

SCUOLA DI INGEGNERIA E ARCHITETTURA

DIPARTIMENTO DI  
INGEGNERIA DELL'ENERGIA ELETTRICA E DELL'INFORMAZIONE  
“GUGLIELMO MARCONI” - DEI  
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA ELETTRONICA

Tesi in  
BIOIMMAGINI E INGEGNERIA CLINICA M

PROGETTO E SVILUPPO DI UNA  
INTERFACCIA GRAFICA E DI  
SOFTWARE PER LA  
SEGMENTAZIONE DI BIO-IMMAGINI

*Candidato:*  
Marco Mongitore

*Relatore:*  
Chiar.mo Prof. Claudio Lamberti



*La differenza tra quello che sei e quello che vorresti essere sta  
in quello che fai.*



# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Bioimmagini</b>	<b>3</b>
2.1	Immagini ad ultrasuoni . . . . .	6
2.1.1	Generazione degli ultrasuoni . . . . .	7
2.1.2	Propagazione degli ultrasuoni . . . . .	9
2.1.3	Tecnica ad eco-impulsi, scansione e modi di visualizzazione . . . . .	12
2.2	Radiazioni Ionizzanti . . . . .	15
2.2.1	Sorgenti . . . . .	17
2.2.2	Dose . . . . .	21
2.2.3	Effetti delle Radiazioni Ionizzanti . . . . .	24
2.2.4	Radioprotezione . . . . .	26
2.3	Immagini a raggi X . . . . .	27
2.3.1	Apparecchiatura a Raggi X . . . . .	28
2.4	La Tomografia Computerizzata (CT) . . . . .	31
2.4.1	Geometrie di scansione . . . . .	32
2.4.2	Evoluzioni e spiral CT . . . . .	34
2.5	Tomografia ad Emissione di Positroni (PET) . . . . .	36
2.6	Immagini di risonanza magnetica (MR) . . . . .	39
2.6.1	Principi fisici . . . . .	41
2.6.1.1	Spin e momento magnetico . . . . .	41
2.6.1.2	Segnale MR . . . . .	46
2.6.2	Sequenze di eccitazione . . . . .	52
2.6.2.1	Saturation Recovery . . . . .	52

2.6.2.2	Spin echo . . . . .	53
2.6.2.3	Inversion Recovery . . . . .	54
2.6.3	Risonanza Magnetica Funzionale . . . . .	54
<b>3</b>	<b>Elaborazione delle immagini digitali</b>	<b>55</b>
3.1	Generalità . . . . .	55
3.1.1	L'imaging dalle radiazioni EM . . . . .	56
3.1.2	Processi tipici del digital image processing . . . . .	57
3.2	Segmentazione . . . . .	59
3.2.1	Metodi basati sulla soglia . . . . .	59
3.2.1.1	Sogliatura globale . . . . .	61
3.2.1.2	OTSU . . . . .	63
3.2.1.3	Soglie multiple . . . . .	67
3.2.2	Metodi basati sull'individuazione di edge, linee, punti . . . . .	69
3.2.2.1	Laplaciano . . . . .	72
3.2.2.2	Operatori di gradiente: Sobel, prewitt, roberts . . . . .	73
3.2.2.3	Log . . . . .	76
3.2.2.4	Canny . . . . .	77
3.2.3	Segmentazione basata sulle regioni: region growing . . . . .	81
3.2.4	Segmentazione morfologica watershed . . . . .	82
3.2.5	Segmentazione basata sui contorni attivi . . . . .	87
3.2.5.1	Chan-Vese . . . . .	89
3.2.5.2	Geodesic . . . . .	92
3.2.5.3	Snakes . . . . .	93
3.2.5.4	Fast Marching . . . . .	95
<b>4</b>	<b>Segmentatore automatico</b>	<b>99</b>
4.1	Guida all'utilizzo, interfaccia . . . . .	100
4.1.1	Tecniche di base . . . . .	100
4.1.2	Tecniche avanzate . . . . .	109
4.1.3	Tecniche manuali . . . . .	114
4.1.4	Strumento di confronto . . . . .	114
4.2	Il software . . . . .	117

4.2.1	Generalità . . . . .	117
4.2.2	Implementazione algoritmi . . . . .	117
4.2.2.1	Tecniche di base . . . . .	117
4.2.2.2	Tecniche avanzate . . . . .	120
4.2.2.3	Tecniche manuali . . . . .	123
4.2.3	Implementazione dell'interfaccia . . . . .	123
4.3	Parametri di confronto . . . . .	125
<b>5</b>	<b><i>Risultati</i></b>	<b>129</b>
<b>6</b>	<b>Conclusioni</b>	<b>133</b>
6.1	Applicazioni . . . . .	134
6.2	Sviluppi futuri . . . . .	135



## Elenco delle figure

2.1	<i>Schema di un'onda ultrasonica.</i>	7
2.2	<i>Rappresentazione di un fascio ultrasonico: a) Non focalizzato, b) focalizzato.</i>	8
2.3	<i>Profilo dell'intensità di un fascio ultrasonico.</i>	9
2.4	<i>Interazione tra ultrasuoni e materia.</i>	10
2.5	<i>Legge di Snell per gli ultrasuoni.</i>	11
2.6	<i>Risoluzione spaziale di un fascio di ultrasuoni.</i>	12
2.7	<i>Ecografia di un cuore realizzata con tecniche che sfruttano l'acustica non lineari.</i>	15
2.8	<i>Spettro elettromagnetico.</i>	16
2.9	<i>Decadimento <math>\alpha</math>, <math>\beta^-</math> e raggi <math>\gamma</math>.</i>	18
2.10	<i>Potere penetrante dei diversi tipi di radiazioni ionizzanti.</i>	18
2.11	<i>Schema di produzione dei raggi X.</i>	29
2.12	<i>Radiografia al torace.</i>	30
2.13	<i>Diverse generazioni di scanner per CT, rispettivamente da A a D dalla 1<sup>a</sup> alla 4<sup>a</sup>.</i>	32
2.14	<i>Apparecchiatura per tomografia computerizzata.</i>	35
2.15	<i>Esempi di immagini solitamente fornite da una TC: in alto a sinistra la ricostruzione volumetrica, a destra il piano assiale, in basso i piani sagittale e frontale.</i>	35
2.16	<i>TAC di un cervello umano, dalla base del cranio fino in cima. Eseguita con mezzo di contrasto intravenoso.</i>	36
2.17	<i>Schema di un processo di acquisizione PET.</i>	37
2.18	<i>Tipica apparecchiatura PET.</i>	38
2.19	<i>PET cerebrale.</i>	38

2.20	<i>Una apparecchiatura medica per la risonanza magnetica.</i>	39
2.21	<i>Immagine sagittale di un ginocchio ottenuta mediante risonanza magnetica.</i>	40
2.22	<i>Momento di spin.</i>	41
2.23	<i>Effetto di un campo magnetico statico <math>\vec{B}_0</math>.</i>	43
2.24	<i>Analogia meccanica, moto di precessione.</i>	44
2.25	<i>Momento angolare di una particella di spin <math>1/2</math>.</i>	45
2.26	<i>Precessione di Larmor.</i>	46
2.27	<i>Effetto dell'eccitazione a radiofrequenza, in a) si evidenzia l'inclinazione di <math>\vec{M}</math> dovuta all'impulso, nelle successive il vettore compie un movimento a spirale ritornando al punto di equilibrio.</i>	48
2.28	<i>Andamento componente longitudinale e trasversale di <math>\vec{M}</math>.</i>	49
2.29	<i>Segnale Free Induction Decay (FID).</i>	50
2.30	<i>Sequenza di eccitazione spin-echo.</i>	53
3.1	<i>Spettro elettromagnetico ordinato per frequenza.</i>	56
3.2	<i>Istogramma dell'intensità che può essere partizionato da una sola soglia.</i>	60
3.3	<i>Rappresentazione di un'immagine, suo profilo di intensità e derivate prima e seconda.</i>	71
3.4	<i>Rappresentazione del negativo del LoG.</i>	78
3.5	<i>Significato Watershed.</i>	83
3.6	<i>Esempio di procedimento watershed.</i>	84
3.7	<i>Illustrazione del metodo Level-Set.</i>	87
3.8	<i>Punti della griglia in Fast Marching.</i>	96
3.9	<i>Evoluzione del Fast Marching.</i>	96
4.1	<i>Schermata di apertura.</i>	101
4.2	<i>Tecniche di base.</i>	101
4.3	<i>Interfaccia Otsu.</i>	102
4.4	<i>Interfaccia Otsu multilivello/multisoglia.</i>	103
4.5	<i>Immagine di partenza della segmentazione di Fig.4.6.</i>	104

4.6	<i>Otsu con diversi livelli di segmentazione, rispettivamente 2, 4, 7 e 20.</i>	105
4.7	<i>Selezione dei diversi livelli con la tecnica di Otsu multisoglia.</i>	105
4.8	<i>Interfaccia Logic Operations.</i>	106
4.9	<i>Immagine con ROI segmentata in Figura 4.10.</i>	107
4.10	<i>Segmentazione con Logic Operations - Sobel.</i>	107
4.11	<i>Segmentazione Watershed.</i>	108
4.12	<i>Interfaccia Region Growing.</i>	108
4.13	<i>Interfaccia tecniche avanzate.</i>	109
4.14	<i>Segmentazione con Chan-Vese.</i>	110
4.15	<i>Segmentazione con Geodesic.</i>	110
4.16	<i>Segmentazione con Fast Marching.</i>	112
4.17	<i>Sovrapposizione della segmentazione di Fig. 4.16 con l'immagine originale.</i>	112
4.18	<i>Segmentazione con Snakes.</i>	113
4.19	<i>Interfaccia per il confronto.</i>	116
5.1	<i>Confronto tra Otsu multilivello e Ground Truth di tessuto adiposo cardiaco su un'immagine CT.</i>	130
5.2	<i>Confronto tra Ground Truth (sinistra) e Geodesic (destra) di una MRI rappresentante un Tumore al cervello.</i>	130
5.3	<i>Confronto tra Fast Marching (sinistra) e Geodesic (destra) di un'immagine CT di un ictus al cervello.</i>	131



# Capitolo 1

## Introduzione

**N**ell'era della digitalizzazione è sempre più importante la gestione e la manipolazione delle immagini digitali. Questa importanza è direttamente legata alle applicazioni, che possono spaziare in vari campi, dall'intrattenimento (e.g. nelle moderne tv), all'editing e il fototocco, o ancora in ambito biomedicale. Quest'ultimo è un campo di applicazione importante dell'elaborazione delle immagini.

Le bioimmagini sono delle immagini generate da un essere vivente, cioè una rappresentazione dell'anatomia o della fisiologia delle parti interne del corpo e possono essere generate in diversi modi, ma sempre facendo interagire una qualche forma di energia (vibrazioni acustiche, raggi X, campi elettromagnetici, etc...) con la struttura da analizzare, generando così, a seguito di elaborazioni, un'immagine 2D o 3D. Un'elaborazione che riveste particolare interesse è la segmentazione, in quanto permette di dividere l'intera immagine in diverse regioni ed isolare quelle di maggiore interesse.

Il presente lavoro ha lo scopo di progettare e sviluppare un software di segmentazione automatica con interfaccia grafica, il quale permette di acquisire immagini (anche in formato DICOM) ed effettuarne la segmentazione utilizzando diverse tecniche disponibili in letteratura.

Inoltre vi è la possibilità di confrontare tra loro diversi metodi o confrontare un metodo con una ground truth. Questo confronto avviene sia visivamente che attraverso dei parametri classici come ad esempio la differenza di area, l'accuratezza, coefficiente di Dice.

Di seguito verranno descritte le maggiori tecniche di acquisizione di bio-immagini e la loro importanza dal punto di vista clinico. Dopo dei cenni introduttivi all'elaborazione delle immagini, verranno illustrate le tecniche utilizzate nel segmentatore automatico. Poi si passerà alla descrizione del software dal punto di vista dell'interfaccia e dell'utilizzo da parte dell'utente, quindi alla sua implementazione e verranno illustrati i parametri di confronto utilizzati e il loro calcolo. Infine saranno presentati dei risultati di comparazione tra diversi metodi e ground truth.

## Capitolo 2

### Bioimmagini

**C**on il termine immagine ci si riferisce alla rappresentazione di un oggetto che appartiene al mondo reale e che può avere diverse caratteristiche, come ad esempio un paesaggio o un volto umano. Le immagini hanno una notevole efficacia per un osservatore umano e sono un importante mezzo per trasmettere informazioni e conoscenza. Infatti l'utilità delle tecniche di *imaging* in medicina è un esempio eclatante dell'importanza delle immagini.

Il desiderio dell'uomo di conoscere la morfologia e i meccanismi interni degli esseri viventi risale a tempi molto antichi, quando venivano studiate le forme di organi interni di persone e animali. Già Leonardo da Vinci, grazie ai suoi studi e dissezioni su cadaveri, era in grado di disegnare parti interne del corpo umano con un dettaglio degno delle conoscenze moderne. Tuttavia la conoscenza morfologica non andava di pari passo con la conoscenza fisiologica.

Le bioimmagini hanno caratteristiche e peculiarità che le distinguono da altre forme di rappresentazione come ad esempio i segnali temporali. Il mondo osservato possiede un alto grado di regolarità che si riflette nell'immagine in una forte *ridondanza*, questo si traduce nella possibilità di

utilizzare dei processi di compressione per compattare le rappresentazioni permettendo così di diminuire la quantità di dati utilizzabili per codificare un'immagine senza perdere informazione. Ovviamente la compressione, vista l'incognita sull'informazione utile, deve essere reversibile consentendo la ricostruzione dell'immagine nella sua condizione originaria. Questa elevata quantità di informazione contenuta nelle bioimmagini costituisce uno degli elementi di specificità<sup>2.1</sup> rispetto ai segnali temporali. Questo discorso si riflette in maniera analoga anche in natura, basti pensare che l'occhio umano ha 100 milioni di fotorecettori, tra coni e bastoncelli, quindi una grande quantità di informazioni (*ridondanza*) e che il cervello ha un'ampia area dedicata alla ricezione e alla interpretazione delle immagini, la corteccia visiva, che è la regione più studiata.

Dal punto di vista fisico le immagini sono il risultato dell'interazione tra l'energia prodotta da una sorgente e la materia costituente l'oggetto o la scena da studiare, questo risultato viene catturato da dei sensori. L'interazione è associata ad una trasformazione che rappresenta lo spazio tridimensionale in una superficie bidimensionale. Le immagini da sole non bastano, per la fruizione e l'estrapolazione delle immagini sono fondamentali due fattori, la *percezione* e l'*interpretazione*. La prima riguarda le capacità visive del soggetto che osserva l'immagine, mentre la seconda implica delle azioni cognitive frutto di una conoscenza soggettiva ed individuale acquisita nel corso delle proprie esperienze.

L'interpretazione umana delle immagini tende a raggruppare gli oggetti simili e vicini dandoci l'idea di una sola entità, la percezione visiva ci permette di distinguere gli oggetti dallo sfondo e collocarli mentalmente nello spazio

---

<sup>2.1</sup>Capacità di un test di dare un risultato normale ("negativo") nei soggetti sani, ovvero la specificità del test verrà calcolata come i veri negativi diviso il totale dei sani:  $S_p = \frac{V^-}{T S^+} = \frac{V^-}{V^- + F^+}$ .

tridimensionale. Però anche il sistema di vista umano commette spesso degli errori di interpretazione (e.g. nelle illusioni ottiche). L'occhio umano nonostante abbia un range dinamico molto esteso ( $10^6 - 10^8$ ), è in grado di distinguere un numero limitato di livelli di grigio; invece è in grado di percepire e classificare correttamente un elevato numero di colore, questa particolarità ha una scarsa importanza in termini di immagini mediche perchè esse sono in generale prive di colori. In certi casi, come ad esempio nell'ecocardiografia, l'osservazione di immagini in movimento rende più agevole l'interpretazione da parte dell'osservatore. Le immagini sono sempre affette da rumore quindi ogni sistema di trattamento deve ottenere le informazioni utili nascoste dal rumore. Nel caso di immagini fotoniche si ottiene maggiore informazione aumentando il numero di fotoni, ma questo spesso produce degli effetti collaterali (radiazioni ionizzanti) non trascurabili che possono mettere a rischio la salute del paziente. Di contro nella diagnostica per immagini i dettagli sono molto importanti, per questo motivo il campo è in continua evoluzione cercando di ottenere il maggior numero di dettagli con il minor numero di esposizione a rischi. Il corpo umano è costituito da tessuti e organi i quali sono composti principalmente da acqua, sostanze organiche, calcio e gas. Questi componenti, grazie alle diverse caratteristiche come densità e peso specifico, sono alla base delle tecniche di visualizzazione in ambito medico. Come già anticipato le immagini scaturiscono dall'interazione tra energia e materia, la classificazione delle diverse tecniche di *imaging* può, ad esempio, avvenire proprio in base alla natura dell'energia coinvolta nel procedimento, di seguito verranno esposte le tecniche più importanti ed usate.

## 2.1 Immagini ad ultrasuoni

Il successo delle tecniche a ultrasuoni è dovuto al fatto che sono totalmente innoque per il paziente, hanno un'ottima risoluzione temporale e il funzionamento basato sulla diffusione e riflessione delle onde all'interfaccia acustica (differenti tessuti). Gli *ultrasuoni* sono delle onde vibrazionali che hanno una frequenza che va oltre la capacità uditiva umana (maggiore di  $16 \div 20\text{kHz}$ ), e per propagarsi hanno bisogno di un mezzo, proprio le caratteristiche del mezzo di propagazione permettono di ricavare delle informazioni sulla regione osservata. In questo caso, quindi, l'energia presa in considerazione è quella meccanica posseduta dalle onde ultrasoniche. La tecnica ultrasonografica è resa possibile dalla limitata velocità delle onde acustiche, che permette di misurarne i tempi di propagazione. La velocità di propagazione dipende dalla natura del mezzo, se  $\lambda$  è la lunghezza d'onda,  $v$  la velocità di propagazione del suono nel mezzo e  $f$  la frequenza dell'onda, risulta:

$$v = \lambda \cdot f \Rightarrow f = \frac{v}{\lambda}.$$

Nei mezzi biologici si considera una velocità di propagazione circa costante, vicino a quella nell'acqua  $1500\frac{m}{s}$ . Fa eccezione però il tessuto osseo nel quale la velocità di propagazione è sensibilmente maggiore  $4000\frac{m}{s}$ . Per convenzione si assume un valore medio della velocità di propagazione per il tessuto molle, indipendentemente dalla frequenza dell'onda ultrasonica, questo valore è di  $1540\frac{m}{s}$  corrispondente a circa  $1.5\frac{mm}{\mu s}$ . Avendo l'esigenza, nella diagnostica medica, di una risoluzione dell'ordine del millimetro o meno, viene utilizzata una frequenza che va da  $1\text{MHz}$  a  $20\text{MHz}$ , dove

si ha che

$$f = 1MHz, \quad \lambda = 1.5mm; \quad f = 20MHz, \quad \lambda = 0.077mm.$$

L'impedenza acustica è caratteristica del mezzo ed è data da

$$Z = \rho v,$$

dove  $\rho$  è la densità del mezzo e  $v$  è la velocità dell'onda ultrasonica. L'impedenza acustica dipende quindi proporzionalmente dalla densità del mezzo. In Figura 2.1 si ha una rappresentazione di un'onda sonora.

