corrisponderà il centro del cerchio originale.

Un insieme di più cerchi con lo stesso raggio può essere trovato con la stessa tecnica. (Fig. 2.3) L'intersezione dei cerchi può corrispondere ai centri cercati (celle rosse) o a falsi positivi, (celle blu). Questi falsi positivi possono essere rimossi osservando la compatibilità con i cerchi dell'immagine originale.

Nel secondo approccio al problema, se il raggio non è noto a priori ma è compreso in un range di valori ammissibili, un cerchio è univocamente determinato da una terna di parametri (a, b, r) , il che comporta che lo SP e

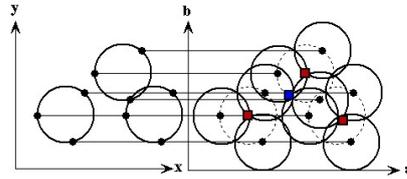


Figura 2.3: CHT

l'array accumulatore diventino tridimensionali: un piano per identificare il centro del cerchio, e l'asse z a cui corrisponde un insieme di valori discretizzati dei possibili raggi.

In particolare, ogni punto di coordinate (x, y) nello SI genera un cono circolare retto nello SP:

$$f((x, y), (a, b), r) = (x - a)^2 + (y - b)^2 - r^2 = 0 \quad (2.7)$$

In questo caso, per tutti i punti dello SI appartenenti ad una circonferenza di raggio r , la terna (a, b, r) da stimare si troverà in corrispondenza della cella accumulatrice in cui si interseca il maggior numero di coni.

Nel procedimento di voto viene quindi scandito l'intervallo discreto dei raggi e per ogni punto vengono incrementate le celle attraversate dalla circonferenza centrata nel punto ed avente raggio pari al valore corrente considerato.

Questo si traduce in una richiesta maggiore dal punto di vista computazionale. Un modo per ridurre la richiesta di memoria consiste nell'utilizzare il

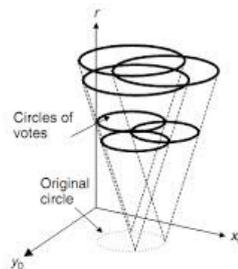


Figura 2.4: CHT

cosiddetto metodo del gradiente, cioè nello sfruttare le informazioni riguardo la direzione del gradiente per ogni punto dello SI.

Se conosciamo l'informazione associata al gradiente in un punto (Fig. 2.5), possiamo ridurre il luogo geometrico nello SP da un cono ad una retta (Fig. 2.5), poiché il centro del cerchio per quel punto dovrà giacere ad una distanza r lungo la direzione del gradiente.

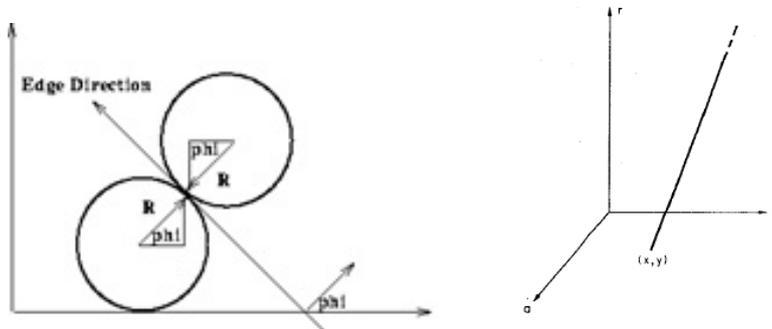


Figura 2.5: Gradiente e Retta

Quindi, se normalmente la parametrizzazione del cerchio richiede tre parametri, utilizzando il gradiente ne serve solo uno. Per far ciò si calcola, attraverso degli operatori, il gradiente locale per ogni punto dello spazio immagine, le cui componenti sono: (g_x, g_y) , la cui intensità è data da $g = \sqrt{g_x^2 + g_y^2}$ e la cui orientazione è $\theta = \arctan(\frac{g_y}{g_x})$. Essendo poi $\cos(\theta) = g_x/g$ e $\sin(\theta) = g_y/g$ e considerando l'equazione parametrica (2.5) possiamo stimare le coordinate del luogo dei centri con:

$$\begin{cases} a = x - r(\frac{g_x}{g}) \\ b = y - r(\frac{g_y}{g}) \end{cases} \quad (2.8)$$

Se la direzione è precisa e si conosce il valore di r allora nell'accumulatore, che è diventato ora bidimensionale, vengono aumentati i punti per i valori (a, b) che soddisfano la (2.8). Tenendo però conto che è improbabile che sia precisa la stima della direzione del gradiente in un punto possiamo aggiungere una tolleranza dovuta all'errore.

In questo accumulatore bidimensionale vengono poi selezionati i centri candidati, i quali devono superare una certa soglia impostata inizialmente. In molti casi non si conosce con certezza il valore del raggio, perciò si procede con il metodo generale di voto, considerando però che stavolta ogni punto dello spazio immagine non contribuisce ad un cono circolare ma bensì ad una retta.

In questo paragrafo ho utilizzato [11, 12, 13, 14].

2.5 Confronto tra i metodi

Prima di passare alla trasformata di Hough, sono state introdotte tre nozioni: la dualità proiettiva, gli spazi di configurazioni e gli spazi di moduli. Questi tre concetti spiegano, attraverso punti di vista diversi, l'idea di base

che c'è dietro il metodo di Hough e nel seguito si evidenzieranno i parallelismi e le differenze tra quest'ultimo e i precedenti.

La dualità proiettiva considera il problema dal punto di vista geometrico: la trasformata di Hough si serve di essa nel momento in cui ad un punto dello spazio immagine, associa un'ipersuperficie nello spazio dei parametri. Facendo, quindi, un parallelismo, lo spazio proiettivo equivale allo spazio immagine mentre lo spazio proiettivo duale allo spazio dei parametri.