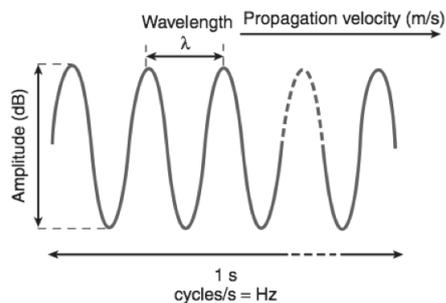


Figura 2.1: *Schema di un'onda ultrasonica.*

### 2.1.1 Generazione degli ultrasuoni

Gli ultrasuoni vengono generati ed emessi attraverso l'effetto piezoelettrico, ovvero la proprietà che hanno alcuni materiali di generare una tensione elettrica proporzionale alla pressione esercitata su di essi e viceversa di modificare le loro dimensioni, generando onde di pressione, a seguito dell'applicazione di una tensione. Consideriamo il caso di un singolo trasduttore costituito da una piastrina di ceramica piezoelettrica eccitata da due elettrodi posti sulle facce parallele. La velocità dell'onda ultrasonica nel piezo

è circa  $4000 \frac{m}{s}$  e verificandosi il fenomeno della risonanza per un valore  $h$  dello spessore pari a  $\frac{\lambda}{2}$  si ha, per una frequenza  $f = 5MHz$

$$h = \frac{\lambda}{2} = \frac{v}{2f} = \frac{4 \times 10^6}{10 \times 10^6} = 0.4mm.$$

Quindi la frequenza espressa in  $MHz$  degli ultrasuoni generati per effetto piezoelettrico da una piastrina dello spessore di  $h = \frac{\lambda}{2}$  millimetri, è  $f = \frac{2}{h}$ . Il trasduttore può funzionare da *generatore* di ultrasuoni quando alimentato da una tensione ad una frequenza opportuna oppure da *ricevitore* convertendo la sollecitazione meccanica in segnale elettrico.

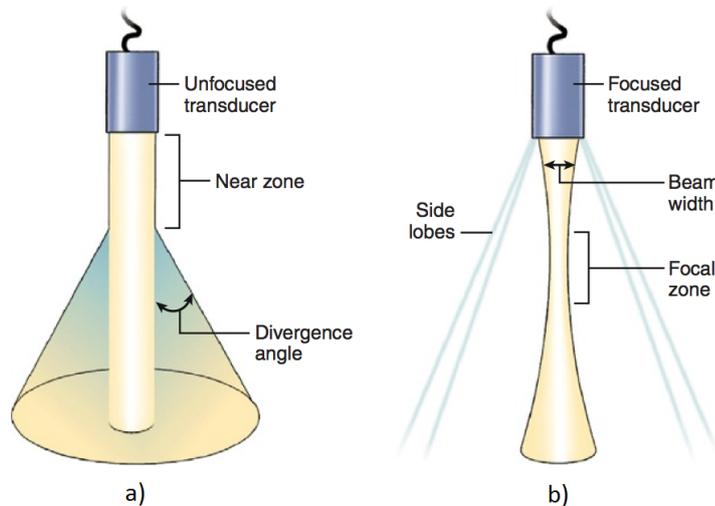


Figura 2.2: *Rappresentazione di un fascio ultrasonico: a) Non focalizzato, b) focalizzato.*

La forma tipica di un fascio ultrasonico si può osservare in Figura 2.2. Nel caso di un fascio non focalizzato (Figura 2.2a) si hanno due zone, quella prossimale caratterizzata da un diametro circa costante e la zona distale dove il fascio diverge. Per concentrare maggiormente l'energia in un ben preciso punto si usa la focalizzazione statica ottenuta con una lente posta nella

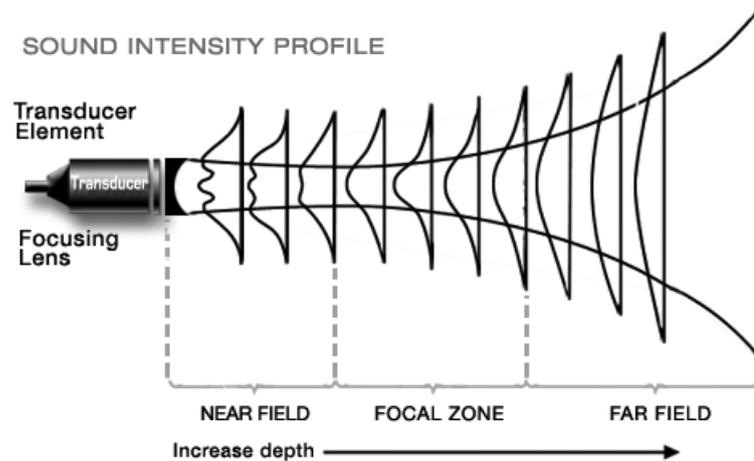


Figura 2.3: Profilo dell'intensità di un fascio ultrasonico.

parte frontale del trasduttore, Figura 2.2b. Mentre in Figura 2.3 si ha una rappresentazione del profilo di intensità di un fascio ultrasonico focalizzato.

### 2.1.2 Propagazione degli ultrasuoni

Per la propagazione di ultrasuoni valgono le stesse leggi che descrivono il comportamento dei suoni udibili, quindi si avranno dei fenomeni di assorbimento, riflessione e rifrazione, in Figura 2.4 si può osservare una rappresentazione della interazione tra ultrasuoni e materia. In un mezzo omogeneo un'onda ultrasonica viaggia in linea retta. Il trasduttore viene appoggiato alla pelle interponendo un gel che assicura la corretta trasmissione delle onde, quando il fronte d'onda incontra un cambiamento di materiale, *interfaccia*, dove cambia il valore dell'impedenza acustica, da  $Z_1$  a  $Z_2$ . In questo passaggio parte dell'energia viene trasmessa, proseguendo ed interagendo con altre interfacce e parte viene riflessa o diffusa a causa delle irregolarità. Possiamo descrivere il fenomeno attraverso la legge di *Snell*, Figura 2.5, facendo riferimento con il pedice  $i$  all'onda incidente,  $r$  riflessa e  $t$  trasmessa,

si ha:

$$\frac{\sin \theta_i}{\sin \theta_t} = \frac{v_1}{v_2},$$

mentre  $v_1$  e  $v_2$  sono rispettivamente le velocità di propagazione delle onde ultrasoniche nel mezzo 1 e nel mezzo 2; si ha che  $\theta_i = \theta_r$ . Ovviamente

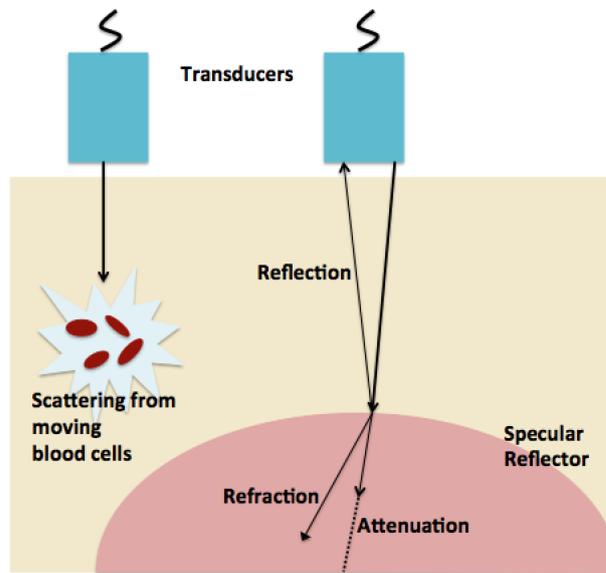


Figura 2.4: *Interazione tra ultrasuoni e materia.*

se le velocità nei due mezzi sono uguali non si ha nessuna rifrazione (deviazione) e gli ultrasuoni proseguono indisturbati. Se l'impedenza acustica dei mezzi è la stessa non si ha riflessione, tutta l'energia viene trasmessa al secondo mezzo. Quindi se le impedenze sono diverse una parte dell'energia viene riflessa. Il coefficiente di riflessione  $R$  si esprime come

$$R = \left( \frac{Z_1 \cos \theta_t - Z_2 \cos \theta_i}{Z_1 \cos \theta_t + Z_2 \cos \theta_i} \right)^2,$$

che nel caso dell'incidenza normale, in cui l'angolo di incidenza e l'angolo

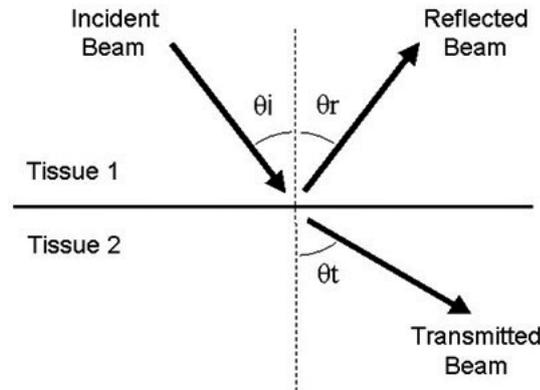


Figura 2.5: Legge di Snell per gli ultrasuoni.

di trasmissione sono zero, diventa

$$R = \left( \frac{Z_1 - Z_2}{Z_1 + Z_2} \right)^2 = \left( \frac{1 - \frac{Z_2}{Z_1}}{1 + \frac{Z_2}{Z_1}} \right)^2,$$

con  $0 \leq R \leq 1$ . Invece il coefficiente di trasmissione è dato da

$$T = 1 - R.$$

Nel caso in cui  $Z_1 \gg Z_2$  o  $Z_2 \gg Z_1$ , come nel caso dell'interfaccia osso-muscolo, si ha che tutta l'energia incidente viene riflessa. Dal momento che non si riesce ad identificare particolari morfologici distanti tra loro meno della lunghezza d'onda degli ultrasuoni utilizzati, si cerca di adoperare frequenze più alte possibile, questo va a scontrarsi però con il coefficiente di attenuazione che cresce proporzionalmente con la frequenza. Ad esempio in cardiologia si usano frequenze di  $2.5 \div 5 MHz$ , con una profondità di penetrazione dei tessuti molli di  $12 - 6 cm$ . Invece in oftalmologia si usano frequenze più elevate, di  $15 \div 20 MHz$ , ma con limitata penetrazione. Tra le tecniche di imaging ad ultrasuoni la più diffusa è sicuramente la tecni-

ca ad *eco-impulsi* (ecografia), ma anche la tecnica *eco-doppler*, che sfrutta l'effetto Doppler, è molto usata.

### 2.1.3 Tecnica ad eco-impulsi, scansione e modi di visualizzazione

Un classico sistema ecografico è formato da un cristallo piezoelettrico pilotato, un'amplificatore di segnale e un display. Il cristallo emette impulsi e si pone in ricezione per percepire l'eco generato dalle interfacce acustiche, il tempo che intercorre dall'emissione dell'impulso al rilevamento dell'eco viene usato per misurare la profondità a cui si trova la discontinuità che ha generato l'eco. L'ampiezza attenuata del segnale di eco indica la differenza di impedenza acustica tra i due mezzi. La risoluzione spaziale di uno strumento siffatto viene divisa in tre tipi, Figura 2.6, *assiale* lungo la direzione longitudinale di propagazione, *laterale* lungo la direzione perpendicolare a quella di propagazione e infine la risoluzione di *elevazione*. Solitamente viene generato un pacchetto d'onde dal trasduttore e subito

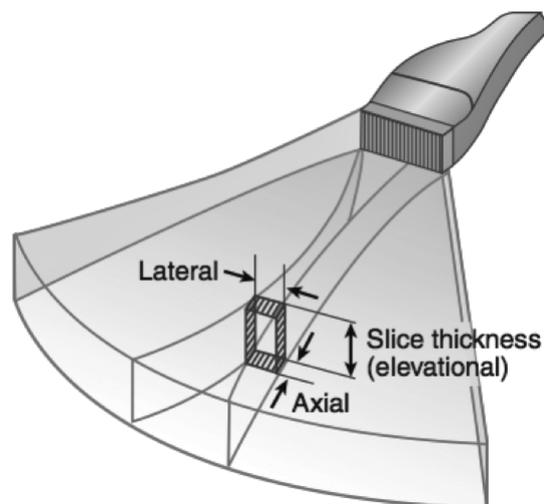


Figura 2.6: *Risoluzione spaziale di un fascio di ultrasuoni.*

dopo va in ricezione, la durata del pacchetto viene resa il più breve possibile per migliorare la risoluzione assiale. Questo però comporta un aumento della frequenza delle onde del pacchetto che incide sulla profondità di penetrazione, bisogna quindi trovare un trade-off. Successivamente il segnale eco viene amplificato con un guadagno proporzionale al ritardo per compensare l'attenuazione per assorbimento. Infatti un segnale che proviene da maggiore distanza ha più necessità di essere amplificato rispetto ad un segnale generato da un'interfaccia prossima al trasduttore. La scansione delle strutture può essere ottenuta con un trasduttore singolo (a movimento meccanico) o un trasduttore multiplo, che a sua volta può essere *linear-array* o *phased-array*.

Per quanto riguarda i modi di visualizzazione dell'informazione si suddividono in:

- Modo A (*Amplitude*): la posizione e l'ampiezza sullo schermo rappresentano rispettivamente la distanza dell'interfaccia dal trasduttore e l'intensità della riflessione, questo modo viene utilizzato ad esempio per esami semplici come la valutazione delle dimensioni della testa di un feto;
- Modo B (*Brightness*): il segnale eco va a modulare l'intensità di un monitor in funzione della distanza, quindi maggiore è l'ampiezza del segnale riflesso e maggiore luminosità avrà il punto sullo schermo;
- Modo M (*Time Motion*): usato nello studio delle strutture cardiache, atri, ventricoli, valvole. Si ottiene affiancando linee di modo B riprese in rapida successione temporale.

I flussimetri basati sull'effetto Doppler consentono di effettuare una misura non invasiva (dall'esterno) della velocità del sangue nei vasi. Il ruolo di riflettori in questo caso è giocato dai globuli rossi, i quali per effetto Doppler provocano una variazione della frequenza  $f_0$  della portante

$$\delta f = f_D - f_0 = \frac{2f_0\nu_0 \cos \theta}{v},$$

dove  $f_D$  è la frequenza modificata,  $\nu_0$  la velocità del sangue,  $v$  la velocità degli ultrasuoni,  $\theta$  l'angolo fra il flusso e la direzione di propagazione dell'onda ultrasonica. In caso di stenosi del vaso si ha un aumento della velocità del flusso in quella sede e un flusso turbolento a valle. In generale l'esame Doppler consente di individuare la presenza di restringimenti od occlusioni a carico dei vasi arteriosi periferici.

Il metodo *Eco-color-Doppler* associa le informazioni morfologiche date da un'ecografia ad alta risoluzione ad informazioni emodinamiche del flusso sanguigno ottenute sfruttando l'effetto Doppler, la visualizzazione a colori permette una più facile identificazione dei vasi e dell'eventuali patologie ad esso legate.

L'ecografia può essere utilmente impiegata nello studio di numerosi organi permettendo di rilevare alterazioni strutturali e quindi malattie. Sono molto importanti le applicazioni in cardiologia ed ostetricia. Uno dei limiti è che non è indicata per lo studio di determinati distretti circondati da osso oppure aria, come ad esempio la teca cranica e i polmoni. Tuttavia grazie alle basse potenze in gioco, l'ecografia garantisce una completa innocuità dell'esame, che pertanto può essere eseguito anche in casi particolari come la gravidanza. In Figura [fig:Ecografia-di-un-cuore] è mostrata una ecografia del cuore effettuata utilizzando tecniche che sfruttano l'acustica non

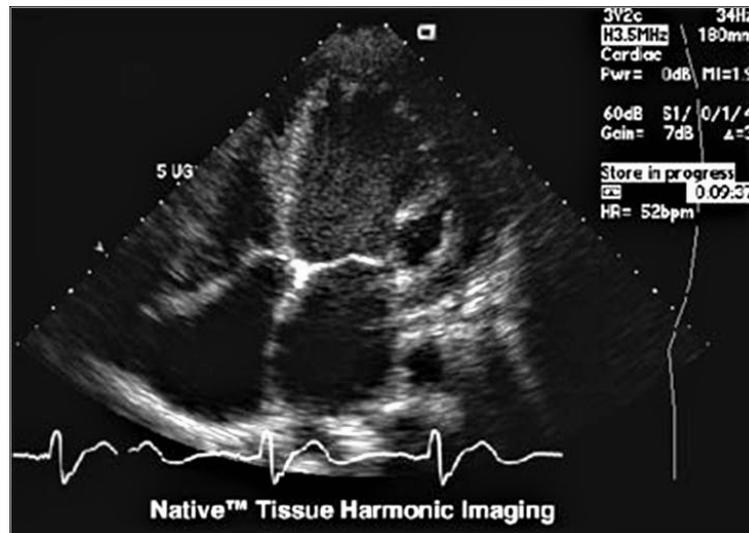


Figura 2.7: *Ecografia di un cuore realizzata con tecniche che sfruttano l'acustica non lineari.*

lineare, chiamata anche *Tissue Harmonic Imaging*.

## 2.2 Radiazioni Ionizzanti

Prima di presentare le tecniche di imaging basate sull'utilizzo dell'energia elettromagnetica è bene fare una panoramica sulle radiazioni ionizzanti, che costituiscono la sorgente più importante per ottenere immagini in medicina. In Figura 2.8 è mostrata un'interessante rappresentazione dello spettro elettromagnetico. Per *radiazione* si intende la propagazione di energia attraverso lo spazio sotto forma di onde o energia cinetica propria di alcune particelle. Il termine *ionizzante* invece indica la capacità di queste radiazioni di rompere i legami atomici e molecolari della materia con cui interagiscono, andando a strappare degli elettroni, modificandone lo stato chimico. In breve le radiazioni ionizzanti sono radiazioni aventi un'energia tale da ionizzare gli atomi o le molecole con cui interagiscono; le energie di soglia

dei processi di ionizzazione sono dell'ordine di alcuni  $eV$ . Le radiazioni ionizzanti quando incontrano un mezzo materiale possono essere assorbite e quindi cedere la loro energia alle molecole costituenti il materiale, provocando l'emissione di un elettrone orbitale di un atomo, formando una coppia di ioni di carica opposta. Le radiazioni corpuscolari, ad esempio  $\alpha$  e  $\beta$ , hanno una scarsa capacità di penetrazione e vengono assorbite da sottili strati di materiale. Invece i raggi  $X$  e i raggi gamma riescono a penetrare materiali come l'acqua o i muscoli e sono fortemente attenuati da altri, quali l'osso o il piombo. Questa diversa misura di penetrazione viene sfruttata per ottenere le immagini radiografiche.

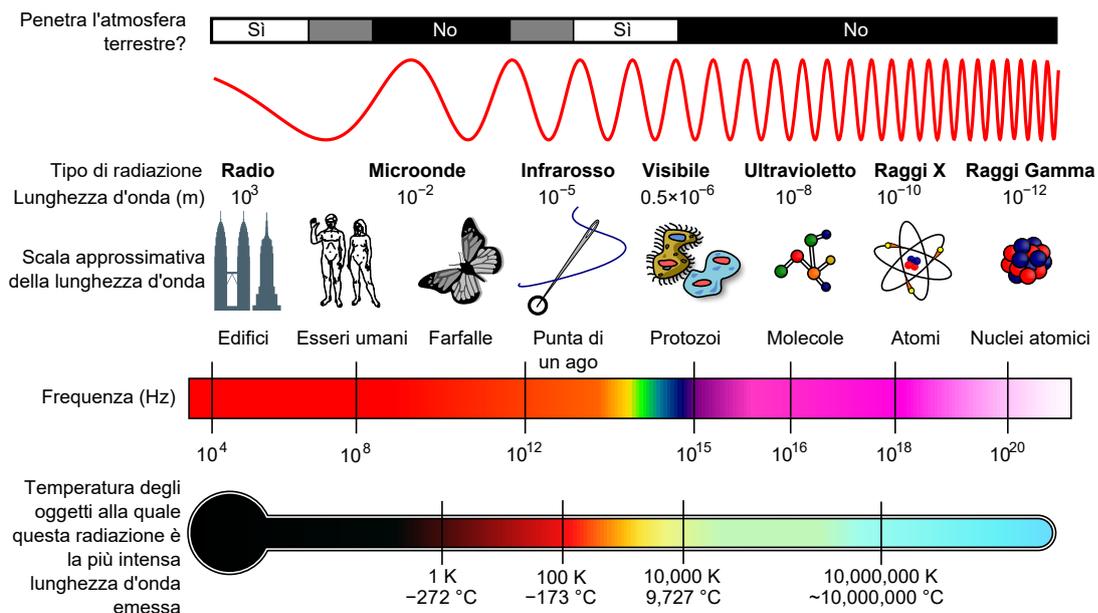


Figura 2.8: *Spettro elettromagnetico.*

### 2.2.1 Sorgenti

Oltre ad essere prodotte in natura da materiali radioattivi (uranio, polonio e radio), le radiazioni ionizzanti, possono essere generate artificialmente da radioisotopi o apparecchiature radiogene. I radioisotopi sono atomi che emettono radiazioni e hanno un nucleo instabile dovuto ad uno sbilanciamento del rapporto neutroni-protoni. Il termine *radioattività* indica proprio questo fenomeno. In generale un radionuclide emettendo radiazione si trasforma in un atomo stabile, non producendo più radiazione. La trasformazione radioattiva non avviene allo stesso istante per tutti gli atomi ma in maniera da far rimanere costante la percentuale di radionuclidi che si trasformano nell'intervallo di tempo. Il numero  $N$  di nuclei radioattivi al tempo  $t$  è dato da:

$$N = N_0 e^{-\lambda t},$$

dove  $N_0$  è il numero di nuclei al tempo  $t = 0$  e  $\lambda$  è la *costante di decadimento* del nuclide. Con il passare del tempo la quantità di materiale radioattivo diminuisce, quindi cala anche la quantità di radiazione emessa, finché non diventa trascurabile. Il tempo necessario a dimezzare la radioattività è detto *tempo di dimezzamento* o *emivita*, ed è dato da

$$T_{\frac{1}{2}} = \frac{\ln 2}{\lambda} \approx \frac{0.693}{\lambda}.$$

L'emivita varia molto tra gli elementi, ad esempio il Carbonio 14 ( $^{14}\text{C}$ ) ha un tempo di dimezzamento di un migliaio di anni, mentre l'isotopo artificiale Elio 3 ( $^3\text{He}$ ) di  $10^{-20}$  secondi.

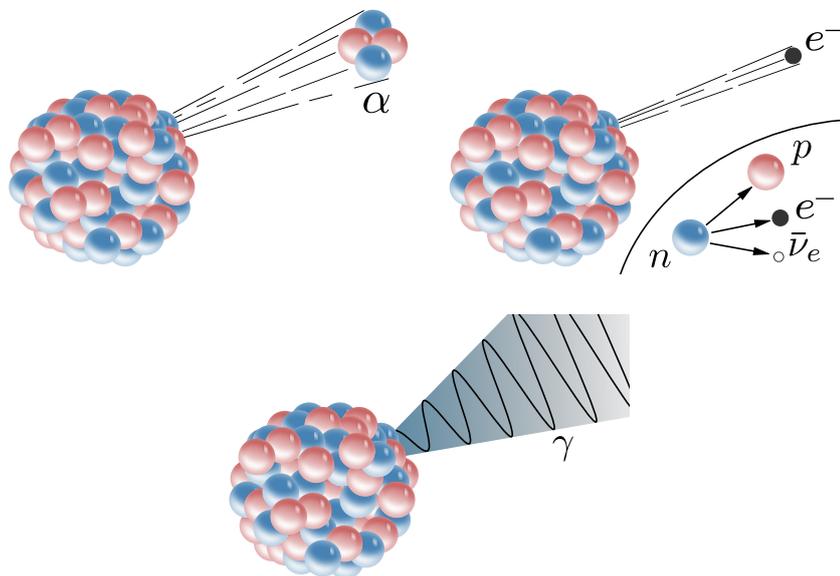


Figura 2.9: *Decadimento  $\alpha$ ,  $\beta^-$  e raggi  $\gamma$ .*

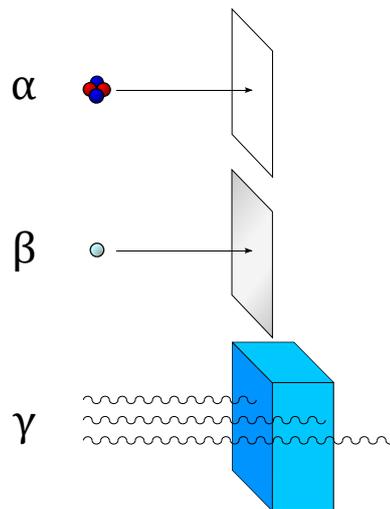


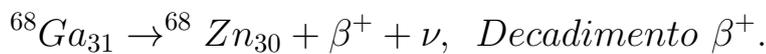
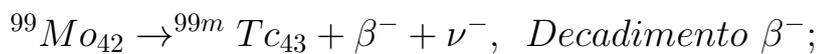
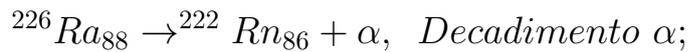
Figura 2.10: *Potere penetrante dei diversi tipi di radiazioni ionizzanti.*

Il *decadimento radioattivo* è accompagnato dall'emissione di particelle  $\alpha$ ,  $\beta$  e raggi  $\gamma$ . In Figura 2.9 si può apprezzare una rappresentazione dei decadimenti radioattivi, e in Figura 2.10 è rappresentato<sup>2.2</sup> il potere penetrante

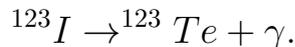
<sup>2.2</sup>Fonte:[[https://it.wikipedia.org/wiki/File:Alfa\\_beta\\_gamma\\_radiation.svg](https://it.wikipedia.org/wiki/File:Alfa_beta_gamma_radiation.svg)]

delle radiazioni ionizzanti. Nel *decadimento*  $\alpha$  vengono emesse particelle  $\alpha$ , ovvero 2 protoni + 2 neutroni (i cosiddetti elioni). In seguito a questa emissione il peso atomico diminuisce di 4 e il numero atomico di 2. Le particelle  $\alpha$  non sono adatte alla generazione di immagini un quanto hanno uno scarso potere penetrante, tuttavia vengono utilizzate in radioterapia. Il *decadimento*  $\beta$  si divide in  $\beta^-$  e  $\beta^+$ , nel primo si ha che un neutrone si trasforma in un protone + un elettrone (+ un antineutrino che non interagisce con la materia). Il peso atomico rimane invariato e il numero atomico aumenta di 1, gli elettroni emessi ( $\beta^-$ ) a causa della loro bassa energia non riescono ad uscire dal corpo e quindi non possono essere usati per tecniche di imaging, sono usati soltanto nella autoradiografia su reperti istologici. Nel *decadimento*  $\beta^+$  un protone si trasforma in neutrone + un positrone (+ un neutrino). Il peso atomico è invariato e il numero atomico diminuisce di 1, questo tipo di decadimento avviene soltanto quando la differenza di energia tra nuclide padre e nuclide figlio è  $> 1022keV$ . Quando un positrone ( $\beta^+$ ) dopo essere stato emesso, si propaga in un corpo per circa un millimetro prima di annichilirsi con un elettrone ( $\beta^-$ ), generando due fotoni da  $511keV$  in direzioni quasi opposte diametralmente. Il processo di annichilazione deve rispettare delle leggi di conservazione, della carica, dell'energia, della quantità di moto. La conservazione dell'energia e della quantità di moto impediscono la generazione di un solo fotone anzichè due, questi due fotoni hanno energia uguale all'energia a riposo del positrone e dell'elettrone ( $511keV$  appunto). Se la quantità di moto delle due particelle è inizialmente nulla si ha che i due fotoni vengono emessi collinearmente in versi opposti, viceversa se la quantità di moto iniziale è diversa da zero si ha la quasi collinearità. L'interazione delle particelle  $\beta$  con la materia ha ge-

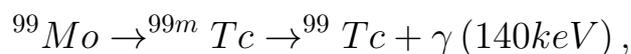
neralmente un raggio d'azione dieci volte superiore, e un potere ionizzante pari a un decimo rispetto all'interazione delle particelle alfa, vengono bloccate completamente da pochi millimetri di alluminio. Di seguito riportiamo alcuni esempi di decadimenti  $\alpha$  e  $\beta$ :



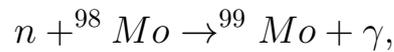
Dove con  $\nu^-$  e  $\nu$  si è indicato rispettivamente l'antineutrino e il neutrino. L'emissione di raggi gamma viene innescata dalla *cattura di un elettrone orbitale*, che viene appunto catturato da nucleo, un protone si trasforma in neutrone + un neutrino con conseguente emissione di raggi  $\gamma$  in seguito al riequilibrio dei livelli elettronici. Un nucleo prodotto da un decadimento radioattivo può rimanere in uno stato eccitato (*isomerico*) per un certo periodo di tempo che può variare da qualche *femto* secondo ad alcune ore prima di decadere ad un livello energetico inferiore emettendo radiazioni  $\gamma$ , questo stato è detto *metastabile*. Questo comportamento è determinante in medicina nucleare. Un esempio di emissione di raggi  $\gamma$  può essere il seguente:



Si possono produrre materiali radioattivi artificialmente facendo interagire sostanze stabili con neutroni o particelle cariche ad alta energia, oppure tramite procedimenti di radiochimica. Ad esempio con la radiochimica:

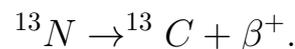
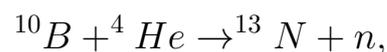


attraverso la cattura di elettroni:



dove i neutroni ( $0.1 < E < 100\text{eV}$ ) vengono ottenuti grazie all'utilizzo di un reattore nucleare.

Mentre nel bombardamento dei nuclei di atomi stabili con particelle cariche, si ha che le particelle devono avere energia cinetica sufficientemente elevata da vincere la repulsione elettrostatica positiva del nucleo; si usano quindi gli acceleratori lineari o ciclotroni, con energie di  $1 - 100\text{MeV}$



Le macchine radiogene invece sono quelle apparecchiature in grado di produrre radiazioni ionizzanti. Mentre i radioisotopi emettono radiazioni continuamente, le macchine radiogene emettono radiazioni soltanto quando vengono attivate. Le macchine radiogene utilizzate in ambito medico sono le macchine a Raggi  $X$ , per la radiodiagnostica, gli acceleratori lineari per la radioterapia e i coclotroni per la produzione di farmaci radioattivi.

### 2.2.2 Dose

Un aspetto molto importante è la valutazione della quantità di radiazione impartita ai pazienti ed al personale e dei relativi effetti biologici che comporta. Questo implica la misura di diverse grandezze che esprimono le capacità ionizzanti della radiazione e dei possibili danni biologici ad essa legati. Il *roentgen* (R) misura le ionizzazioni prodotte in una certa quantità

di aria:  $1 R = 1.61 \times 10^{15}$  ionizzazioni per grammo d'aria, che equivale a  $2.58 \times 10^{-4}$  coulomb per chilogrammo di aria. Il roentgen è stato gradualmente sostituito dal *Coulomb/Kg* di aria  $1 \frac{C}{Kg} = 3876R$ .

La *dose assorbita* ( $D$ ) misura la quantità di energia assorbita da una massa unitaria di tessuto, l'unità di misura in questo caso è il *gray* ( $Gy$ ) pari a  $1 \text{ Joule}/Kg$ , precedentemente si usava il *rad* corrispondente ad un centesimo di  $Gy$  ( $1 Gy = 100 rad$ ). Il rapporto tra dose assorbita ed esposizione dipende dall'energia dei fotoni, per il tessuto muscolare è circa costante mentre per il tessuto osseo varia con l'energia, a basse energia l'osso ha una maggiore esposizione. In ogni caso, a parità di esposizione, la dose assorbita dall'osso è sempre maggiore di quella assorbita dai muscoli.

La *dose equivalente* ( $H$ ) è una grandezza convenzionale che si ottiene moltiplicando la dose assorbita  $D$  per un fattore  $WR$  di ponderazione per la radiazione,  $H = D \cdot WR$  esprime la capacità della radiazione di generare effetti biologici nei tessuti, l'effetto non dipende soltanto dall'energia ceduta ma anche dal tipo di radiazione. In questo caso l'unità di misura è il *sievert* ( $Sv$ ) che ha sostituito il *rem* corrispondente ad un centesimo di sievert. Il fattore  $WR$  per i raggi X, gamma e beta è pari ad 1, quindi per tali radiazioni la dose equivalente coincide con la dose assorbita ( $H = D$ ). Dal momento che diversi tessuti hanno una diversa suscettibilità alle radiazioni, si è avuta la necessità di introdurre un'altra unità di misura, la *dose efficace* ( $DE$ ). Quest'ultima esprime la gravità del danno che un organo può subire, si misura sempre in  $Sv$  e tiene conto anche di fattori ( $WT$ ) di ponderazione specifici per i vari organi, infatti si ha che  $DE = H \cdot WT$ . Nella pratica dei raggi X l'unità utilizzata è il *mAs*, ovvero l'esposizione ottenuta facendo scorrere una corrente di  $1 mA$  nel tubo radiogeno per un tempo di esposi-

zione di 1 *secondo*. La misura della dose viene effettuata con dei *dosimetri*, i quali sono dispositivi che rilevano la quantità di radiazioni alle quali una persona è stata esposta. Ad esempio possiamo avere delle pellicole sensibili ai raggi  $X$ , oppure dei cristalli di litio o di calcio che se esposti a radiazioni avvengono degli spostamenti permanenti degli elettroni in livelli energetici diversi dalla norma.

Giornalmente siamo esposti a radiazioni ionizzanti di cui il 90% è dovuto a cause naturali (radiazione naturale di fondo), come ad esempio la radiazione cosmica, da materiali che ci circondano, da radioisotopi presenti all'interno del nostro corpo, il restante contributo del 10% è dovuto ad attività umane. La radiazione naturale di fondo a cui una persona è esposta<sup>2.3</sup> è quantificata in circa  $1 \div 3.5 mSv/anno$ <sup>2.4</sup>. L'esposizione medica costituisce la principale causa di fonte di irraggiamento dovuto ad attività umane, la dose efficace nei paesi più avanzati, valore medio pro capite, è di  $1 mSv/anno$  per attività radiodiagnostica di cui il 20% dovuto alla tomografia computerizzata,  $0.09 mSv/anno$  per attività di medicina nucleare e  $0.73 mSv/anno$  per attività di radioterapia. Facendo una media a livello mondiale, tali valori si abbasserebbero ancora di più. A titolo di esempio si tenga conto che, una comune radiografia al torace comporta per il paziente una dose di circa  $0.02 mSv$ , mentre una all'addome o una mammografia comportano dosi inferiori a  $1 mSv$  ( $0,4 - 0,7 mSv$ ). Una *TAC* addominale  $8 mSv$  mentre una *PET* o una scintigrafia dai 10 ai 20 *milliSievert*. In radioterapia si forniscono invece dosi molto più massicce di radiazioni, ma concentrate limitatamente

---

<sup>2.3</sup>La media mondiale della dose equivalente di radioattività assorbita da un essere umano e dovuta al fondo naturale è di  $2,4 mSv/anno$ , in Italia circa  $3 mSv/anno$ .

<sup>2.4</sup>Esistono aree geografiche dove il fondo naturale è significativamente più alto della media mondiale, in india è di  $13 mSv/anno$ . Addirittura a Ramsar vi sono sorgenti termali dove la radioattività dell'acqua è di  $260 mSv/anno$  ovvero  $0.03 mSv/ora$ .

ed esclusivamente sul tumore da distruggere. Ad esempio per cancro alla gola vengono somministrati  $2Gy$  a seduta per 30 sedute. In totale  $60Gy$  circa pari a  $60Sv^{2.5}$ .

### 2.2.3 Effetti delle Radiazioni Ionizzanti

Le radiazioni ionizzanti passando attraverso le cellule ionizzeranno atomi e molecole della struttura cellulare, causando la rottura dei legami. Questo produrrà ioni e radicali liberi, ovvero atomi neutri dal punto di vista della carica ma con forte tendenza a formare legami con altri atomi o molecole, particolarmente con molecole di acqua. Questi ioni e radicali liberi interagiscono con altri atomi provocando danni che alcune volte sono riparabili ma in altri casi portano alla morte delle cellule. I radicali liberi infatti agiscono sulle altre molecole che si trovano nella soluzione acquosa, creando composti chimici non presenti prima. Molti di questi nuovi composti chimici sono biologicamente dannosi (o generano una conseguenza dannosa). Si verificano quindi nella cellula modifiche che possono determinarne la morte o il danneggiamento. I danni causati possono essere riparati dalle cellule in modo corretto, quindi senza conseguenze, oppure erroneamente con la conseguenza di possibili difficoltà durante la duplicazione cellulare. Queste difficoltà di duplicazione possono riflettersi in un danno per l'organo di cui la cellula fa parte e quindi per l'intero organismo. Gli effetti delle radiazioni dipendono da quanta energia viene ceduta all'organismo. In Tabella 2.2.1 si ha un'indicazione degli effetti biologici delle radiazioni ionizzanti sull'organismo. Dal punto di vista del tempo in cui si manifestano possiamo distinguere due tipi di effetti, a lungo termine ed acuti. Gli effetti a

---

<sup>2.5</sup>Fonte: <https://it.wikipedia.org/wiki/Sievert>

lungo termine si manifestano anche dopo anni dall'esposizione o addirittura in generazioni successive, dovute a mutazioni somatiche a seguito di danni al DNA cellulare. Se la cellula modificata, di tipo somatico, è ancora in grado di riprodursi può dar luogo ad una patologia neoplastica maligna, per questi tipi di effetti non esiste un valore di soglia della dose assorbita. Se invece la cellula modificata è di tipo germinale (spermatozoi o cellule uovo), le radiazioni ionizzanti possono provocare degli effetti che interessano non soltanto il soggetto esposto, ma anche i suoi figli. Questi effetti conseguono ad un danno provocato dalle radiazioni ionizzanti sul DNA delle cellule germinali o all'irradiazione del prodotto del concepimento durante la vita uterina. Gli effetti acuti invece si presentano in seguito alla morte di cellule di tessuti in cui, in genere, proliferano rapidamente. Questi si manifestano nell'arco di 2-3 settimane dall'esposizione, o più rapidamente nel caso di danno al midollo osseo (febbre, emorragie, infezioni), al sistema gastrointestinale (nausea, vomito, diarrea, dolore addominale) oppure al sistema nervoso centrale (encefalite, convulsioni, letargia). In quest'ultimo caso i sintomi si manifestano nel giro di ore o addirittura minuti. Possono presentarsi anche casi di invecchiamento accelerato.