Come già accennato, sono coinvolti solo alcuni concetti dell'intero metodo, infatti, parlando delle differenze, ad esempio, la dualità proiettiva coinvolge anche i sottospazi all'infinito, che invece non sono previsti nel metodo di Hough.

Gli spazi di configurazioni studiano il problema dal punto di vista fisico, con l'obiettivo di capire la posizione di un corpo rigido nello spazio.

Nel metodo di Hough per la ricerca di rette, considerando quest'ultima un corpo rigido, si hanno bisogno di due parametri: uno per la distanza dall'origine e uno per la direzione. Così le coordinate generalizzate sono proprio (ρ, θ) e lo spazio delle configurazioni non è altro che lo spazio dei parametri. Anche in questo caso, il concetto non viene ripreso in toto, in quanto gli spazi di configurazioni sono utilizzati per lo studio del moto fisico dei corpi, e si servono di coordinate generalizzate differenziabili, ad esempio, non necessarie per il metodo di Hough.

Infine gli spazi di moduli, che sono utilizzati per problemi di classificazione, vengono anch'essi utilizzati come concetto di base per Hough, considerando la parte geometrico-algebrica del problema. Negli spazi di moduli, come già visto, partendo dalla curva che si vuole studiare, per esempio retta o circonferenza, si definiscono i parametri necessari per identificarla univocamente, per poi creare lo spazio dei parametri o di moduli.

Parallelamente, nel metodo di Hough si considera lo spazio immagine e si definisce lo spazio dei parametri, in questo caso moduli, in base alla curva che si vuole ricercare. Anche in questo caso, a prova che solo una parte del concetto di spazi di moduli viene ripreso, si osserva che non è considerata la parte della classificazione in base a isomorfismo.

In conclusione, la trasformata di Hough utilizza concetti matematici molto simili tra loro, per quanto differenti, poiché preposti alla risoluzione di problemi in campi di ricerca diversi, come fisica o geometria, e quindi con particolari esigenze che li rendono differenti.

Capitolo 3

Applicazioni del metodo della trasformata di Hough

La trasformata di Hough, come abbiamo visto, è un metodo matematico per la visione artificiale. In particolare, grazie ad essa un computer è capace di studiare l'immagine a livello di rilevazione di curve particolari. La sua importanza è data dal fatto che molte volte la descrizione dei contorni e delle forme in un'immagine è l'unica cosa interessante; un esempio sono queste tre applicazioni, riguardanti ambiti completamente diversi come la geologia e la medicina, che verranno presentate di seguito.

3.1 Applicazione alla geologia: allineamento dei vulcani

3.1.1 Introduzione al problema generale

La prima applicazione che viene presa in esame riguarda lo studio della distribuzione vulcanica lungo l'arco vulcanico della Sonda.

Quest'ultimo è un arco vulcanico insulare, che interessa la zona indonesiana estendendosi dall'isola di Sumatra a est di Flores, attraversando Giava e Nusa Tenggara. Si è formato a causa dello scivolamento della Placca Indo-Australiana al di sotto della Placca Sonda dell'Eurasia e comprende, oltre le varie isole, ben 35 vulcani. (Fig. 3.1)

Quest'ultimi sono posti a scaglione in quattro gruppetti di catene (distribuzione *en echelon*) ognuna lunga dai 500 ai 700 km e si trovano ai margini delle placche al limite del continente asiatico. Generalmente i vulcani tendono a localizzarsi nelle zone di compressione, dove le placche tendono a convergere, ma esistono casi in cui si trovano in zone di divergenza o all'interno delle

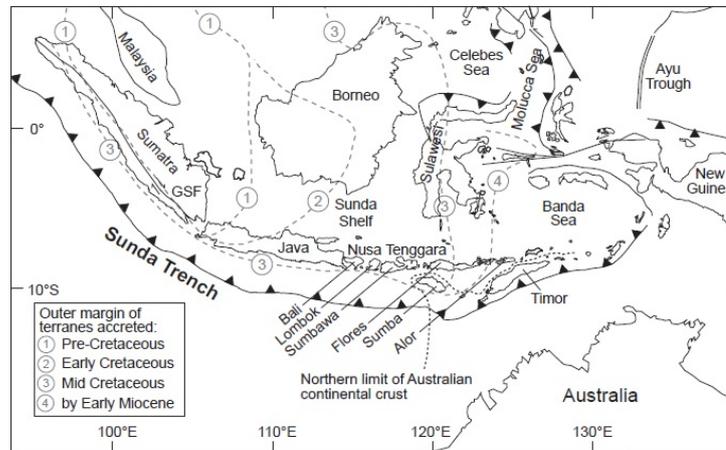


Figura 3.1: L'arco della Sonda centrale

placche stesse (punti caldi).

Cercare di capire la distribuzione dei vulcani negli archi della litosfera ci consente di capire le variazioni di distanza tra i vulcani e la zona delle placche sismiche al centro dell'arco della Sonda, dove ci sono piccole variazioni nella geometria delle placche o del loro grado di convergenza.

A partire da ciò si riesce poi a capire l'era di formazione di ogni singolo vulcano, la durata della sua attività e il tipo di materiale magmatico che fuoriesce al momento dell'eruzione.

Lo studio dei vulcani della Sonda fornisce un importante modello per capire la distribuzione vulcanica, per molte ragioni. Infatti, l'arco della Sonda ha ospitato abbondanti attività vulcaniche nella storia e la maggioranza dei vulcani sono completamente emersi dalle acque marine e perciò possono essere prontamente identificati attraverso una mappa locale o un'immagine satellitare.

Inoltre ci sono differenti tipi di tensione lungo tutto l'arco insulare (importanti in quanto delineano la distribuzione dei vulcani): a ovest un'alta convergenza obliqua causa significanti sforzi sull'isola di Sumatra, ad est l'arco è andato a collidere con la litosfera del continente australiano.

Gli archi vulcanici sono già modellizzati come una disposizione allungata di centri vulcanici, quindi lo scopo di questo studio è determinare quale delle modellizzazioni per tali centri sia la migliore, tra piccoli cerchi o grandi cerchi in una superficie sferica, cioè se sono singoli segmenti arcuati o lineari. La trasformata di Hough è un metodo che fornisce lo strumento più appropriato per tale obiettivo.

3.1.2 Applicazione della trasformata di Hough

Le trasformate di Hough sono applicate alle immagini per riconoscere le figure geometriche che possono essere espresse matematicamente, in questo caso rette. Le moderne applicazioni di questo metodo sono molto diffuse, soprattutto in fisica e nella tecnologia della visione della robotica, mentre la tecnica ha ricevuto solo delle piccole applicazioni come strumento per analizzare lo spazio e le strutture geologiche.

Alcuni tentativi sono stati fatti per identificare gli allineamenti vulcanici, in modo maggiore per i *monogenetic volcanic fields* (gruppi di vulcani che eruttano una sola volta). A tal fine è stata utilizzata la trasformata semplice di Hough, in quanto ciò che si vuole evidenziare nella mappatura dell'arco vulcanico sono gli allineamenti presenti nella struttura.

Si rimanda al par. 2.4.1 per gli approfondimenti, in cui è presente una spiegazione schematica sul modello della trasformata semplice di Hough.

Il processo utilizzato nell'applicazione in esame si serve, inizialmente, di una mappa in proiezione dell'arco della Sonda centrale in scale di grigio, in cui ogni centro vulcanico è rappresentato con un cerchio di un'uniforme grandezza, posto in corrispondenza del luogo in cui è posizionata la camera magnetica. (Fig. 3.3a)

Esso consiste in un algoritmo in Matlab volto ad attuare l'analisi dell'immagine attraverso le trasformate di Hough (Fig. 3.2), e si compone di tre parti: la prima produce $(\rho - \theta)$, spazio dei parametri, la seconda evidenzia e sceglie la più alta frequenza di picchi da questo grafico e la terza proietta le corrispondenti linee in un'immagine in scala di grigi. È necessario esprimere due valori manualmente: il minimo della lunghezza della linea accettabile, stabilito in modo che se una linea di tale lunghezza connette tre o meno punti dati, non è accettabile come un allineamento significativo e il massimo valore dell'intervallo tale per cui la linea può connettere due vulcani separati da tale intervallo.

Successivamente si analizza una seconda mappa, in cui i vulcani non hanno però una grandezza uniforme: ogni vulcano è scalato in un cerchio di $3/4$ il diametro della sua base, come quanto determinato dalle mappe geologiche e dalle immagini satellitari (Fig. 3.3b).

Questo approccio effettivamente pesa nell'analisi, però permette l'ipotesi di andamenti lineari fra strutture, anche se i centri geometrici veri e propri non sono allineati e inoltre minimizza il problema della rappresentazione dovuto al fatto che le strutture vulcaniche potrebbero non giacere direttamente sopra il principale condotto di alimentazione in profondità. Si sceglie una scala di $3/4$ del diametro poiché il cerchio scalato rispetto alla grandezza completa del diametro potrebbe generare una sovrapposizione di molte strutture che

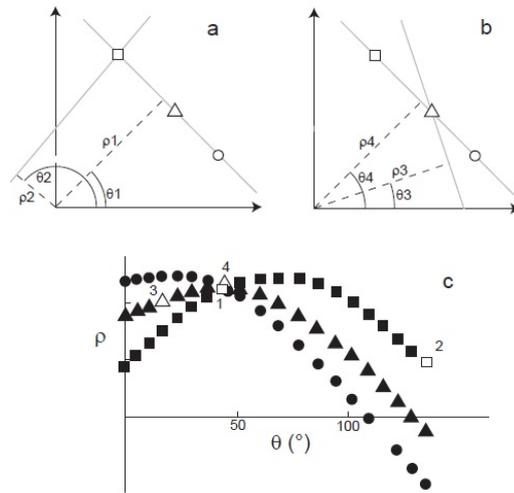


Figura 3.2: Trasformata di Hough applicata ai vulcani

fanno parte di sistemi vulcanici differenti. I risultati ottenuti dall'analisi delle due mappe sono poi confrontati (Fig. 3.3) per capire quale sia la migliore modellizzazione per la distribuzione dei vulcani. Come si vede dalla figura, la lunghezza dei segmenti e il numero di punti presenti in ognuno di essi variano e molti punti sono inclusi in più di un segmento. Ci focalizziamo solo in sei di essi che sono comuni ad entrambi gli approcci e per ognuno di questi segmenti, all'interno delle ellissi disegnate (Fig. 3.3c), è calcolato un coefficiente di determinazione che denoteremo con R^2 .

Quest'ultimo assume valori compresi tra 0 e 1 ed è calcolato in modo tale che più risulti alto, più il segmento tende a minimizzare la deviazione dalla posizione reale dei vulcani. Analizzando l'immagine (Fig. 3.3c), il segmento più a ovest ($R^2 = 0.947$) comprende lo stesso e unico insieme di punti in entrambi gli approcci e dunque possiamo considerare che questi definiscano una valida struttura geologica.

Ci sono, invece, due possibili segmenti nella parte più a est, ma dal momento che le differenze tra essi sono trascurabili, possiamo trattarli come un singolo elemento. ($R^2 = 0.774$). Al centro ci sono tre potenziali segmenti, in cui il più piccolo dei tre ha il più basso valore del coefficiente ($R^2 = 0.917$) e tutti i suoi punti possono essere approssimati in uno degli altri due possibili segmenti di questa regione.