Dose equivalente	Effetti biologici
$1Sv$	Alterazioni temporanee dell'emoglobina
$2 \div 3Sv$	Nausea, perdita di capelli, emorragie
$4Sv$	Morte nel 50% dei casi
$> 5 \div 6Sv$	Sopravvivenza improbabile

Tabella 2.2.1: *Effetti delle radiazioni ionizzanti sulla salute.*

### 2.2.4 Radioprotezione

L'impiego di radiazioni ionizzanti può essere dannoso per gli operatori, quindi si ha la necessità di proteggere tali soggetti attraverso le cosiddette *norme di radioprotezione*. Particolare preoccupazione suscita l'esecuzione di un esame radiologico su una donna in gravidanza, in Tabella 2.2.2 sono indicate le implicazioni di un irraggiamento del feto in base al periodo di gestazione. Nell'uomo questi effetti sono deterministici e hanno una soglia che comunque è molto maggiore della dose assorbita in radiodiagnostica. Per ovvi motivi l'esecuzione materiale delle radiografie e dei trattamenti terapeutici è affidata per legge esclusivamente a personale specializzato. I principi che deve rispettare la radioprotezione sono regolati dall'art. 2 del citato D. Lgs. 230/95, il quale dice brevemente che i principi sono di giustificazione, di ottimizzazione e limitazione delle dosi. Inoltre, oltre che l'esposizione alle radiazioni, esiste anche il pericolo di contaminazione, infatti negli ambienti in cui si trovano sorgenti di radiazioni ionizzanti e in cui il lavoratore è tenuto a rispettare un regolamento interno di radioprotezione sono dette zone *classificate*. In queste zone non si devono superare determinate soglie di dose in un anno solare. Dipendentemente dall'entità del rischio si hanno diverse categorie di appartenenza.

Tempo di gestazione	Effetti
Prime 3 settimane	Nessuno
Tra la 3 <sup>a</sup> e 8 <sup>a</sup> settimana	Rischio di malformazione di organi
Tra la 8 <sup>a</sup> e 25 <sup>a</sup> settimana	Rischio di gravi ritardi mentali
Dalla 4 <sup>a</sup> settimana in poi	Rischio di tumori letali

Tabella 2.2.2: *Effetti dell'irraggiamento di un feto.*

## 2.3 Immagini a raggi X

I raggi x sono una radiazione elettromagnetica (fotoni) con lunghezza d'onda compresa approssimativamente tra  $10nm$  e  $1pm$ . Quelli con una lunghezza d'onda maggiore di  $0.1nm$  sono chiamati raggi X molli, mentre quelli aventi lunghezze minori, sono chiamati raggi X duri. I raggi X duri si affiancano ai raggi  $\gamma$  in quanto ad energie e quindi frequenza, vengono distinti in base alla loro origine, i fotoni X sono prodotti da variazioni della cinetica degli elettroni, mentre quelli gamma da transizioni e decadimenti all'interno di un nucleo atomico (origine nucleare), o dall'annichilazione tra un positrone ed un elettrone. La relazione che lega la lunghezza d'onda con l'energia è la seguente

$$E = h\nu = \frac{hc}{\lambda},$$

dove  $\nu$  è la frequenza dell'onda elettromagnetica,  $c$  è la velocità della luce nel vuoto e  $h$  la costante di Plank. I raggi X sono molto importanti in ambito diagnostico e terapeutico, grazie alle seguenti proprietà:

- *Penetrazione*: consentono di ottenere immagini d'ombra delle strutture interne del corpo umano, grazie al contrasto risultante fra ossa e tessuti vari, dovuto alla loro diversa densità, che causa diversa opacità al fascio di raggi X che investe il corpo,
- *Fluorescenza*: i raggi X, invisibili all'occhio umano, rendono fluorescenti alcuni materiali, questa proprietà è usata per convertire l'immagine a raggi X in un'immagine luminosa osservabile,
- *Azione chimico-fisica*: i raggi X impressionano le pellicole fotografiche in maniera simile ai fotoni della luce visibile. Interagiscono anche

con la materia organica, volutamente nel caso di utilizzi terapeutici o involontariamente, provocando danni biologici a tessuti.

### 2.3.1 Apparecchiatura a Raggi X

Sommariamente un sistema a raggi X comprende un tubo radiogeno, un bersaglio (paziente) e un rivelatore. La produzione di raggi X viene fatta mediante un apparato (tubo radiogeno), Figura 2.11, costituito da un ampolla contenente il catodo e l'anodo, il primo formato da un filamento che verrà riscaldato, e il secondo da una superce metallica obliqua. Il filamento riscalda il catodo che inizia ad emettere elettroni per effetto termoionico, accelerati poi da un'alta tensione che li farà andare verso l'anodo e colpire contro la superficie metallica, nell'impatto la maggior parte l'energia cinetica degli elettroni si trasforma in calore e soltanto l'1% in forma di raggi X, questa emissione elettromagnetica avviene per *Bremsstrahlung* (frenamento), ovvero gli elettroni interagiscono con il materiale dell'anodo e vengono rallentati o fermati. I tubi radiogeni utilizzati per la diagnostica medica sono costituiti da un contenitore termoresistente di vetro al boro in cui stato creato un vuoto spinto. Il vuoto serve per ottenere un cammino libero medio delle molecole di gas dell'ordine di alcuni  $Cm$  e consentire di conseguenza di applicare al tubo le alte differenze di potenziale necessarie. Il tubo è avvolto da un manicotto riempito di olio per raffreddarlo. Il filamento riscaldato dalla corrente emette un fascio di elettroni secondo la legge di Richardson

$$\Phi = AT^2 e^{-\frac{a}{K_b T}},$$

dove  $T$  è la temperatura in Kelvin e  $K_b$  la costante di Boltzmann.  $A$  dipende dalla superficie del materiale del catodo,  $a$  è il lavoro di estrazione

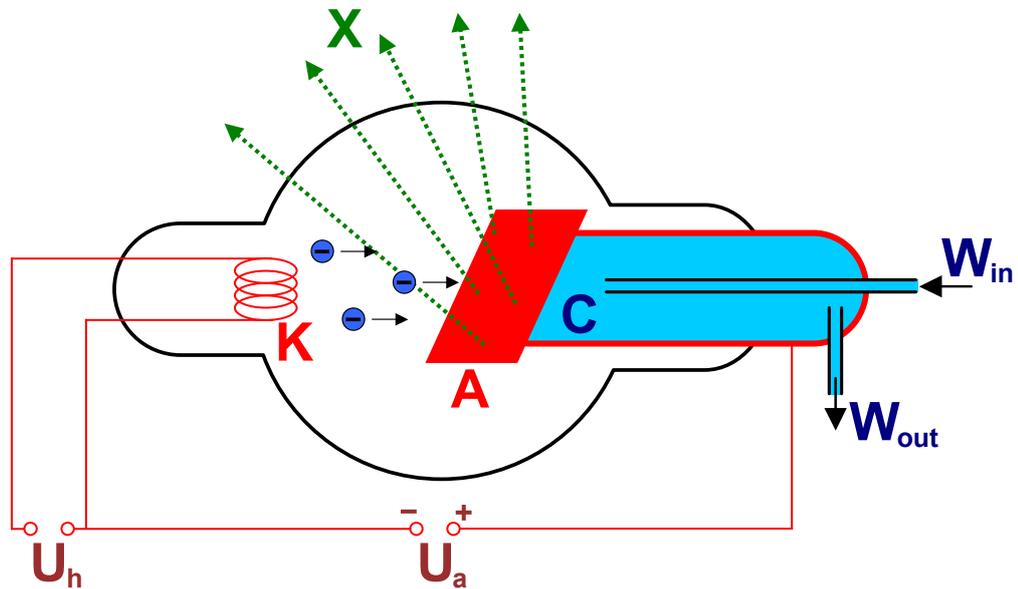


Figura 2.11: *Schema di produzione dei raggi X.*

degli elettroni e  $\Phi$  è il flusso degli elettroni che sono accelerati da una differenza di potenziale di alcune decine di  $kV$ , fino a  $150kV$ , e colpiscono l'anodo. Nella radiologia convenzionale il rivelatore è costituito da una lastra fotografica. La lastra deve essere impressionata, sviluppata e stampata, non è quindi una indagine real-time. L'intensificatore di brillantezza nasce come rivelatore alternativo alla lastra. E' costituito da un grande schermo di ingresso (fluorescente), un fotocatodo, un elevato campo elettrico generato all'interno e un piccolo schermo di uscita, il tutto all'interno di una ampolla di vetro sotto vuoto. Il fotocatodo, per effetto fotoelettrico e in corrispondenza della radiazione luminosa, emette degli elettroni che, accelerati dal campo elettrico, sbattono contro il secondo schermo convertendo la loro energia cinetica in energia luminosa. Ottengo quindi una immagine nel visibile che corrisponde all'intensità dei raggi X che hanno investito lo schermo in ingresso, di dimensioni ridotte rispetto a quella d'entrata ma

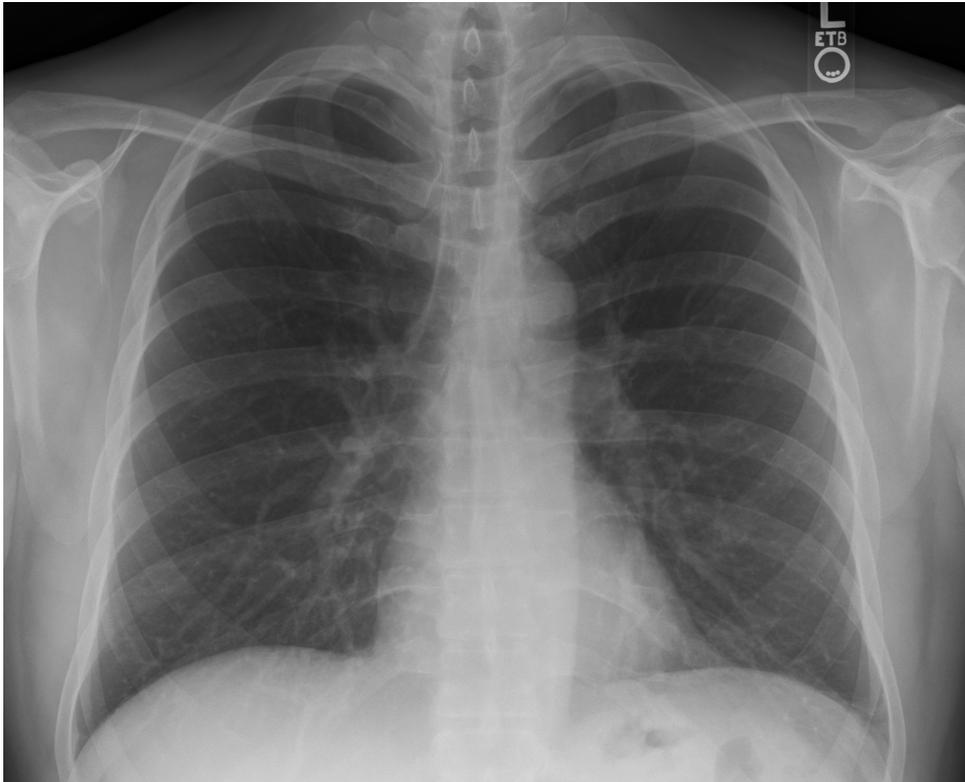


Figura 2.12: *Radiografia al torace.*

di luminosità superiore. In Figura 2.12 è mostrata una classica radiografia al torace. I vantaggi sono che si riesce a visualizzare in tempo reale l'immagine, non si ha bisogno di una lastra perchè l'immagine compare sullo schermo e si possono fare scansioni utilizzando minore energia (*radioscopia e fluoroscopia*) e quindi minore esposizione per il paziente. In questo caso si ha però svantaggio in termini di risoluzione, dovuto sia alle minori quantità di energia in gioco, sia alla presenza di uno schermo di uscita piccolo. Nella pratica clinica spesso si fa prima una radioscopia per vedere se la regolazione dell'apparecchiatura è corretta, in seguito a questa verifica si effettua la radiografia, in una indagine più mirata. Campionando il segnale TV in uscita e acquisendo l'immagine punto per punto e riga per riga, avrò

una uscita numerica, con grandi vantaggi per l'archiviazione, l'elaborazione e la trasmissione. Il rivelatore è quello che ha subito una evoluzione più importante nel tempo ed è quello che ha ancora ampi margini di sviluppo. I metodi di imaging diagnostico a raggi X più diffusi e più significativi dal punto di vista dello sviluppo tecnologico sono i Film + schermo di rinforzo, la Computed Radiology (CR) e la Digital Radiography (DR).

## **2.4 La Tomografia Computerizzata (CT)**

La radiografia convenzionale ha dei limiti, tra cui il fatto avere una proiezione 2D di una realtà 3D e la limitata risoluzione di contrasto. La Tomografia Computerizzata (CT) consente di superare molte delle limitazioni imposte. Le immagini CT rappresentano la distribuzione del coefficiente di attenuazione dell'oggetto in una sezione predefinita ed in esse possiamo osservare differenze di attenuazione altrimenti non percepibili con altre tecniche. Un sistema CT è costituito dal tomografo (scanner), da un sistema computerizzato di calcolo con dei processori ausiliari per ridurre i tempi di calcolo richiesti per la ricostruzione dell'immagine, e da una stazione di consultazione dalla quale attraverso display e dispositivi interattivi si può manipolare le immagini al fine di estrarre informazioni utili nelle varie procedure diagnostiche. In Figura 2.14 è mostrata<sup>2.6</sup> un'apparecchiatura per tomografia computerizzata completa di lettino porta paziente. Il risultato di un esame CT è una serie di matrici (slice) allineate perpendicolarmente all'asse definito dalla colonna vertebrale del paziente. Ogni slice rappresenta una fetta del corpo del paziente. I moderni tomografi riescono a produrre una slice in una frazione di secondo. Un intero esame richiede mediamente

---

<sup>2.6</sup>Fonte:[[https://it.wikipedia.org/wiki/File:UPMCEast\\_CTscan.jpg](https://it.wikipedia.org/wiki/File:UPMCEast_CTscan.jpg)]

un tempo che va dai 5 ai 15 minuti, ma una scansione avviene in una decina di secondi. La dose di radiazioni da CT è simile quella di una serie di immagini tradizionali a raggi X.

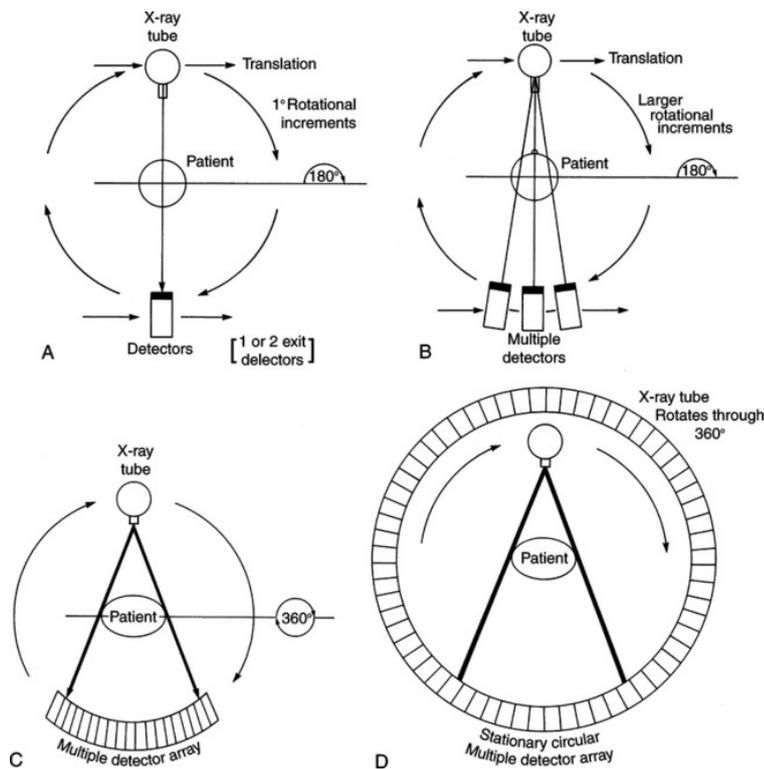


Figura 2.13: *Diverse generazioni di scanner per CT, rispettivamente da A a D dalla 1<sup>a</sup> alla 4<sup>a</sup>.*

### 2.4.1 Geometrie di scansione

La strumentazione di misura (scanner) ha subito numerose evoluzioni nella sua geometria, tutte quelle adottate presentano una sorgente di radiazione costituita da un tubo a raggi X ed un sistema di rilevamento. Sia sorgente che rivelatori effettuano movimenti rispetto all'oggetto per generare un

insieme di proiezioni. Facendo riferimento<sup>2.7</sup> alla Figura 2.13 distinguiamo quattro diverse generazioni:

- *Prima generazione*: il sistema è formato da una sorgente ed un rivelatore. L'acquisizione di una proiezione comporta un movimento di traslazione del tubo e del rivelatore, per cambiare angolo di proiezione è necessaria una loro rotazione. Essendoci di due movimenti si hanno lunghi tempi di acquisizione (circa dieci minuti).
- *Seconda generazione*: si hanno  $n$  rivelatori, di solito 10, disposti lungo una retta. In questo modo con una traslazione completa del sistema sorgente-rilevatori si acquisiscono simultaneamente  $n$  proiezioni. Richiede sempre due movimenti, il tempo di acquisizione viene diminuito di un fattore  $n$ , arrivando quindi ad un valore di un minuto.
- *Terza generazione*: viene introdotto un array di rivelatori (un migliaio) disposti ad arco, capaci di un movimento solidale con la sorgente. La geometria delle proiezioni cambia, si passa da raggi paralleli a raggi a ventaglio. La presenza del solo movimento rotatorio permette un'ulteriore diminuzione dei tempi di acquisizione, un secondo.
- *Quarta generazione*: in questo caso i sensori sono fissi e sono disposti ad anello attorno al paziente e l'unico elemento di movimento è il tubo radiogeno che può ruotare attorno all'asse del sistema. I tempi di acquisizione scendono ulteriormente, sotto al secondo.

---

<sup>2.7</sup>Fonte:[<http://radiologykey.com/computed-tomography-4/>]

### 2.4.2 Evoluzioni e spiral CT

Successivamente sono state introdotte le CT a spirale, che ha sostituito gli scanner che acquisivano una fetta per volta. La tecnica consente di acquisire in un unico passaggio i dati relativi ad un intero volume. Questo è possibile aggiungendo ad uno scanner con tecnologia *slip ring* (il tubo radiogeno ruota continuamente) un movimento di traslazione del lettino porta paziente. Un'ulteriore evoluzione della TC a raggi X è costituita dai sistemi *multislice* in grado di acquisire simultaneamente i dati di più sezioni utilizzando un array di sensori multipli. Nelle Figure 2.15 e 2.16 sono mostrati<sup>2.8</sup> degli esempi di immagini ottenute da CT.