Un potenziale segmento ha un alto valore del coefficiente ($R^2 = 0.958$), ma non copre interamente la parte centrale dell'area in quanto si estende solamente dal punto A a B, sfiorando solamente il punto C. Se lo si termina manualmente, si crea un segmento forzato centro-est ($R^2 = 0.974$) che com-

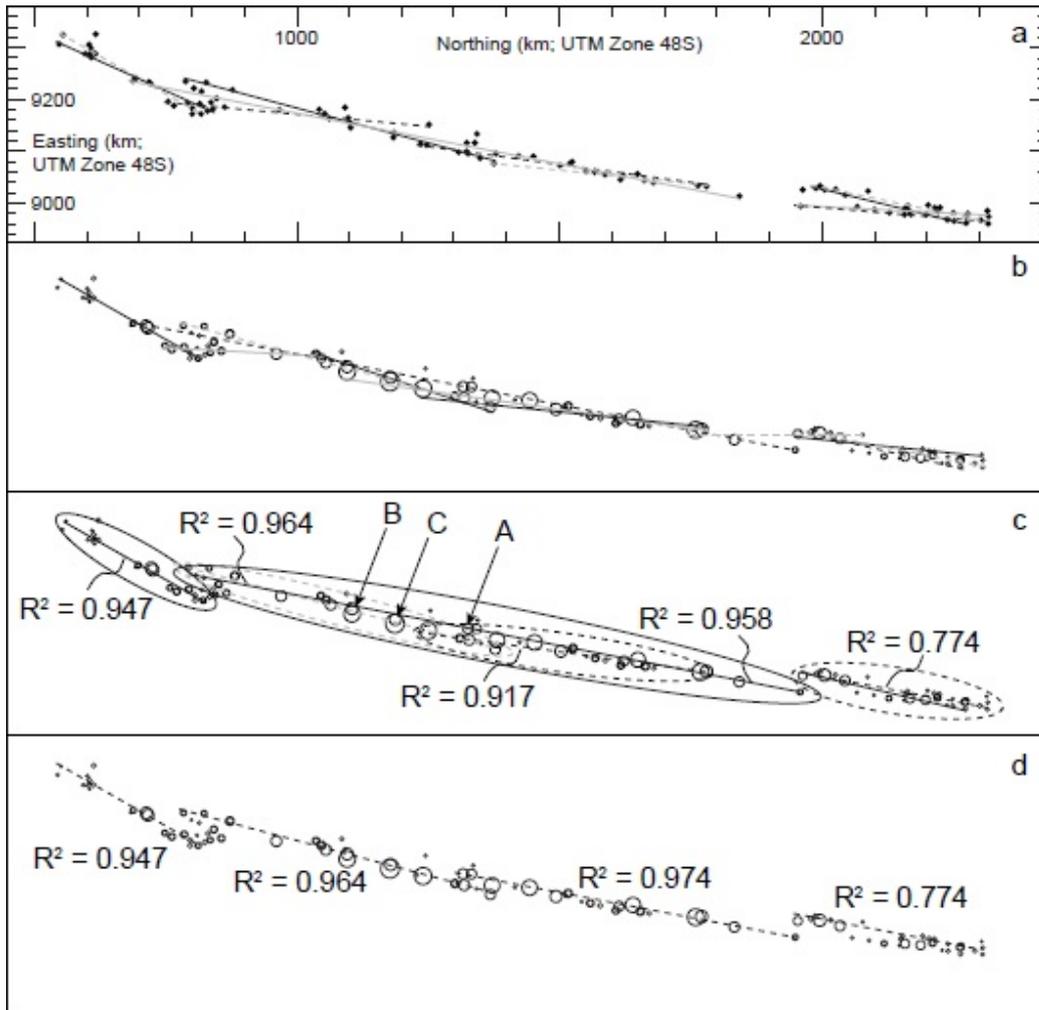


Figura 3.3: Confronto

pleti il potenziale segmento centro-ovest ($R^2 = 0.964$). Quindi un iniziale tentativo di un'oggettiva definizione di allineamento identifica quattro segmenti, come raffigurato in (Fig. 3.3d). A partire da questa analisi, è possibile osservare che la distribuzione dei vulcani nell'arco della Sonda non si adatta all'uso dei piccoli cerchi, in quanto la deviazione risulta assai elevata, cosa che non accade se si utilizzano i grandi cerchi.

Un'ulteriore miglioramento nella ricerca della retta più adatta per modellizzare la distribuzione dei vulcani, si è ottenuta escludendo dall'analisi quei vulcani che si trovano in strutture locali parallele ai segmenti presi in esame (triangolini arancioni in Fig. 3.4). Ciò che ne risulta (Fig. 3.5) è un modello in cui il coefficiente R^2 si alza e quindi la mappa risulta più precisa. La stessa precisione potremmo ottenerla, modellizzando i vulcani con dei cerchi un po' più piccoli dei grandi cerchi, e dividendo i segmenti in modo che diventino più numerosi e più piccoli.

In questo paragrafo ho utilizzato [15].

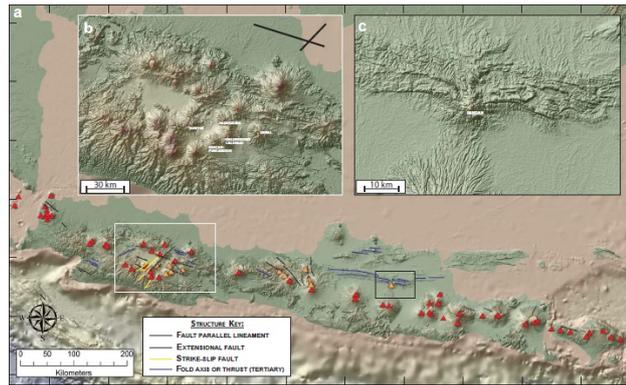


Figura 3.4: Disposizione dei vulcani nell'arco della Sonda

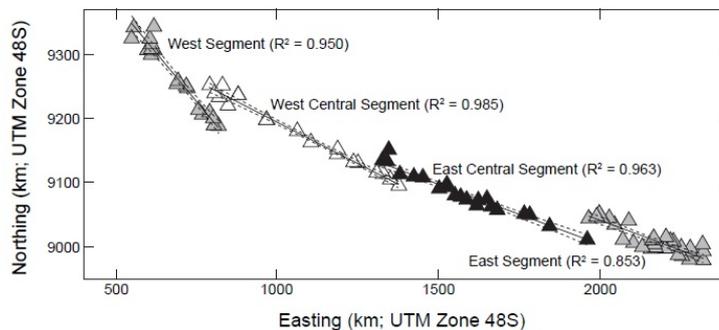


Figura 3.5: Allineamenti nell'arco della Sonda

3.2 Applicazione alla medicina: identificazione dell'aorta

3.2.1 Introduzione al problema generale

Nello studio in esame la trasformata di Hough è utilizzata nelle immagini ecografiche per l'identificazione dell'aorta. Quest'ultima è la più grande arteria del corpo umano, si origina dal ventricolo sinistro del cuore, con un diametro di circa 3 cm, e trasporta il sangue ossigenato a tutte le parti del corpo.

Il fine della ricerca è diagnosticare la presenza di un eventuale tumore all'esofago, con conseguente espansione (metastasi) negli organi vicini. La malattia, generalmente, si forma nella parete dell'esofago e si espande verso l'esterno, negli organi circostanti (per esempio l'aorta o i gangli) e lo studio delle metastasi ci fornisce molte informazioni riguardo l'inizio e l'evoluzione della stessa. Nel caso specifico viene presa in esame l'aorta, essendo uno degli organi che potrebbe ammalarsi, ed in particolare si vuole localizzarla per poi accertarsi della presenza o meno di una metastasi. Il sistema utilizzato nel rilevamento sfrutta le immagini ecografiche ad ultrasuoni, in quanto tale tecnica è passiva e non intrusiva, quindi meno pericolosa per i pazienti.

L'interpretazione delle immagini ecografiche è computata con due tipi di interferenze durante l'acquisizione dei dati, le quali inquinano i risultati con: macchie, le quali derivano dalle diffrazioni degli ultrasuoni da parte di micro strutture aventi una grandezza comparabile alla lunghezza d'onda degli ultrasuoni e luci brillanti che si sovrappongono intorno al centro dell'esofago. Tutti questi fattori contribuiscono al fatto che l'interpretazione delle immagini risulti difficile, anche per medici specialisti. A tal fine la trasformata di Hough è utilizzata in modo da avere una precisa rilevazione e localizzazione dell'aorta, riducendo le interferenze.

3.2.2 Applicazione della trasformata di Hough

La trasformata di Hough si è dimostrata un metodo efficiente per cercare curve parametrizzate in un'immagine e in questo caso sarà discussa la sua applicazione alle immagini a ultrasuono della regione intorno all'esofago, per la localizzazione dell'aorta. In particolare si farà uso della trasformata generalizzata di Hough, i cui dettagli sono spiegati nel par. 2.4.2.

L'algoritmo proposto ha come caratteristica quella di integrare cose conosciute o fonti elementari, come i punti nello spazio immagine, a due tipi di incertezze: l'imprecisa stima della direzione del gradiente e il raggio preciso dell'aorta.

In particolare, la fase di integrazione ha luogo nel momento del processo di voto (par. 2.4.2), in quello che viene chiamato l'accumulation kernel (AK). L'AK è definito come un'area ristretta nello spazio dei parametri, delimitata dal contributo di ogni fonte elementare (nel caso della ricerca di rette è una retta, o nella ricerca di cerchi dal raggio fissato, un cerchio) e ciò che spinge a darne una definizione e denominazione precisa sta nel fatto che esso varia, in base al tipo e al numero di incertezze su cui è basato l'algoritmo.

Nel caso in esame, la localizzazione dell'aorta prevede la ricerca di cerchi dal raggio fissato, in quanto tendenzialmente essa ha raggio costante. Per minimizzare i costi si usa il metodo del gradiente (par. 2.4.2), dunque ogni punto nello spazio immagine è univocamente determinato da 4 parametri: la posizione, la direzione e l'intensità del gradiente.

Nella fase di voto, se utilizziamo come informazione solo la posizione del punto nello spazio immagine, allora per ogni punto p_k , viene incrementata di un'unità ogni cella corrispondente ad AK_k : *total ignorance accumulation*. In questo caso non c'è nessuna preferenza di incremento per le celle di accumulazione di AK_k . Se invece aggiungiamo come informazione l'intensità e la direzione del gradiente (ϕ_k), ogni cella AK_k sarà aumentata in funzione del gradiente: *accumulation evidence*.

Supponendo che le informazioni sulle direzioni del gradiente siano precise, allora per ogni punto dello spazio immagine, solo una cella di AK_k dovrà essere aumentata: *total knowledge accumulation*.

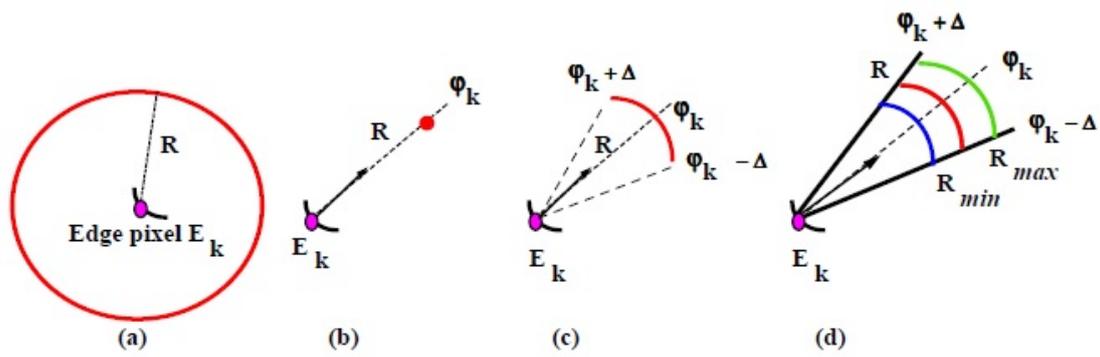
Dato che non si è mai in grado di conoscere con precisione la direzione del gradiente, esiste un altro metodo di incremento, intermedio: si incrementano solo le celle situate in un arco con direzione ϕ_k e centrate in x_k e y_k . All'incertezza sulla direzione del gradiente, si aggiunge anche l'incertezza riguardo il raggio del cerchio.