---

<sup>2.8</sup>Fonte:[<https://it.wikipedia.org/wiki/File:Ct-workstation-neck.jpg>]



Figura 2.14: *Apparecchiatura per tomografia computerizzata.*

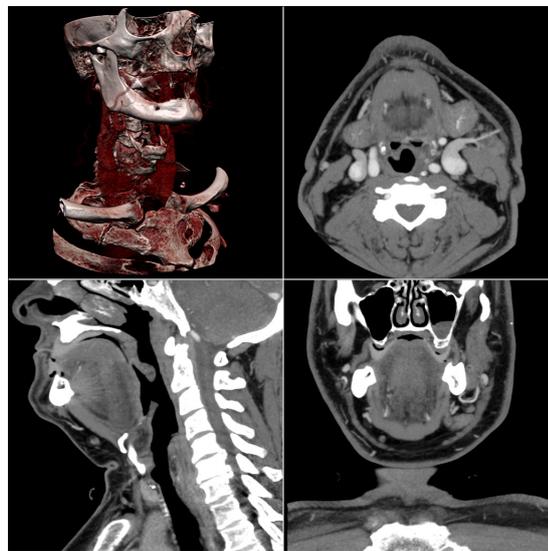


Figura 2.15: *Esempi di immagini solitamente fornite da una TC: in alto a sinistra la ricostruzione volumetrica, a destra il piano assiale, in basso i piani sagittale e frontale.*

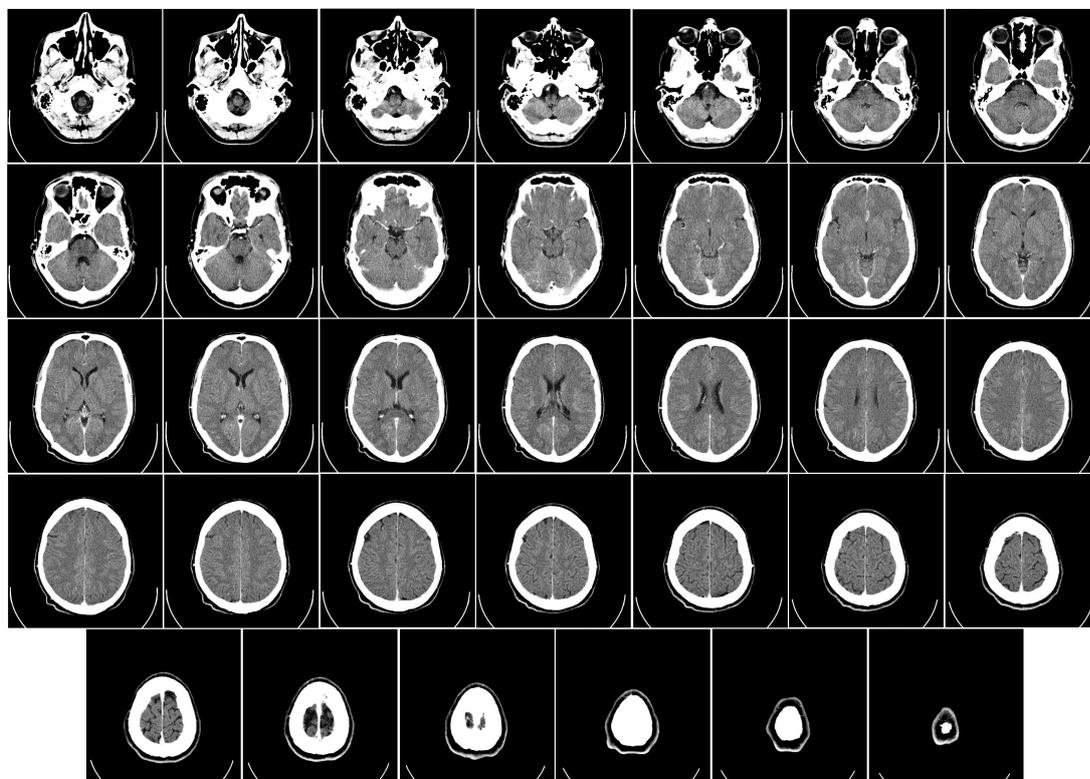


Figura 2.16: *TAC di un cervello umano, dalla base del cranio fino in cima. Eseguita con mezzo di contrasto intravenoso.*

## 2.5 Tomografia ad Emissione di Positroni (PET)

La tomografia ad emissione di positroni (PET) fornisce informazioni sulla chimica del corpo che nessun'altra tecnica è in grado di dare. La *PET* utilizza isotopi radioattivi emittenti positroni, poichè i traccianti utilizzati hanno una emivita molto breve, la quantità di radiazioni ricevute dal paziente è relativamente bassa, pari a circa 2 radiografie al torace. L'emivita particolarmente bassa dei radiofarmaci implica che la loro produzione deve avvenire in loco, mediante *ciclotrone*. Si tratta di isotopi di elementi costituenti la materia biologica, quindi ossigeno, carbonio, azoto e fluoro

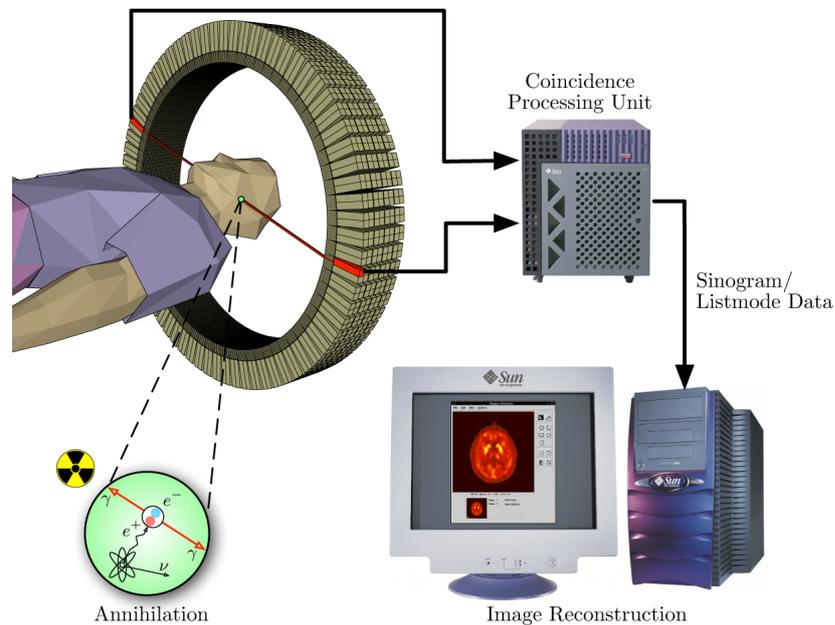


Figura 2.17: *Schema di un processo di acquisizione PET.*

(che è chimicamente sostituibile all'idrogeno). Potenzialmente si potrebbero marcare tutte le molecole del corpo umano. Un tomografo PET, mostrato in Figura 2.17, consiste in un anello di rivelatori che circondano l'organo studiato, collegati a circuiti di rivelazione di coincidenza. Il rilevamento di due fotoni contemporanei, definisce una linea lungo la quale l'evento si è verificato. Gli scanner più moderni sono dotati di più anelli di rivelatori che formano un cilindro di rivelatori, con questa geometria si possono ottenere sia ricostruzioni 2D, dove ogni anello opera separatamente; e ricostruzioni 3D dove gli anelli operano come un unico sistema e vengono rivelate sia coincidenze all'interno dei singoli anelli che fra anelli diversi. Le immagini PET forniscono informazioni funzionali che devono essere integrate con informazioni anatomiche generate con altri metodi (CT). Da qui nascono gli scanner ibridi PET-CT. La PET è usata soprattutto in cardiologia, neurologia ed oncologia. I traccianti usati mimano il comportamento meta-



Figura 2.18: *Tipica apparecchiatura PET.*

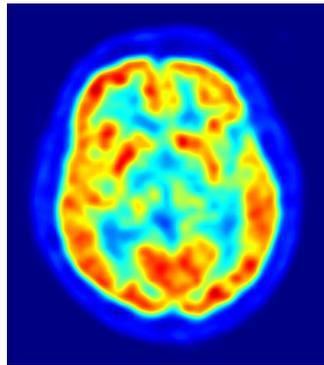


Figura 2.19: *PET cerebrale.*

bolico di sostanze utilizzate dall'organismo, come il glucosio o l'ossigeno, accumulandosi nelle aree di maggior consumo.

## 2.6 Immagini di risonanza magnetica (MR)

La Risonanza Magnetica Nucleare è un fenomeno fisico quantistico nel quale nuclei atomici immersi in un campo magnetico esterno assorbono e riradono una radiazione elettromagnetica. Tale radiazione è ad una specifica frequenza (di risonanza) che dipende dall'intensità del campo magnetico e dalle proprietà intrinseche dei nuclei. Il fenomeno della risonanza magne-



Figura 2.20: *Una apparecchiatura medica per la risonanza magnetica.*

tica ha numerosi campi di applicazione: spettroscopia, imaging, magnetometria, analisi movimento macromolecole. L'imaging a risonanza magnetica (*MRI*) ha riscosso un grande successo, dovuto ai seguenti fattori fondamentali:

- osservazione di grandezze fisiche caratteristiche della struttura tissutale con notevole capacità di discriminare i tessuti molli,

- uso di radiazione elettromagnetica non ionizzante
- acquisizione di immagini di sezioni, secondo piani con orientazione qualsiasi, e in generale di volumi,
- possibilità di indagare aspetti sia morfologici che funzionali, in particolare modo dell'encefalo.

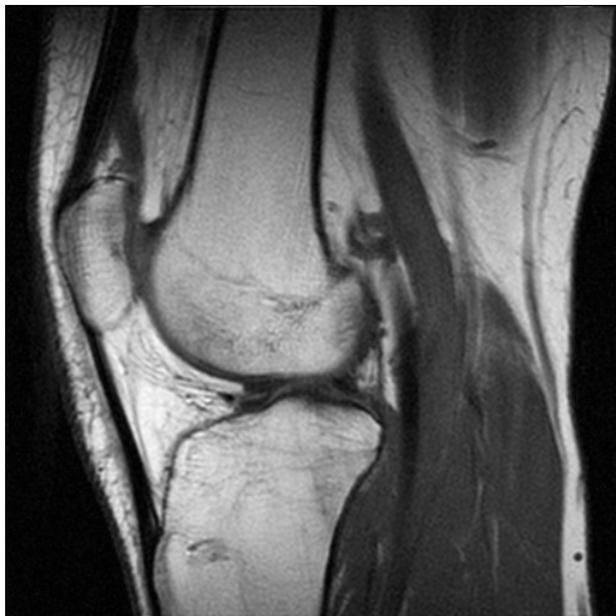


Figura 2.21: *Immagine sagittale di un ginocchio ottenuta mediante risonanza magnetica.*

In Figura 2.20 è mostrata<sup>2.9</sup> una classica apparecchiatura per Imaging a Risonanza Magnetica, mentre in Figura 2.21 è mostrata<sup>2.10</sup> una tipica immagine ottenuta attraverso questa tecnica.

---

<sup>2.9</sup>Fonte:[<https://it.wikipedia.org/wiki/File:MRI-Philips.JPG>]

<sup>2.10</sup>Fonte:[[https://it.wikipedia.org/wiki/File:MR\\_Knee.jpg](https://it.wikipedia.org/wiki/File:MR_Knee.jpg)]

## 2.6.1 Principi fisici

### 2.6.1.1 Spin e momento magnetico

La descrizione dei fenomeni della risonanza magnetica nucleare coinvolge il formalismo della meccanica quantistica, tuttavia dal momento che nei sistemi macroscopici aventi spin  $\frac{1}{2}$  le descrizioni del modello classico sono in accordo con quelle della teoria quantistica, si può attuare una trattazione più semplice semi-classica seguendo l'approccio di Bloch. Ad ogni

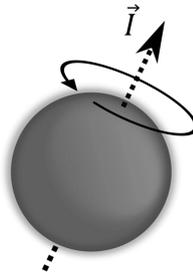


Figura 2.22: *Momento di spin.*

nucleo atomico possiamo associare una quantità misurabile detta *spin*, il *momento di spin*  $\vec{I}$  si può assimilare ad un momento intrinseco si quantità di moto rotazionale, Figura 2.22, ma si tratta di una proprietà intrinseca del nucleo, che in realtà non ruota, come la carica o la massa. Al numero di spin  $I$  possono essere associati  $2I + 1$  livelli energetici quando al nucleo è applicato un campo magnetico statico esterno. Per in nuclei che hanno un numero pari sia di protoni che neutroni si ha  $I = 0$ . Tra gli altri nuclei, quelli che son attivi dal punto di vista della risonanza magnetica nucleare, e quindi quelli utili ai fini dell'imaging *MR*, hanno  $I = \frac{1}{2}$ , caratterizzati da due livelli energetici. L'idrogeno  $^1H$  per la sua abbondanza nei tessuti è il nuclide solitamente utilizzato per l'imaging basato sulla risonanza magnetica nucleare, quindi sarà quello preso in riferimento. L'elettromagnetismo

insegna che il moto rotatorio di una carica elettrica produce un momento magnetico  $\vec{\mu}$  legato a quello meccanico dalla seguente relazione

$$\vec{\mu} = \gamma \vec{I},$$

dove  $\gamma$  è il *rapporto magnetico*, ed è una costante caratteristica che varia per ogni nucleo. Il momento magnetico avrà quindi lo stesso verso del momento di spin. Il modulo di  $\vec{\mu}$  si può esprimere come

$$\mu = \gamma \frac{h}{2\pi} \sqrt{I(I+1)},$$

dove  $h$  è la costante di Plank. L'orientazione di  $\vec{\mu}$  in assenza di eccitazioni esterne è casuale, per via dei fenomeni di agitazione termica. A livello macroscopico non si osserva nessun campo magnetico. Si può osservare una magnetizzazione macroscopica soltanto applicando un campo magnetico esterno  $\vec{B}_0$  che ipotizzeremo orientato secondo l'asse  $z$ . Quindi si avrà che la componente di  $\vec{\mu}$  rispetto al campo magnetico  $\vec{B}_0$  sarà

$$\mu_z = \gamma m_I \frac{h}{2\pi},$$

dove  $m_I$  è il numero quantico magnetico che assume i valori  $m_I = -I, \dots, I$ , corrispondenti a  $2I + 1$  possibili orientazioni di  $\vec{\mu}$  rispetto al campo esterno. Bisogna far presente che la componente di  $\vec{\mu}$  nel piano  $(x, y)$  è casuale. Nel caso di un unico protone sono possibili soltanto due elettroni,  $^1H, I = \frac{1}{2}$ , una nel verso del campo magnetico (*spin-up*) e una in verso opposto (*spin-down*), come in Figura 2.23. Su ogni protone agisce un momento meccanico dato da  $\vec{\mu} \times \vec{B}_0$  e possiamo scrivere l'equazione del moto

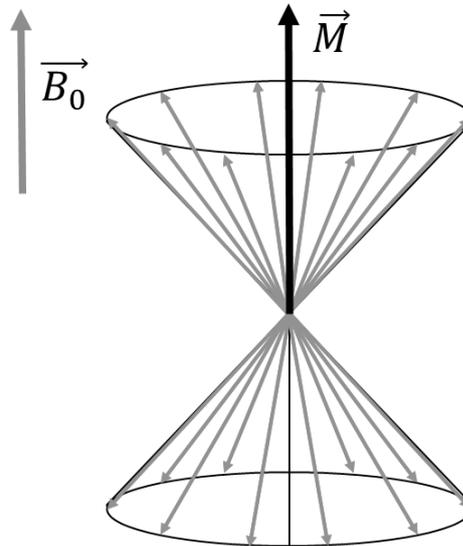


Figura 2.23: Effetto di un campo magnetico statico  $\vec{B}_0$ .

come

$$\frac{d\vec{I}}{dt} = \vec{\mu} \times \vec{B}_0,$$

quindi

$$\frac{d\vec{\mu}}{dt} = \gamma \vec{\mu} \times \vec{B}_0;$$

che descrive un moto di precessione del vettore  $\vec{\mu}$  attorno a  $\vec{B}_0$  con una velocità angolare definita dalla *legge di Larmor*:

$$\vec{\Omega} = -\gamma \vec{B}_0.$$

L'analogia con il moto di precessione di una trottola attorno all'asse gravitazionale è evidente (Figura 2.24)<sup>2.11</sup>. In campo medico si utilizzano tipicamente valori<sup>2.12</sup> del campo magnetico di  $0.2 \div 3T$  (Tesla), mentre il campo

<sup>2.11</sup> Modified from: [https://en.wikipedia.org/wiki/File:PrecessionOfATop.svg]

<sup>2.12</sup> In alcuni casi si può arrivare a  $7T$ .

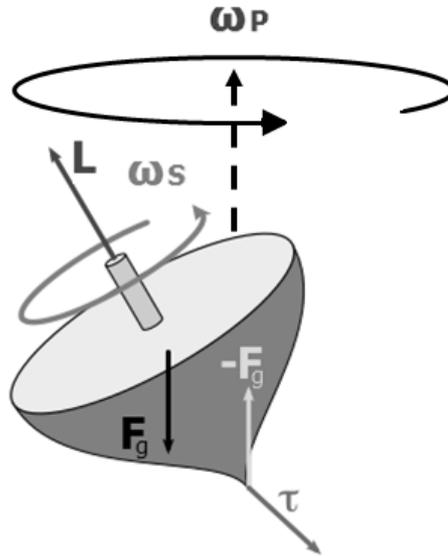


Figura 2.24: Analogia meccanica, moto di precessione.

magnetico terrestre è di soli  $0.3 \div 0.6 \text{ Gauss}^{2,13}$  e quindi non ha un effetto significativo sullo spin. La presenza del campo magnetico induce due diversi livelli energetici nella popolazione degli spin, il cosiddetto *Effetto Zeeman*, che corrispondono a due diverse orientazioni aventi energia

$$E = -\vec{\mu} \cdot \vec{B}_0 = \mu_z B_0 = -\gamma \frac{h}{2\pi} m_I B_0,$$

ovvero per ogni orientazione:

$$E_{\uparrow} = -\frac{1}{2} \gamma \frac{h}{2\pi} B_0, \quad E_{\downarrow} = \frac{1}{2} \gamma \frac{h}{2\pi} B_0.$$

Lo stato *spin-up* corrisponde al più basso livello energetico. Il salto energetico fra i due livelli è quindi

$$\Delta E = E_{\downarrow} - E_{\uparrow} = \gamma \frac{h}{2\pi} B_0 = h\nu, \quad \nu = \frac{\gamma}{2\pi} B_0$$

---

<sup>2,13</sup>  $1T = 10^4 \text{ Gauss}$ .

dove  $\nu$  prende il nome di *frequenza di risonanza o di Larmor*.

In Figura 2.25 sono mostrati i due possibili valori dello spin. La distribu-

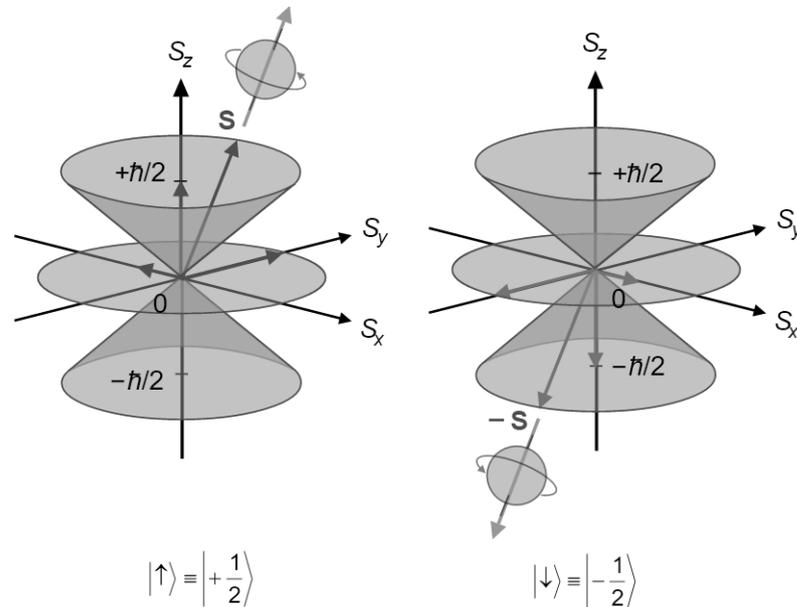


Figura 2.25: *Momento angolare di una particella di spin 1/2.*

zione della popolazione degli spin nei due livelli energetici  $E_{\downarrow} - E_{\uparrow}$  è data dalla *legge di distribuzione di Boltzmann*:

$$\frac{N_{\uparrow}}{N_{\downarrow}} = e^{\frac{\Delta E}{k_B T}},$$

dove  $N_{\uparrow}$  e  $N_{\downarrow}$  sono rispettivamente il numero degli *spin-up* e il numero degli *spin-down*,  $T$  la temperatura espressa in *kelvin*, mentre  $k_B$  è la costante di Boltzmann. Questo vuol dire che esiste un leggerissimo eccesso di protoni nello stato energetico più basso (*spin-up*) che causa la magnetizzazione macroscopica. Facendo i calcoli risulta che per  $B_0 = 1T$  e  $T = 300K$  si ha un eccesso di protoni nello stato *spin-up* di dell'ordine di qualche protone per milione rispetto allo stato *spin-down*. Nella MRI interessa studiare non il comportamento dei singoli, ma la magnetizzazione risultante  $\vec{M}$  a livello

macroscopico generata dalla popolazione di spin osservata

$$\vec{M} = \sum_i \vec{\mu}_i,$$

che come abbiamo detto è diverso da zero solo in presenza di un campo magnetico esterno e la sua intensità cresce con l'aumentare della differenza tra le popolazioni *spin-up* e *spin-down*, ovvero al crescere dell'intensità di  $\vec{B}_0$ . All'equilibrio la direzione di  $\vec{M}$  è la stessa di  $\vec{B}_0$ .