Questa è dovuta al fatto che cerchiamo una forma "vicina" al cerchio, dove il raggio è definito come un intervallo di valori e non un singolo valore, e inoltre i valori dei raggi non possono essere conosciuti con precisione. In ognuna di queste casistiche, l' AK_k avrà una diversa conformazione. (Fig. 3.6)

Nell'ultimo caso considerato, l'accumulation evidence sarà una funzione decrescente di parametro $|\phi - \phi_k|$, in cui si nota che l'accumulation evidence è uguale a 1 se $\phi_i = \phi_k$ per tutti i raggi compresi tra r_1 e r_2 . (Fig. 3.7)

Il raggio ottimale R_{opt} , può essere determinato usando un semplice accumulatore 1-dimensionale o un istogramma dei possibili raggi: per ogni punto la cui normale punta verso il centro del cerchio, è misurata la distanza dal centro e immagazzinata nell'istogramma, alla fine di questo processo si trova un massimo nell'istogramma; il quale rappresenta il raggio ottimale. (Fig. 3.8)

Se si ha anche un'incertezza sull'informazione riguardante la posizione dell'aorta nello spazio immagine, si utilizzano delle funzioni nello spazio dei



a) Total Ignorance Accumulation, b) Total Knowledge Accumulation, c) Imprecise Direction, accumulation and d) Imprecise Direction-Radius Accumulation.

Figura 3.6: Tipologie di accumulazione

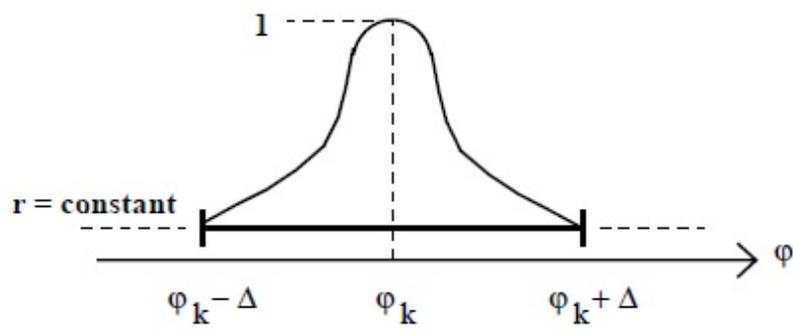


Figura 3.7: Grafico dell'*accumulation evidence*

parametri che assumono il valore zero nelle aree dove l'aorta non può essere rilevata, e il valore uno nel resto dell'immagine.

I risultati ottenuti in termini di rilevamento e localizzazione dell'aorta sono estremamente promettenti e mostrano la robustezza dell'approccio proposto anche se ci sono delle interferenze nelle immagini a ultrasuoni.

Due esempi di risultati sono dati dalla Fig. 3.9: la a mostra una immagine endoscopica di un paziente normale e la b una regione dell'aorta di un paziente che ha un tumore all'esofago esteso che comporta una degradazione delle pareti dello stesso e della precedente. In entrambi i casi, i cerchi risultanti approssimati dell'aorta sono di gran precisione.

In questo paragrafo ho utilizzato [16].

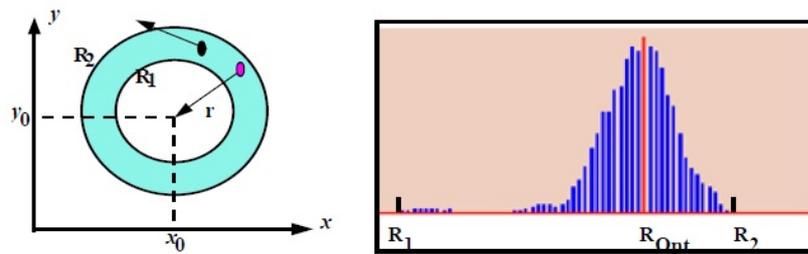


Figura 3.8: Istogramma Raggio ottimale

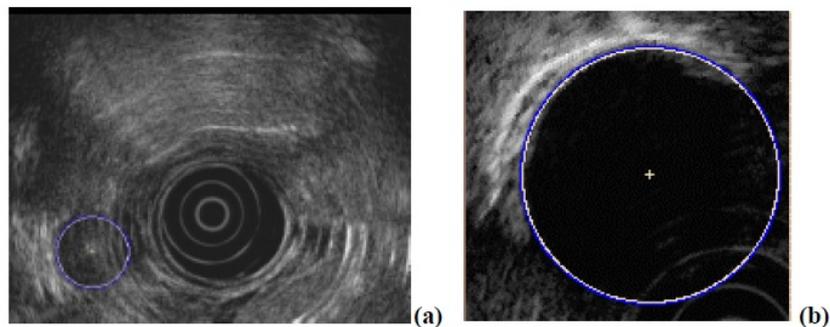


Figura 3.9: Sezioni dell'aorta

3.3 Applicazione alla medicina: analisi delle vescicole di membrana

3.3.1 Introduzione al problema generale

Negli ultimi anni, le vescicole di membrana sono diventate potenti strumenti ampiamente utilizzati per studiare i processi di membrana. Al tal fine si utilizzano vescicole più grandi in modo che possano essere rilevate tramite immagini di microscopia digitale, dell'ordine di $1 - 10\mu m$, chiamate vescicole unilamellari, le quali esibiscono una grande varietà di forme differenti, in particolare sferiche: le GUVs (Giant Unilamellar Vesicles).