### 2.6.1.2 Segnale MR

La generazione del segnale *MR* è basata sulla precessione di *Larmor* vista poco sopra, a grandi linee la si può descrivere nel modo seguente. Supponiamo di spostare in qualche modo il vettore  $\vec{M}$  dalla sua posizione di equilibrio (Figura 2.23) ottenendo la situazione in Figura 2.26. Il vettore

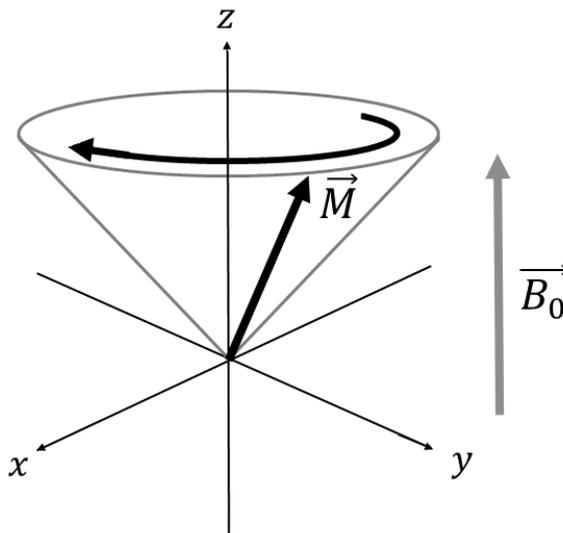


Figura 2.26: *Precessione di Larmor*.

$\vec{M}$  comincia un moto di precessione attorno all'asse del campo magnetico

$\vec{B}_0$  con frequenza pari alla frequenza di Larmor  $\nu$ . Per l'idrogeno ( $^1H$ ) e un campo magnetico dell'ordine del Tesla ( $\vec{B}_0 = 0.1 \div 2T$ ) questa frequenza appartiene alle radiofrequenze,  $1 \div 100MHz$  quindi non sono radiazioni ionizzanti e attraversano i tessuti senza attenuazione apprezzabile. Ponendo un'opportuna bobina sul piano  $(x, y)$  si può rilevare la proiezione di  $\vec{M}$  su tale piano, questo costituisce il *segnale di risonanza magnetica*. Per deflettere il vettore di magnetizzazione  $\vec{M}$  bisogna applicare un'ulteriore perturbazione che si ottiene con un campo magnetico  $\vec{B}_1$  rotante in un piano ortogonale a  $\vec{B}_0$ , quindi sul piano  $(x, y)$ . Scegliendo la pulsazione di  $\vec{B}_1$  coincidente con quella di Larmor si stabilisce una condizione di risonanza detta *risonanza magnetica nucleare*. Si verifica infatti che la radiazione elettromagnetica interagente con il sistema è formata da fotoni aventi energia

$$h\nu = h\frac{\gamma}{2\pi}B_0,$$

corrispondente, come abbiamo già visto, al salto di energia  $\Delta E$  tra i due livelli ammissibili. La frequenza  $\nu$  del campo  $\vec{B}_1$ , come detto poco sopra, corrisponde alle radiofrequenze, quindi il campo  $\vec{B}_1$  è detto *eccitazione a radiofrequenza*. Il campo  $\vec{B}_1$  può essere generato utilizzando la stessa bobina di acquisizione, si invia una corrente oscillante tale da generare un campo rotante di intensità  $2B_1$  lungo l'asse  $x$ . A livello macroscopico, in un sistema di riferimento rotante solidalmente con  $\vec{B}_1$ , come si vede in Figura 2.27, l'eccitazione a radio frequenza provoca la rotazione del vettore  $\vec{M}$  attorno a  $\vec{B}_1$  di un angolo  $\alpha$  detto *flip angle*, che assumendo un impulso rettangolare, è dato da

$$\alpha = \Omega\delta t = \gamma B_1\delta t,$$

con  $\delta t$  il tempo di applicazione del campo rotante. Se  $\delta t$  è abbastanza lungo

da produrre una rotazione pari a  $\alpha = \frac{\pi}{2}$  si parla di *impulso a 90°*, mentre se  $\alpha = \pi$  si parla di *impulso a 180°*. Ad esempio per  $B_1 = 0.6 \text{ gauss}$  si ottiene  $\alpha = \frac{\pi}{2}$  per  $\delta t = 0.1 \text{ ms}$ . Prendendo  $M_0$  come valore all'equilibrio di  $\vec{M}$ ,

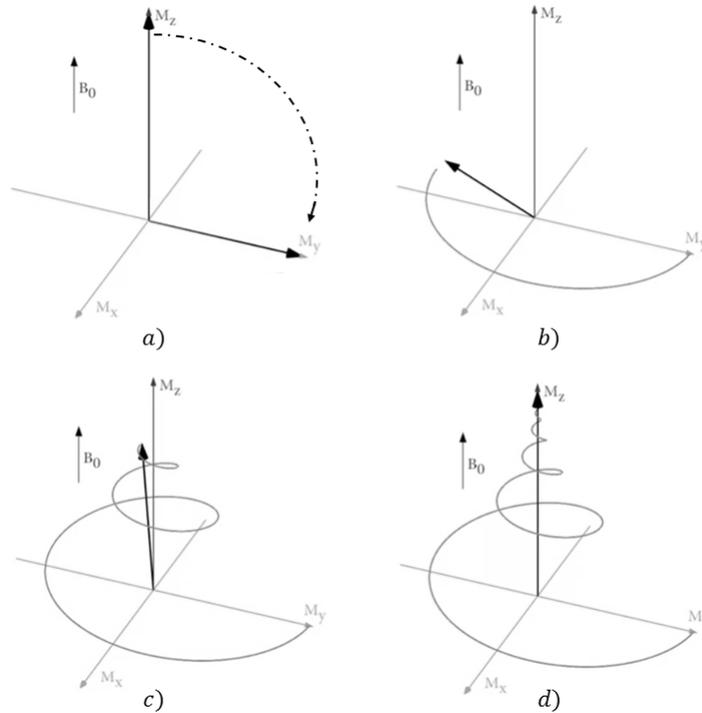


Figura 2.27: *Effetto dell'eccitazione a radiofrequenza, in a) si evidenzia l'inclinazione di  $\vec{M}$  dovuta all'impulso, nelle successive il vettore compie un movimento a spirale ritornando al punto di equilibrio.*

abbiamo che, come si può osservare in Figura 2.27, il vettore  $\vec{M}$  descrive un moto a spirale con velocità angolare  $\Omega$ .

Quindi se  $\alpha \neq \pi$  la *componente trasversale* di  $\vec{M}$ , sul piano  $(x, y)$ , sarà data da

$$M_{xy} = M_0 \sin \alpha \cdot e^{j(\Omega t + \varphi)}$$

che produce, per induzione elettromagnetica, un segnale elettrico sulla bobina ricevente. Per  $\alpha = \pi$  si ha soltanto un ribaltamento della magnetiz-

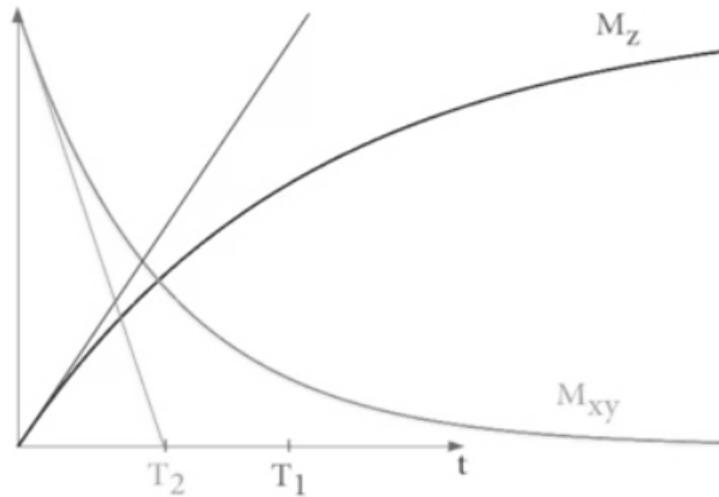


Figura 2.28: Andamento componente longitudinale e trasversale di  $\vec{M}$ .

zazione, senza componente trasversale. Quanto termina l'eccitazione, la *componente longitudinale*, lungo l'asse  $z$ , tende a recuperare il suo valore di equilibrio e la componente trasversale decade a zero, questo andamento si può osservare in Figura 2.28, dove sono indicati anche i tempi  $T_1$  e  $T_2$  che introdurremo tra poco. Il segnale rilevato sulla bobina ricevente tende a zero quasi esponenzialmente, viene detto *Free Induction Decay (FID)*. Il segnale FID, il cui andamento è mostrato in Figura \* può essere espresso in questo modo:

$$S_x(t) = M_x(t) \cdot \text{Rilass. Trasv.} = k \cdot M_0 \cos(2\pi\nu t) e^{-\frac{t}{T_2}},$$

con  $k$  costante dipendente dalla bobina. Il FID è un'oscillazione smorzata con ampiezza massima dipendente da  $M_0$ , una volta fissati  $\vec{B}_0$  ed  $\alpha$ . Complessivamente, la componente longitudinale della magnetizzazione ritorna al valore iniziale per via degli scambi energetici tra i nuclei di idrogeno e le molecole circostanti (reticolo) con una costante di tempo  $T_1$ , questo pro-

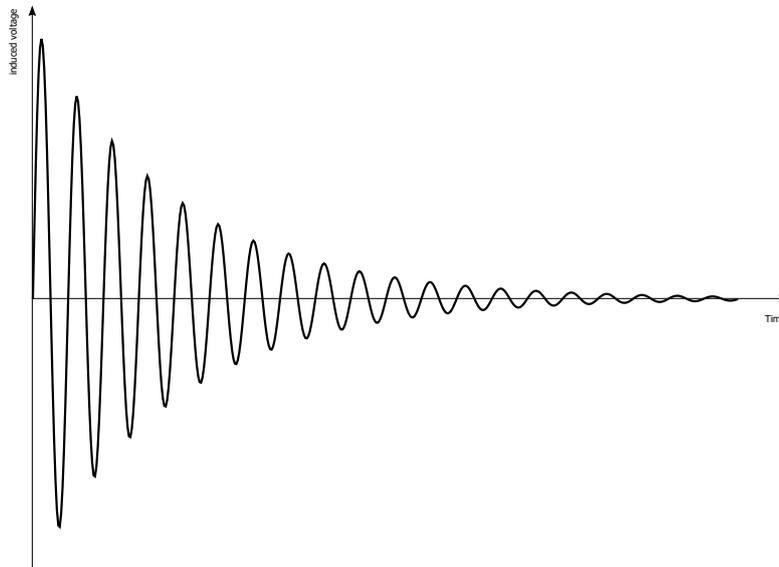


Figura 2.29: *Segnale Free Induction Decay (FID)*.

cesso è detto *rilassamento spin-reticolo*. La componente trasversale decade a causa di interazioni tra gli spin con una costante di tempo  $T_2$ , questo invece è il *rilassamento spin-spin*. In realtà, essendo  $\vec{B}_0$  non omogeneo, si ha un maggior sfasamento tra gli spin e quindi una riduzione della componente trasversale, perciò il FID è caratterizzato da una costante di tempo  $T_2^*$  minore di  $T_2$ . Sia  $T_1$  che  $T_2$  dipendono dalla mobilità delle molecole, dal peso molecolare, dall'intensità del campo magnetico e dalla temperatura. La relazione  $T_2 \leq T_1$  è sempre verificata, in genere  $T_1$  varia molto di più di  $T_2$ . In Tabella 2.6.1 si possono osservare i tempi di rilassamento tipici per alcuni tessuti. Il segnale FID è intrinsecamente multiparametrico perchè dipende dalla densità protonica, da  $T_1$  e da  $T_2$ . Il peso di questi termini all'interno del segnale FID può essere gestito cambiando sequenza di eccitazione, ovvero cambiando il numero, il tipo e la durata degli impulsi a RF che deflettono  $\vec{M}$ . Nei tessuti biologici a  $37^\circ\text{C}$  i tempi di rilassamento

Tessuto	$T_1$ (ms)	$T_2$ (ms)
Grasso	240	85
Fegato	420	45
Sotanza bianca	680	90
Sostanza grigia	810	100
Acqua	$\sim 3000$	$\sim 3000$

Tabella 2.6.1: *Tempi di rilassamento  $T_1$  e  $T_2$  tipici per alcuni tessuti per  $B_0 = 1.5T$ .*

sono più brevi che nei liquidi perchè la mobilità è minore e ciò facilita le interazioni e gli scambi energetici. Riassumendo quindi, la generazione di un segnale MR prevede:

- la presenza di un campo magnetico statico  $\vec{B}_0$ ,
- l'applicazione di una eccitazione a radiofrequenza mediante il campo  $\vec{B}_1$
- la rilevazione in una bobina ricevente del segnale FID.

Il comportamento della magnetizzazione  $\vec{M}$  può essere descritta attraverso l'equazione di Bloch, che in riferimento all'asse  $z$  parallelo a  $\vec{B}_0$  assume la forma

$$\frac{d\vec{M}}{dt} = \gamma \vec{M} \times \vec{B} - \vec{R} (\vec{M} - \vec{M}_0),$$

dove  $\vec{B} = \vec{B}_0 + \vec{B}_1$  e  $\vec{M}_0$  è la magnetizzazione di equilibrio, mentre  $\vec{R}$  è la matrice di rilassamento definita come

$$\vec{R} \begin{pmatrix} \frac{1}{T_2} & 0 & 0 \\ 0 & \frac{1}{T_2} & 0 \\ 0 & 0 & \frac{1}{T_1} \end{pmatrix}.$$

## 2.6.2 Sequenze di eccitazione

L'utilizzo di *disequenze di eccitazione* opportunamente progettate consente di enfatizzare la dipendenza di FID dai tre parametri  $\rho, T_1, T_2$  e di controllare di conseguenza il contrasto dell'immagine. Da notare che l'immagine della sola densità protonica ( $\rho$ ) è poco significativa perchè permette di differenziare soltanto limitatamente i tessuti. Un parametro caratteristico a tutte le sequenze di eccitazione è il *tempo di ripetizione*  $T_R$ , ovvero l'intervallo temporale tra l'inizio di una sequenza e la successiva. Per descrivere le caratteristiche di una sequenza di eccitazione si utilizza la seguente notazione

$$(\alpha_1 - t_1 - \alpha_2 - t_2 \dots \alpha_m - t_m)_n,$$

dove  $\alpha_i$  indica i gradi dell'impulso e  $t_i$  il tempo tra un impulso e il successivo  $\alpha_{i+1}$ , si noti che  $\sum_i t_i = T_R$ , mentre  $n$  è il numero di ripetizioni della sequenza. Esistono diverse sequenze di eccitazione, di seguito illustreremo brevemente le più importanti.

### 2.6.2.1 Saturation Recovery

Con questo tipo di eccitazione viene enfatizzata la dipendenza da  $T_1$ , applicando ripetutamente impulsi a  $90^\circ$  distanziati di un tempo di ripetizione  $T_R$ . Dall'equazione di Bloch si ha che l'ampiezza  $S$  dei segnali FID successivi al primo risulta proporzionale a

$$\rho \left[ 1 - e^{\left(-\frac{T_R}{T_1}\right)} \right].$$

Per  $T_R$  lunghi rispetto a  $T_1$  si ottiene praticamente una mappa della densità protonica, mentre viceversa per  $T_R$  brevi si ha un'immagine pesata dal

contributo di  $T_1$ .

### 2.6.2.2 Spin echo

In questo caso oltre ad un primo impulso a  $90^\circ$  segue un altro impulso a  $180^\circ$  dopo un tempo di  $\frac{T_E}{2}$ .

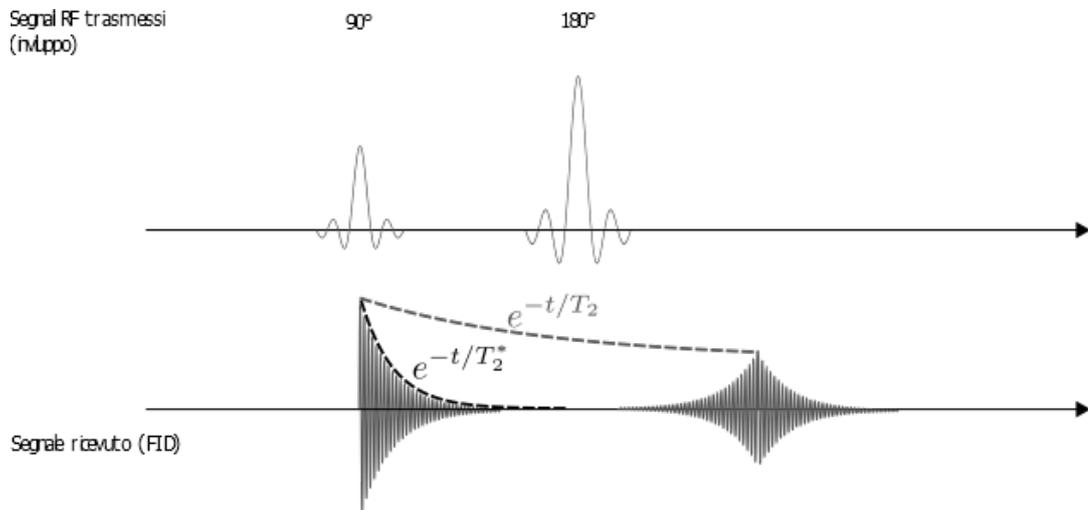


Figura 2.30: Sequenza di eccitazione spin-echo.

L'effetto del secondo impulso (rotazione di  $180^\circ$  intorno all'asse  $y$ ) è quello di riallineare gli spin producendo un *segnale echo* dopo un ulteriore tempo di  $\frac{T_E}{2}$ , che viene acquisito, la cui ampiezza  $S$  ricavata dall'equazione di Bloch, è approssimativamente proporzionale a

$$\rho \left[ 1 - e^{-\frac{T_R}{T_1}} \right] e^{-\frac{T_E}{T_2}},$$

per  $T_E \ll T_R$ . Scegliendo  $T_R \gg T_1$  e  $T_E \ll T_2$  si ottengono immagini di densità protonica, con  $T_R \approx T_1$  e  $T_E \ll T_2$  si hanno immagini pesate su  $T_1$ , mentre scegliendo  $T_R \gg T_1$  e  $T_E \geq T_2$  si hanno immagini pesate su

$T_2$ . In Figura 2.30 è mostrata<sup>2.14</sup> la sequenza di eccitazione *spin-echo* e il corrispondente segnale FID.

### 2.6.2.3 Inversion Recovery

In questo caso ad un impulso iniziale di  $180^\circ$  segue dopo un tempo  $T_i$  un impulso a  $90^\circ$  e successivamente la lettura del segnale. Il primo impulso provoca l'inversione della magnetizzazione e il secondo consente di rilevare la componente longitudinale parzialmente ripristinata durante  $T_i$ . Dall'equazione di Bloch risulta che l'ampiezza  $S$  del segnale FID è proporzionale a

$$\rho \left[ 1 - 2e^{-\frac{T_i}{T_1}} + e^{-\frac{T_R}{T_1}} \right],$$

dove  $T_R$  è sempre il tempo di ripetizione della sequenza. In questo caso si può ad esempio esaltare i perncimi rispetto ai liquidi ponendo  $T_i \leq T_1$ .

### 2.6.3 Risonanza Magnetica Funzionale

Nei primi anni '90 cominciò lo studio di metodi per la generazione di immagini funzionali mediante MR, che ha prodotto una nuova metodica chiamata *fMRI* (*functional MRI, risonanza magnetica funzionale*). L'osservazione che da origine alla fMRI è che l'emodinamica locale celebrale riflette l'attività neuronale, analogamente al principio sfruttato nell'PET. In particolare viene sfruttata una proprietà dell'emoglobina ossigenata e deossigenata<sup>2.15</sup> unitamente al fatto che l'attività neuronale produce variazioni sulle proporzioni di queste nel sangue (*BOLD: Blood Oxygenation Level Depend*). Uno degli scopi principali della fMRI riguarda l'attribuzione di ruoli funzionali a determinate aree del cervello.

---

<sup>2.14</sup>Fonte:[<https://it.wikipedia.org/w/index.php?curid=1113357>]

<sup>2.15</sup>L'emoglobina ossigenata è diamagnetica, mentre l'emoglobina deossigenata è paramagnetica.

## Capitolo 3

# Elaborazione delle immagini digitali

### 3.1 Generalità

**L**'interesse per i metodi di elaborazione delle immagini digitali (*digital image processing*) nasce da due principali applicazioni: il miglioramento delle informazioni figurate, per una migliore interpretazione da parte dell'uomo e l'elaborazione dei dati dell'immagine per l'immagazzinamento, tra trasmissione e la rappresentazione autonoma da parte delle macchine. Un'immagine può essere definita come una funzione bidimensionale  $f(x, y)$ , dove  $x$  e  $y$  sono le coordinate spaziali, e l'ampiezza di  $f$  in ogni coppia di coordinate viene chiamata intensità o livello di grigio dell'immagine in quel determinato punto, o nel caso di immagini a colori il valore di  $f$  corrisponde ad una terna riferita ai colori principali RGB (*Red Green Blue*). Quando  $x, y$  e  $f$  sono quantità finite, quindi discreti, possiamo definire l'immagine come un'immagine digitale. Gli elementi di un'immagine si definiscono come *picture element*, *image element* o *pixel* che è il termine usato per la maggiore. La vista è il senso più sofisticato, dunque non c'è da stupirsi del fatto che l'immagine occupi un ruolo prioritario nella percezione umana. In ogni caso, diversamente dall'uomo, che ha

una fruizione limitata delle radiazioni visive dello spettro elettromagnetico, i dispositivi di imaging nel trattamento e nell'acquisizione delle immagini digitali riescono a lavorare con tutte le componenti dello spettro elettromagnetico, dai raggi gamma alle onde radio. Alcuni dispositivi possono generare immagini da fonti a cui l'occhio umano non associa nessun'immagine, come ad esempio gli ultrasuoni o il microscopio elettronico. Le immagini generate dalle radiazioni elettromagnetiche sono le più familiari, specie le immagini a raggi X o quelle delle bande visibili dello spettro elettromagnetico. Le onde elettromagnetiche possono essere considerate come onde sinusoidali di lunghezza (quindi frequenza) variabile e possono essere considerate come un flusso di particelle prive di massa, ognuna delle quali viaggia alla velocità della luce. Ogni particella contiene una certa quantità di energia, che viene definita  *fotone* . In Figura 3.1 è rappresentato lo spettro elettromagnetico in base alla frequenza dei fotoni.

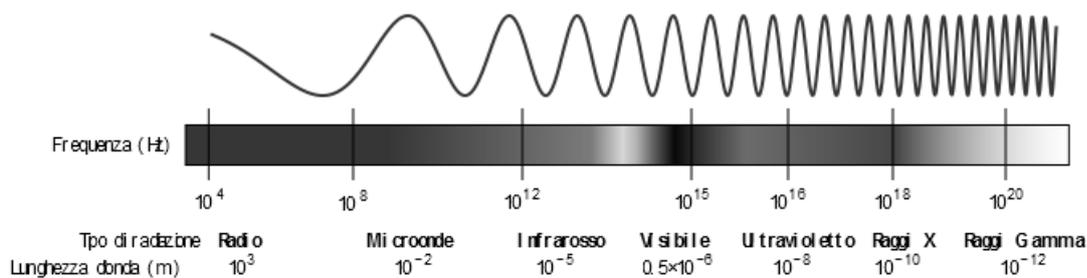


Figura 3.1: *Spettro elettromagnetico ordinato per frequenza.*

### 3.1.1 L'imaging dalle radiazioni EM

L'imaging a raggi gamma è usato principalmente in medicina nucleare e nelle osservazioni astronomiche, nel primo caso si introduce in un paziente un farmaco radioisotopo che quando decade emette raggi gamma. I raggi

X invece sono le fonti più antiche di radiazione utilizzate per l'imaging. Il loro impiego più conosciuto è quello relativo alla diagnosi medica, ma sono largamente usati anche nell'industria e in astronomia. L'ultravioletto è usato invece in litografia, l'ispezione industriale, la microscopia, i laser, la biologia e le osservazioni astronomiche. La banda visibile dello spettro EM è la più familiare quindi non ci sorprende che proprio in quest'area ci sia la gran parte delle applicazioni dell'immagine processing. Tuttavia anche la banda dell'infrarosso è spesso usata insieme all'imaging del visibile. Queste hanno varie applicazioni, tra cui la microscopia, l'astronomia, l'acquisizione remota (i.e. video sorveglianza), l'industria e l'informatica forense. La principale applicazione dell'imaging nella banda delle microonde è il radar. La caratteristica principale del radar è la sua capacità di raccogliere virtualmente informazioni su qualsiasi regione e in ogni circostanza, indipendentemente dalle condizioni meteorologiche e di luminosità. Alla parte opposta dello spettro rispetto ai raggi gamma, abbiamo le onde radio; l'imaging basato su onde radio ha anch'esse applicazioni nel campo medico e astronomico. In campo medico la ritroviamo nella risonanza magnetica (MRI).

### **3.1.2 Processi tipici del digital image processing**

Di seguito illustriamo i processi di base tipici dell'elaborazione delle immagini digitali, attraverso i quali si ottengono delle manipolazioni delle stesse al fine di migliorarne la funzionalità e la fruibilità.

*L'acquisizione delle immagini* è sicuramente il primo processo.

Il miglioramento o l'*enhancement* è un processo di manipolazione che in base alla specifica applicazione ne migliora la qualità.

Il *ripristino di un'immagine*, o *image restoration*, consiste anch'esso nel miglioramento dell'immagine. In questo caso però è un miglioramento oggettivo, basato su modelli matematici o statistici del degrado, un esempio può essere l'*upscaling* ovvero l'aumento della risoluzione di un'immagine. L'*elaborazione di un'immagine a colori* (*color image processing*) è un settore che si è sviluppato grazie al crescente utilizzo delle immagini digitali in internet.

Le *wavelet* sono la base per la compressione e per la rappresentazione piramidale, nella quale le immagini vengono successivamente suddivise in regioni più piccole.

La *compressione* riguarda le tecniche per la riduzione dello spazio occupato in memoria dall'immagine o per la banda occupata durante la sua trasmissione.

L'*elaborazione morfologica* si occupa degli strumenti per l'estrazione di componenti dalle immagini che vengono utilizzate per la rappresentazione e la descrizione della forma.

Le procedure di *segmentazione* (*segmentation*) dividono un'immagine nelle sue parti costituenti (i.e. oggetti). In generale la segmentazione è uno dei problemi più difficili nell'elaborazione delle immagini digitali. Sono processi critici perchè gli algoritmi e i metodi relativi non garantiscono in generale dei risultati ottimali in tutti i campi applicativi; più accurato sarà il processo di segmentazione e migliori saranno le prestazioni degli algoritmo di riconoscimento.

A tutto ciò si deve aggiungere la *conoscenza di base* del problema, questa conoscenza può apparire come una regione dettagliata di un'immagine in cui l'informazione è cercata, così da limitarne la ricerca. La conoscenza di

base può però essere abbastanza complessa come ad esempio una lista dettagliata di tutti i difetti possibili in un problema di controllo del materiale. La conoscenza di base controlla anche l'interazione tra i moduli di elaborazione. Da notare che non tutte le applicazioni dell'elaborazione delle immagini necessitano di tutti i processi sopra elencati, ad esempio nel miglioramento della qualità non abbiamo quasi mai bisogno di far intervenire gli altri stadi.

## **3.2 Segmentazione**

Come già anticipato, la segmentazione suddivide un'immagine nelle regioni o negli oggetti che la compongono. Il livello di dettaglio della segmentazione dipende da ciò che si vuole ottenere, cioè la segmentazione dovrebbe arrestarsi quando gli oggetti o le regioni di interesse sono stati individuati. La segmentazione è uno degli obiettivi più difficili da raggiungere nell'ambito dell'elaborazione delle immagini. La maggior parte degli algoritmi di segmentazione si basa su una delle due proprietà fondamentali dei valori di intensità, ovvero la discontinuità e la similarità. Nel primo caso si partiziona l'immagine basandosi su repentini cambiamenti di intensità, come ad esempio negli edge. Nel secondo caso si segmenta in base alle similarità tra regioni, facendo riferimento ad un insieme di criteri di similarità predefiniti. In quest'ultima tipologia rientrano le tecniche di sogliatura (*thresholding*) e di crescita delle regioni (*region growing*).

### **3.2.1 Metodi basati sulla soglia**

Grazie alle sue proprietà intuitive, alla semplicità di implementazione e alla velocità computazionale, la sogliatura è una delle tecniche più importanti

per la segmentazione.

Si supponga di avere un'immagine avente un istogramma distribuito intorno a due diversi valori, come in Figura 3.2, un'immagine del genere potrebbe consistere in degli oggetti chiari su sfondo scuro o viceversa, in modo che gli oggetti e lo sfondo abbiano dei valori di intensità raggruppati in due mode principali.

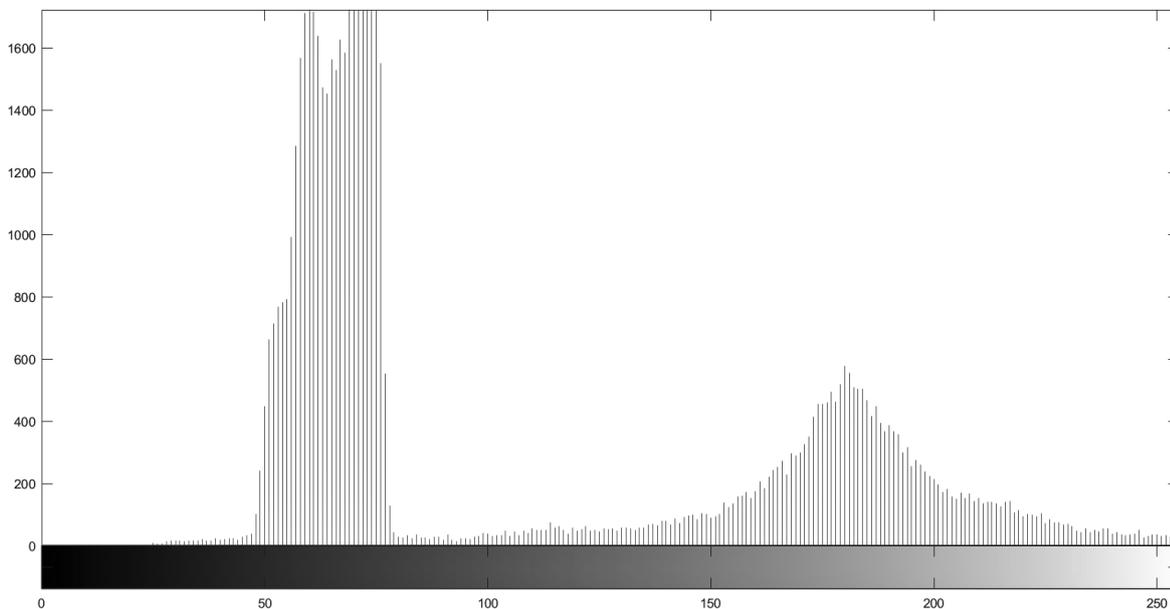


Figura 3.2: *Istogramma dell'intensità che può essere partizionato da una sola soglia.*

Per estrarre gli oggetti si può scegliere una soglia  $T$  che permette di separare queste mode. Successivamente ogni punto  $(x, y)$  nell'immagine che produce  $f(x, y) > T$  viene chiamato *punto dell'oggetto*, mentre se  $f(x, y) < T$  viene chiamato *punto dello sfondo*. Praticamente l'immagine segmentata

$g(x, y)$  sarà data da

$$g(x, y) = \begin{cases} 1, & \text{se } f(x, y) > T \\ 0, & \text{se } f(x, y) \leq T \end{cases}$$

Quando  $T$  è una costante applicabile all'intera immagine, il processo viene definito *sogliatura globale*.

Se invece  $T$  cambia nel corso del processo si usa l'espressione *sogliatura variabile*. Si ha invece la *sogliatura regionale* o *locale* quando si denota una sogliatura variabile il cui valore dipende dalle proprietà dell'intorno del punto considerato. Nel caso in cui avessimo, nell'istogramma dell'immagine, tre o più mode sarebbe necessario l'utilizzo di due o più soglie per avere dei risultati migliori. Questo può succedere, ad esempio nel caso di tre mode, nel caso di due tipi di oggetti chiari su uno sfondo scuro. In questo caso la *sogliatura multipla* viene effettuata in questo modo

$$g(x, y) = \begin{cases} a, & \text{se } f(x, y) > T_2 \\ b, & \text{se } T_1 < f(x, y) \leq T_2 \\ c, & \text{se } f(x, y) \leq T_1 \end{cases}$$

dove  $a, b$  e  $c$  sono tre valori distinti dell'intensità. Il successo della sogliatura dipende dall'altezza delle mode e dalle profondità delle valli.

### 3.2.1.1 Sogliatura globale

Come già accennato, quando le distribuzioni di intensità dei pixel dello sfondo e degli oggetti sono abbastanza distinti, si può utilizzare una soglia singola (sogliatura globale), applicabile a tutta l'immagine.

La sogliatura globale è un approccio molto comune ma risulta evidente la

necessità di dover stimare automaticamente il valore di soglia per ogni immagine.

Per fare ciò può essere usato il seguente algoritmo iterativo:

1. Si stima un valore iniziale della soglia globale  $T$ .
2. Si segmenta l'immagine usando la soglia appena stimata, il risultato di questa operazione darà come risultato due gruppi di pixel  $G_1$ , che comprende tutti i pixel i cui valori di intensità sono sopra soglia, e  $G_2$  in cui i pixel hanno un'intensità minore della soglia  $T$ .
3. Si calcolano i valori di intensità media  $m_1$  ed  $m_2$  per i rispettivi gruppi di pixel.
4. Si calcola un nuovo valore di soglia

$$T = \frac{(m_1 + m_2)}{2}.$$

5. Si ripetono i passi che vanno dal 2 al 4 finchè la differenza tra i valori assegnati a  $T$  non risulti minore di un parametro definito  $\Delta T$ .

Il semplice algoritmo appena descritto è molto efficace in situazioni in cui si ha un divario netto tra le mode nell'istogramma legate agli oggetti e lo sfondo.

L'intensità media dell'immagine risulta un buon punto di partenza per stimare la soglia iniziale  $T$ .

### 3.2.1.2 OTSU

La sogliatura può essere intesa come un problema teorico di statistica, il cui obiettivo è quello di minimizzare l'errore medio ottenuto dividendo i pixel in due o più gruppi (dette anche *classi*). Questo problema ha una soluzione in forma chiusa, conosciuta anche come *regola di decisione di Bayes*, che si basa solo su due parametri: la funzione di densità delle probabilità (PDF, *Probability Density Function*) dei livelli di intensità di ogni singola classe e la probabilità che ciascuna classe ha di essere presente in una data applicazione. Purtroppo la stima della PDF non è semplice e il problema viene semplificato attraverso delle assunzioni sulla forma della PDF (i.e. gaussiana). Il *metodo di Otsu* è una valida alternativa, è ottimale, cioè massimizza la varianza interclasse, una nota misura utilizzata nell'analisi discriminante statistica. Si basa sul fatto che classi ben sogliate possono essere distinte rispetto ai valori di intensità dei loro pixel, e viceversa, una soglia riuscendo ad ottenere la migliore separazione tra le classi rispetto ai valori di intensità, risulterà la migliore. Il metodo di Otsu si basa esclusivamente sull'istogramma dell'immagine, cioè su una matrice monodimensionale facilmente calcolabile. Ipotizziamo che un'immagine digitale di  $M \times N$  pixel, abbia  $L$  differenti livelli di intensità e con  $n_i$  indichiamo il numero di pixel di intensità  $i$ . Il numero totale  $M \cdot N$  di pixel dell'immagine è  $M \cdot N = n_0 + n_1 + n_2 + \dots + n_{L-1}$ .

L'istogramma normalizzato ha componenti  $p_i = \frac{n_i}{MN}$  da cui si ha

$$\sum_{i=0}^{L-1} p_i = 1, \quad p_i \geq 0.$$

Ora ipotizziamo di selezionare una soglia  $T(k) = k$ ,  $0 < k < L - 1$ , e

di utilizzarla come soglia sull'immagine di input ottenendo due classi  $C_1$  e  $C_2$ . La prima contiene tutti i pixel aventi valore compreso nell'intervallo  $[0, k]$  mentre la seconda nell'intervallo  $[k + 1, L - 1]$ .

Usando questa soglia la probabilità  $P_1(k)$  che un pixel venga assegnato alla classe  $C_1$  è data dalla somma cumulativa

$$P_1(k) = \sum_{i=0}^k p_i,$$

in altre parole non è altro che la probabilità di occorrenza della classe  $C_1$ .

Analogamente la probabilità di esistenza della classe  $C_2$  è data da

$$P_2(k) = \sum_{k+1}^{L-1} p_i = 1 - P_1(k).$$

Il valore di intensità medio dei pixel appartenenti a  $C_1$  è

$$m_1(k) = \sum_{i=0}^k iP(i/C_1) = \sum_{i=0}^k iP(C_1/i) \frac{P(i)}{P(C_1)} = \frac{1}{P_1(k)} \sum_{i=0}^k ip_i.$$

Dove il termine  $P\left(\frac{i}{C_1}\right)$  è la probabilità del valore  $i$ , dal momento che  $i$  deriva dalla classe  $C_1$ . Nel secondo passaggio dell'ultima espressione è stato fatto uso della formula di Bayes sulla probabilità condizionata:

$$P(A/B) = P(B/A) \frac{P(A)}{P(B)},$$

mentre nell'ultimo passaggio si è sfruttato il fatto che la probabilità di  $C_1$  dato  $i$ ,  $P(C_1/i)$ , è uguale a 1 dal momento che si ha a che fare con i valori di  $i$  della classe  $C_1$ , infine  $P(i)$  è semplicemente la probabilità dell' $i$ -esimo valore, che corrisponde alla  $i$ -esima componente dell'istogramma  $p_i$ .