Tra le varie applicazioni, le GUVs sono ampiamente utilizzate per studiare i meccanismi di quelle molecole che formano dei pori nella membrana, nonché le proteine relative a quest'ultima tramite i profili radiali delle vescicole e la sua permeabilizzazione.

In particolare, però, in questa applicazione ci si concentrerà sulla localizzazione e sulla cinetica di dispersione delle vescicole, in modo da ricavarne i meccanismi che ne guidano il moto e la funzione. Tuttavia, per ottenere un'elevata precisione statistica riguardo ciò che si vuole analizzare, devono essere considerate diverse centinaia di GUVs per ogni esperimento, e se eseguita manualmente, è un'analisi che richiede molto tempo e può presentare delle soggettività.

Infatti un ricercatore potrebbe tendere a selezionare GUVs che presentano alcune caratteristiche spiccate riguardanti, per esempio, la dimensione e quindi introdurre una distorsione nel campione. Al fine di ovviare a questi inconvenienti, si esegue un rilevamento automatico di una grande popolazione di GUVs, basata sulla trasformata circolare di Hough (CHT, par. 2.4.2), una tecnica robusta per trovare le forme nelle immagini.

L'automatizzazione non è solo vantaggiosa in termini di risparmio di tempo, ma garantisce anche una analisi statistica imparziale nella massima misura possibile. Il software, inoltre, è potente, non solo nel rilevare oggetti circolari in immagini al microscopio, ma anche oggetti che moderatamente discostano dalla forma circolare. Per l'analisi in movimento è stato implementato, invece, un algoritmo di tracciamento che consenta di seguire le singole vescicole, rendendo il software applicabile a GUVs in movimento.

3.3.2 Applicazione della trasformata di Hough

Le immagini ricavate dalla microscopia digitale, acquisite in condizioni reali, non sono di solito adatte per il rilevamento di curve, ma hanno bisogno di una certa pre-elaborazione per aumentarne il contrasto. Infatti la maggior

parte delle immagini rappresentanti le GUVs ha uno scarso contrasto, a causa della bassa concentrazione del colorante utilizzato e della bassa potenza del laser.

Come già anticipato, si passa poi all' algoritmo di identificazione delle GUVs, il quale è basato sul CHT (per i dettagli si rimanda al par. 2.4.2). Le figure 3.10b e 3.10c mostrano l'array di accumulo finale per una tipica immagine delle GUVs (Fig. 3.10a), rispettivamente in due e tre dimensioni, quindi le due immagini si riferiscono all'identificazioni di forme circolari dal raggio fissato, la prima, e dal raggio variabile, la seconda. I picchi indicano chiaramente la posizione spaziale delle GUVs.

Lo strumento non è solo preposto all'individuazione di vescicole perfettamente sferiche, ma può inoltre funzionare con oggetti che deviano leggermente dalla forma sferica. La figura 3.10d, è ottenuta allungando nell'immagine 3.10a le vescicole circolari di un fattore 1,5 nella direzione orizzontale, trasformandole in ellissi.

Dato che le GUVs, per effettuare i trasporti di membrana, debbono muoversi

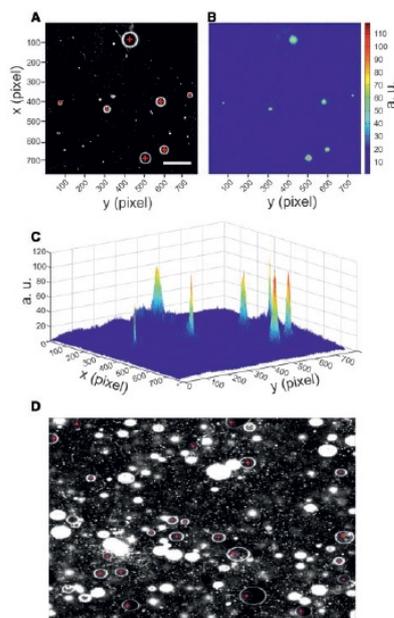


Figura 3.10: Applicazione del CHT alle GUVs

e diffondersi, si vuole ampliare gli studi ai processi dinamici, concentrandosi sugli algoritmi di tracciamento.

In particolare il metodo presuppone la creazione di serie temporali, registrate in file di stack con molti frame (maggiore di 100). Le GUVs, diffondendosi, passano da un frame all'altro e, potendosi spostare dentro e fuori dalla

regione osservata, possono portare ad un cambiamento nel numero iniziale osservato. (Fig. 3.11)

L'algoritmo di riconoscimento dei GUVs è applicato ad ogni frame e il risultato è un insieme di T (numero totale di frame nella sequenza di immagini) matrici C^t con righe $[x_p, y_p]_{p=1}^{N_t}$ dove N_t è il numero di GUVs nel frame t e x_p e y_p sono le coordinate del centro di ogni GUV nel frame t identificata tramite CHT. L'algoritmo di tracciamento si basa sull'identificazione di punti corrispondenti allo stesso GUV nella sequenza di frame.

Sia \mathfrak{F} l'insieme di punti identificati p_i ($i = 1, \dots, N_t$) nel frame t e $R \in \mathbb{Z}$ un parametro che specifica quanti frame futuri saranno considerati. Per ogni insieme \mathfrak{F}_r ($r = 1, \dots, R$) di punti identificati q_j è associata una matrice G_r^t , in cui il generico elemento (i, j) è uguale a 1 se p_i nel frame t e q_j nel frame $t + r$ sono prodotti dalla stessa vescicola, diversamente è uguale a zero.

Quindi, ogni riga e colonna di G_r^t può contare al massimo un 1 e tutti gli altri sono 0. I punti p_i e q_j sono associati allo stesso GUV se la distanza tra i punti che identificano il loro centro è più piccola di ogni altra distanza possibile ed è allo stesso tempo minore di una distanza definita.

Se per una $r > 1$ abbiamo $G_r^t(i, j) = 1$ e $G_{r'}^t(i, j) = 0$ per ogni $r' = 1, \dots, r - 1$, allora una GUV è identificata con successo nel frame t e $t + r$ ma non nei frame intermedi. Comunque assumiamo che tale GUV sia presente anche nei frame intermedi e che la mancanza del risultato sia dovuta al fatto che l'efficienza del metodo è limitata. In questo caso, la mancata localizzazione del punto è automaticamente aggiunta al frame intermedio.

Il centro $\vec{x}_{t+r'}$ della vescicola mancante nel frame $t + r'$ è posto a $\vec{x}_{t+r'} = \vec{x}_t + \frac{r'}{r}(\vec{x}_{t+r} - \vec{x}_t)$ e il raggio $R_{t+r'}$ a $R_{t+r'} = [(r - r')R_t + r'R_{t+r}]/r$. In riferimento a questo poniamo $G_{r'}^t(i, j) = 1$ per ogni $r' = 1, \dots, r - 1$.

Siano t e t' due frame con $t' > t$. Il coefficiente di valore 1 della matrice prodotto $B_{tt'} = \prod_{i=t}^{t'} G_1^i$ connette i punti individuati fra i due frame. $B_{tt'}(i, j) = 1$ significa che l' i -esimo punto individuato nel frame t e il j -esimo individuato nel frame t' sono prodotti dalla stessa GUV.

In questo paragrafo ho utilizzato [17].

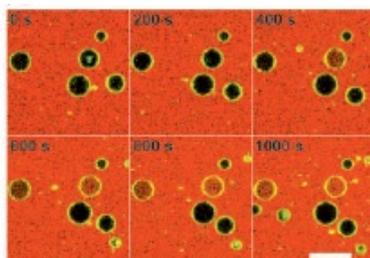


Figura 3.11: GUVs in movimento

Capitolo 4

Conclusione

Questo elaborato si è occupato principalmente di descrivere uno dei metodi utilizzati dalla visione artificiale per l'individuazione di forme particolari nell'immagine, la trasformata di Hough, e di fornire delle applicazioni a prova dell'importanza e del vasto uso di tale tecnica.

Inizialmente ci fornisce una panoramica sui problemi propri della visione artificiale, lasciando aperto il quesito sulla ricerca di un metodo flessibile che possa rendere le macchine capaci di vedere. Successivamente, ci descrive il metodo matematico della trasformata di Hough, sia semplice che generalizzata, a prova che dei passi avanti nella ricerca sono stati fatti.

Questo concetto è ulteriormente evidenziato nell'ultimo capitolo, in cui vengono presentate tre applicazioni recenti, in cui tre problemi assai complicati e urgenti vengono risolti col metodo della trasformata.

Bibliografia

- [1] B. Horn, "*Robot Vision*". MIT Press Cambridge, MA, USA. (1986)
- [2] D. Marr, "*Vision*". MIT Press Cambridge, MA, USA. (1982)
- [3] M. Ferri, "*Visione delle macchine: una sfida anche per i matematici*".
In: Bollettino dell'Unione Matematica Italiana 4.1, pp.85-115. (2001)
- [4] A. Hulbert, T. Poggio, "*Rendere le macchine (e l'intelligenza artificiale) in grado di vedere*". In: Sistemi Intelligenti 1.1, pp.75-102. (aprile 1989)
- [5] <http://www.istanze.unibo.it/oscar/mente/mente04.htm>.
- [6] E. Sernesi, "*Geometria 1*". Bollati Boringhieri, Torino. (1989)
- [7] E. Fortuna, R. Frigerio, R. Pardini "*Geometria proiettiva: Problemi risolti e richiami di teoria*". Springer- Verlag Italia. (2011)
- [8] G.J. Sussman, "*Structure and Interpretation of Classical Mechanics*". MIT Press, London. (2001)
- [9] <https://www.ma.utexas.edu/users/benzvi/math/pcm0178.pdf>.
- [10] David Mumford, "*Geometric invariant theory*". Springer-Verlag, Berlin-New York. (1965)
- [11] K.R. Sloan, D. R. Ballard, "*Experience with the generalized transform*".
In: 5th IJCPR, Miami Beach, pp. 174-179. University of Rochester. (1980)
- [12] Kyewook Lee, "*Application of the Hough transform*". University of Massachusetts, Lowell. (January 2006)
- [13] D.H. Ballard, "*Generalizing the hough transform to detect arbitrary shape*". In: Pattern Recognition 11.2, pp.111-122. (1981)

- [14] O.Duda and Peter E. Hart, "*Use of the Hough transformation to Detect Lines and Curves in Pictures*". In: Communications of the ACM, 15.1, pp. 11-15. (1972)
- [15] Adam Pacey, Colin G. Macpherson, , Ken J.W. McCaffrey, *Linear volcanic segments in the central Sunda Arc, Indonesia, identified using Hough Transform analysis: Implications for arc lithosphere control upon volcano distribution*. In: Earth and planetary science letters, 369-370, pp. 24-33. (2013)
- [16] B. Solaiman, B. Burdsall and Ch. Roux, *Hough transform and uncertainty handling. Application to circular object detection in ultrasound medical images*. In: Proc. Int. Conf. Image Processing (ICIP98), 3, pp. 828-831. (1998)
- [17] Eduard Hermann, Stephanie Bleicken, Yamunadevi Subburaj and Ana J. Garcia-Saez*, *Automated analysis of giant unilamellar vesicles using circular Hough transformation*. In: Bioinformatic 30.12, pp. 1747-1754. (2014)