Analogamente per il valore di intensità medio della seconda classe è

$$m_2(k) = \sum_{i=k+1}^{L-1} iP(i/C_2) = \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} ip_i.$$

Mentre la media cumulativa fino al livello  $k$  è data da

$$m(k) = \sum_{i=0}^k ip_i,$$

e la media globale, l'intensità media dell'intera immagine, è

$$m_G = \sum_{i=0}^{L-1} ip_i.$$

Deve valere che

$$P_1 m_1 + P_2 m_2 = m_G, \quad P_1 + P_2 = 1.$$

L'efficienza della soglia al livello  $k$  può essere espressa come

$$\eta = \frac{\sigma_B^2}{\sigma_G^2},$$

dove  $\sigma_G^2$  è la *varianza globale* e  $\sigma_B^2$  è la *varianza interclasse*, che sono definite come segue:

$$\begin{aligned} \sigma_G^2 &= \sum_{i=0}^{L-1} (i - m_G)^2 p_i, \\ \sigma_B^2 &= P_1 (m_1 - m_G)^2 + P_2 (m_2 - m_G)^2 = \\ &= P_1 P_2 (m_1 - m_2)^2 = \frac{(m_G P_1 - m)^2}{P_1 (1 - P_1)}. \end{aligned}$$

Da notare che  $\sigma_B^2$  cresce all'aumentare della distanza tra le due medie  $m_1$  ed  $m_2$ , mettendo in evidenza che la varianza interclasse è la misura della *separabilità* tra le classi. Essendo la varianza globale,  $\sigma_G^2$ , una costante, si ha che anche il rendimento  $\eta$  è una misura della separabilità e massimizzarlo vuol dire massimizzare  $\sigma_B^2$ . Quindi l'obiettivo è trovare il valore di soglia  $k$  che massimizzi la varianza interclasse. Otteniamo allora

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2},$$

e

$$\sigma_B^2(k) = \frac{(m_G P_1(k) - m(k))^2}{P_1(k)(1 - P_1(k))}.$$

Quindi la soglia ottimale è  $k^*$  che massimizza  $\sigma_B^2(k)$ , ovvero

$$\sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k).$$

Una volta che è stato trovato  $k^*$  l'immagine viene segmentata nel modo seguente

$$g(x, y) = \begin{cases} 1, & \text{se } f(x, y) > k^* \\ 0, & \text{se } f(x, y) \leq k^* \end{cases}$$

L'algoritmo può essere esposto come segue:

1. Calcolare l'istogramma normalizzato dell'immagine, denotando le componenti  $p_i$ ,  $i = 0, 1, 2, \dots, L - 1$ .
2. Calcolare le somme cumulative  $P_1(k)$  per  $k = 0, 1, 2, \dots, L - 1$ .
3. Calcolare le medie cumulative  $m(k)$ .
4. Calcolare l'intensità globale media  $m_G$ .

5. Calcolare la varianza interclasse  $\sigma_B^2$ .
6. Calcolare la soglia si Otsu  $k^*$ , cioè il valore di  $k$  che rende massima la varianza interclasse. Se il massimo non è unico, ottenere  $k^*$  facendo la media dei valori di  $k$  corrispondenti ai differenti massimi calcolati.
7. Ottenere la misura di separabilità  $\eta^*$ .

### 3.2.1.3 Soglie multiple

Il metodo di Otsu può essere esteso ad un numero arbitrario di soglie, dal momento che la misura di separabilità su cui si basa può essere estesa a un numero arbitrario di classi. Nel caso di  $K$  classi abbiamo

$$\sigma_B^2 = \sum_{k=1}^K P_k (m_k - m_G)^2,$$

dove

$$P_k = \sum_{i \in C_k} p_i,$$

e

$$m_k = \frac{1}{P_k} \sum_{i \in C_k} i p_i,$$

ed  $m_G$  è la media globale calcolata allo stesso modo del caso a soglia singola. Le  $K$  classi sono separate da  $K - 1$  soglie i cui valori  $k_1^*, k_2^*, \dots, k_{K-1}^*$ , massimizzano l'equazione

$$\sigma_B^2(k_1^*, k_2^*, \dots, k_{K-1}^*) = \max_{0 \leq k_1 < k_2 < \dots < k_{K-1} \leq L-1} \sigma_B^2(k_1, k_2, \dots, k_{K-1}).$$

Quando il numero delle classi aumenta, il significato di questo risultato sparisce perchè si ha a che fare con una sola variabile, l'intensità. Infatti

questo approccio viene utilizzato soltanto quando si pensa che il problema possa essere risolto con due soglie, per più di due soglie si usano anche altre informazioni oltre ai valori di intensità. Bisogna utilizzare dei descrittori aggiuntivi come ad esempio il colore, e trattare l'applicazione come un problema di pattern recognition. Ad esempio per tre classi che comprendono tre intervalli di intensità, abbiamo che la varianza interclasse è data da

$$\sigma_B^2 = P_1 (m_1 - m_G)^2 + P_2 (m_2 - m_G)^2 + P_3 (m_3 - m_G)^2,$$

dove

$$P_1 = \sum_{i=0}^{k_1} p_i, \quad P_2 = \sum_{i=k_1+1}^{k_2} p_i, \quad P_3 = \sum_{i=k_2+1}^{L-1} p_i;$$

e

$$m_1 = \frac{1}{P_1} \sum_{i=0}^{k_1} ip_i, \quad m_2 = \frac{1}{P_2} \sum_{i=k_1+1}^{k_2} ip_i, \quad m_3 = \frac{1}{P_3} \sum_{i=k_2+1}^{L-1} ip_i.$$

E valgono le seguenti

$$P_1 m_1 + P_2 m_2 + P_3 m_3 = m_G, \quad P_1 + P_2 + P_3 = 1.$$

Le due soglie ottimali  $k_1^*$  e  $k_2^*$  sono quei valori che massimizzano  $\sigma_B^2(k_1, k_2)$ . Quindi le soglie ottimali si trovano individuando

$$\sigma_B^2(k_1^*, k_2^*) = \max_{0 < k_1 < k_2 < L-1} \sigma_B^2(k_1, k_2).$$

Si inizia quindi individuando il primo valore  $k_1$  che è uguale a 1 perchè altrimenti non avrebbe senso cercare una soglia con intensità 0. Successivamente si incrementa  $k_2$  tra i valori maggiori di  $k_1$  e minori di  $L-1$ . Poi viene aumentato  $k_1$  e ripetuta la procedura, continuando finchè non  $k_1 = L-3$ . Il risultato sarà una matrice bidimensionale  $\sigma_B^2(k_1, k_2)$  nella quale bisogna in-

dividuare il massimo. Una volta trovati  $k_1^*$  e  $k_2^*$  che identificano il massimo si passa alla segmentazione

$$g(x, y) = \begin{cases} a, & \text{se } f(x, y) \leq k_1^* \\ b, & \text{se } k_1 < f(x, y) \leq k_2^* \\ c, & \text{se } f(x, y) > k_2^* \end{cases}$$

dove  $a, b$  e  $c$  corrispondono a tre diversi valori di intensità. La misura di separabilità è data da

$$\eta(k_1^*, k_2^*) = \frac{\sigma_B^2(k_1^*, k_2^*)}{\sigma_G^2},$$

dove  $\sigma_G^2$  è sempre la varianza globale calcolata nel modo consueto definito precedentemente.

### 3.2.2 Metodi basati sull'individuazione di edge, linee, punti

Le caratteristiche delle immagini utilizzate da questi metodi sono appunto gli edge, le linee e i punti. I *pixel edge* sono pixel la cui funzione di intensità nell'immagine cambia all'improvviso e gli *edge* sono insiemi di pixel di edge connessi. Gli *edge detector* (individuatori di edge) sono metodi progettati per individuare i pixel di edge. Una linea può essere vista come un segmento di edge in cui l'intensità dello sfondo su entrambi i lati della linea è maggiore o minore dell'intensità dei pixel della linea. Le derivate prime e seconde sono particolarmente adatte a rivelare brusche variazioni locali di intensità. In elaborazione digitale le derivate vengono definite in termini di differenze. Possiamo ottenere un'approssimazione della derivata prima in un punto  $x$  della funzione  $f(x)$  attraverso l'espansione di  $f(x + \Delta x)$  in

serie di Taylor in  $x$ , ponendo  $\Delta x = 1$  e tenendo conto soltanto dei termini lineari, il risultato è la differenza digitale

$$\frac{\partial f}{\partial x} = f'(x) = f(x+1) - f(x).$$

In questo caso si usa la derivata parziale in quanto la funzione ha due variabili  $f(x, y)$ . Per la derivata seconda invece, differenziando la precedente rispetto a  $x$  otteniamo

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial f'}{\partial x} = f'(x+1) - f'(x) = \\ &= f(x+2) - f(x+1) - f(x+1) + f(x) = \\ &= f(x+2) - 2f(x+1) + f(x),\end{aligned}$$

questa espressione vale nel punto  $x+1$ , dal momento che a noi interessa il valore della derivata in  $x$ , sottraiamo 1 dagli argomenti delle espressioni precedenti ottenendo

$$\frac{\partial^2 f}{\partial x^2} = f''(x) = f(x+1) + f(x-1) - 2f(x).$$

In Figura 3.3 è mostrata[2] un'immagine con il suo profilo di intensità misurato orizzontalmente verso il centro dell'immagine, che include il punto di rumore isolato. In basso si hanno un profilo semplificato, la linea dell'immagine corrisponde al profilo di intensità ed i numeri sono i valori che l'intensità assume, poi sono riportati i valori delle derivate prima e seconda. L'individuazione di punti isolati coinvolge quindi la derivata seconda, per le linee dovremmo aspettarci che le derivate seconde diano una risposta più forte e producano linee più sottili rispetto alle derivate prime, quindi

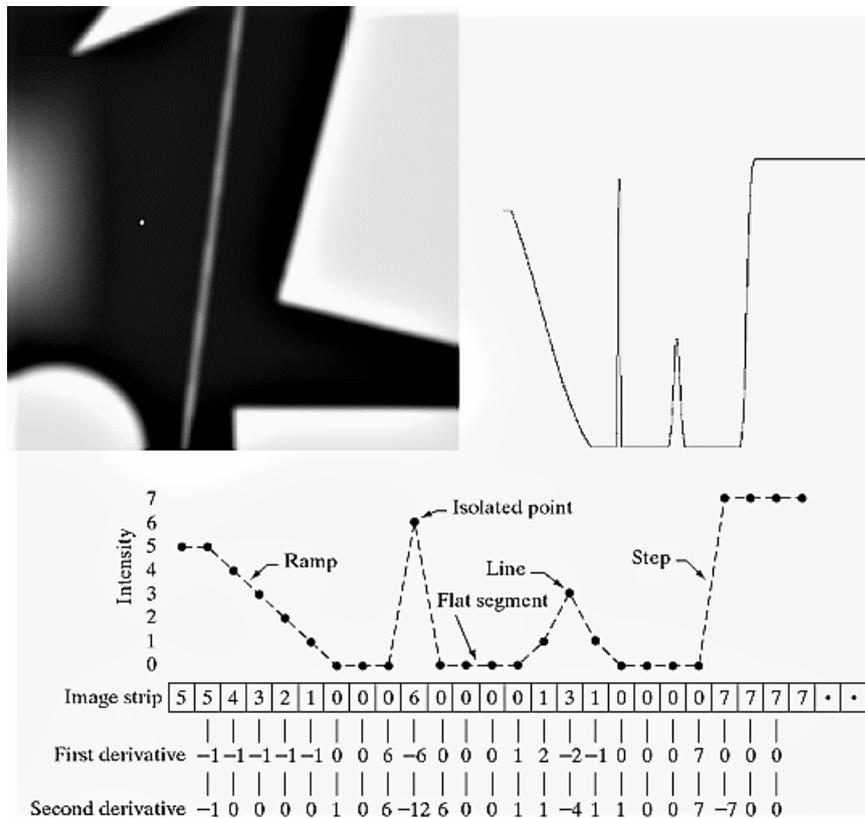


Figura 3.3: *Rappresentazione di un'immagine, suo profilo di intensità e derivate prima e seconda.*

si potrebbe usare una maschera laplaciana. L'individuazione di edge è il metodo più utilizzato per segmentare immagini caratterizzate da bruschi cambiamenti di intensità.

### 3.2.2.1 Laplaciano

Il laplaciano è definito come

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

dove le derivate parziali sono calcolate utilizzando le formule introdotte prima, quindi risulta che

$$\nabla^2 f(x, y) = f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 4f(x, y).$$

Per implementare questa espressione possono essere utilizzate delle maschere per pesare le differenze tra un pixel e l'altro, alcune maschere possono essere le seguenti

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

(a)

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}, \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}, \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}.$$

*Orizzontale*                       $+45^\circ$                       *Verticale*                       $-45^\circ$

Dove la prima maschera, (a), è indicata per la ricerca dei punti isolati, mentre le altre individuano delle linee in direzioni specifiche.

### 3.2.2.2 Operatori di gradiente: Sobel, prewitt, roberts

L'individuazione degli edge ha tre passi fondamentali:

1. *Applicare lo smoothing per la riduzione del rumore.*
2. *Individuazione dei punti di edge.* Si tratta di un'operazione locale che estrae da un'immagine tutti i punti che potenzialmente possono essere di edge.
3. *Localizzazione degli edge.* Tra i punti di edge candidati si selezionano quelli che sono veri membri dell'insieme di punti che comprendono un edge.

Lo strumento adatto a definire l'intensità e la direzione dell'edge nella posizione  $(x, y)$  di un'immagine  $f$  è il gradiente, definito come il vettore

$$\nabla f = \text{grad}(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}.$$

Il gradiente punta nella direzione della maggiore variazione di  $f$  nella posizione  $(x, y)$ . La magnitudio, lunghezza, del vettore gradiente è data da

$$M(x, y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2},$$

e rappresenta il valore del tasso di variazione nella direzione del vettore gradiente. Mentre  $g_x, g_y$  ed  $M(x, y)$  sono immagini della stessa dimensione di  $f(x, y)$ . La direzione del vettore gradiente è data dall'angolo

$$\alpha(x, y) = \tan^{-1} \left[ \frac{g_y}{g_x} \right],$$

rispetto all'asse  $x$ .  $M(x, y)$  è detta immagine gradiente. Per ottenere il gradiente di un'immagine si richiede il calcolo delle derivate parziali, che avendo a che fare con quantità digitali si utilizzano le approssimazione delle derivate parziali in un intorno del punto. Da quanto esposto nei paragrafi precedenti segue che

$$g_x = \frac{\partial f(x, y)}{\partial x} = f(x + 1, y) - f(x, y),$$

$$g_y = \frac{\partial f(x, y)}{\partial y} = f(x, y + 1) - f(x, y).$$

Quando ci interessa la direzione diagonale di un edge è necessario applicare una maschera bidimensionale, i primi tentativi furono gli operatori di gradiente a croce di Roberts che utilizzano maschere 2D con preferenza diagonale

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

*Roberts*

Gli operatori di Roberts si basano sull'implementazione delle differenze diagonali, facendo riferimento alla seguente regione  $3 \times 3$

$$\begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix}$$

e utilizzando le maschere  $2 \times 2$  appena definite, abbiamo

$$g_x = \frac{\partial f}{\partial x} = (z_9 - z_5), \quad g_y = \frac{\partial f}{\partial y} = (z_8 - z_6);$$

queste derivate possono essere implementate semplicemente filtrando l'im-

magine con le maschere. Le maschere di dimensioni  $2 \times 2$  sono concettualmente semplici, ma non sono utili per calcolare la direzione in quanto non sono simmetriche rispetto al punto centrale come lo sono le maschere  $3 \times 3$ . Queste ultime veicolano maggiori informazioni riguardo alla direzione. Per gli operatori di *Prewitt* si utilizzano le seguenti maschere

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

*Prewitt*

ottenendo

$$g_x = \frac{\partial f}{\partial x} = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3),$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7).$$

Una lieve variante di *Prewitt* utilizza un valore 2 nel coefficiente centrale, queste maschere vengono chiamate *operatori di Sobel*:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

*Sobel*

e si ha che

$$g_x = \frac{\partial f}{\partial x} = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3),$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7).$$

Le maschere di *Prewitt* sono più semplici da implementare ma la lieve diffe-

renza computazionale fa preferire le maschere di Sobel che hanno migliori prestazioni. Da notare che i coefficienti di tutte le maschere hanno somma 0, quindi danno una risposta nulla in aree a intensità costante, come previsto da un operatore derivativo.

### 3.2.2.3 Log

I metodi di individuazione degli edge visti fin orasi basano sul semplice filtraggio di immagini con una o più maschere, adesso ci occuperemo di valutare anche fattori come il rumore e la natura degli edge stessi. Marr e Hildreth intuirono che le variazioni di intensità sono dipendenti dalla scala dell'immagine e quindi la loro individuazione richiede l'uso di operatori di dimensioni diverse, e un cambiamento di intensità improvviso darà origine ad un picco sulla derivata prima o equivalentemente a un zero crossing nella derivata seconda. Un operatore dovrebbe avere quindi due caratteristiche, essere differenziale capace di calcolare un'approssimazione della derivata prima e seconda. Poi dovrebbe poter essere regolato per agire ad ogni scala selezionata, in modo tale che gli operatori più grandi possano essere usati per individuare gli edge più sfocati e gli operatori più piccoli per individuare i dettagli più piccoli e poco visibili. Marr e Hildreth intuirono che il miglior operatore che rispettasse queste condizioni fosse il filtro  $\nabla^2 G$  dove  $\nabla^2$  è l'operatore laplaciano  $\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$  e  $G$  è la funzione gaussiana bidimensionale

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$$

con  $\sigma$  deviazione standard. Per trovare un'espressione per tale filtro applichiamo

$$\nabla^2 G(x, y) = \frac{\partial^2 G(x, y)}{\partial x^2} + \frac{\partial^2 G(x, y)}{\partial y^2} =$$

$$\begin{aligned}
 &= \frac{\partial}{\partial x} \left( -\frac{x}{\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \right) + \frac{\partial}{\partial y} \left( -\frac{y}{\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \right) = \\
 &= \left( \frac{x^2}{\sigma^4} - \frac{1}{\sigma^2} \right) e^{-\frac{x^2+y^2}{2\sigma^2}} + \left( \frac{y^2}{\sigma^4} - \frac{1}{\sigma^2} \right) e^{-\frac{x^2+y^2}{2\sigma^2}};
 \end{aligned}$$

mettendo insieme i due termini si ottiene

$$\nabla^2 G(x, y) = \left( \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right) e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Questa espressione viene detta *Laplacian of the Gaussian (LoG)*, *laplaciano del gaussiano*. In Figura 3.4 possiamo notare[2] in a) una rappresentazione 3-D del negativo del LoG, b) negativo del LoG visualizzato come immagine, c) sezione di a) che mostra gli zero crossing, d) maschera di approssimazione  $5 \times 5$  alla forma in a). In pratica si usa il negativo di questa maschera.

L'algoritmo di individuazione degli edge di Marr-Hildreth può essere esposto brevemente come segue.

1. Filtrare l'immagine di input con un filtro passa basso gaussiano  $n \times n$  ottenuto dal campionamento della funzione  $G(x, y) e^{-\frac{x^2+y^2}{2\sigma^2}}$ .
2. Calcolare il laplaciano dell'immagine che risulta dal passo 1 usando ad esempio la maschera  $3 \times 3$  del laplaciano esposta poc'anzi.
3. Trovare gli zero crossing dell'immagine dal passo 2.

#### 3.2.2.4 Canny

Anche se l'algoritmo è più complesso, il metodo di *Canny*, è superiore a tutti gli altri approcci.

Si basa su tre obiettivi di base.

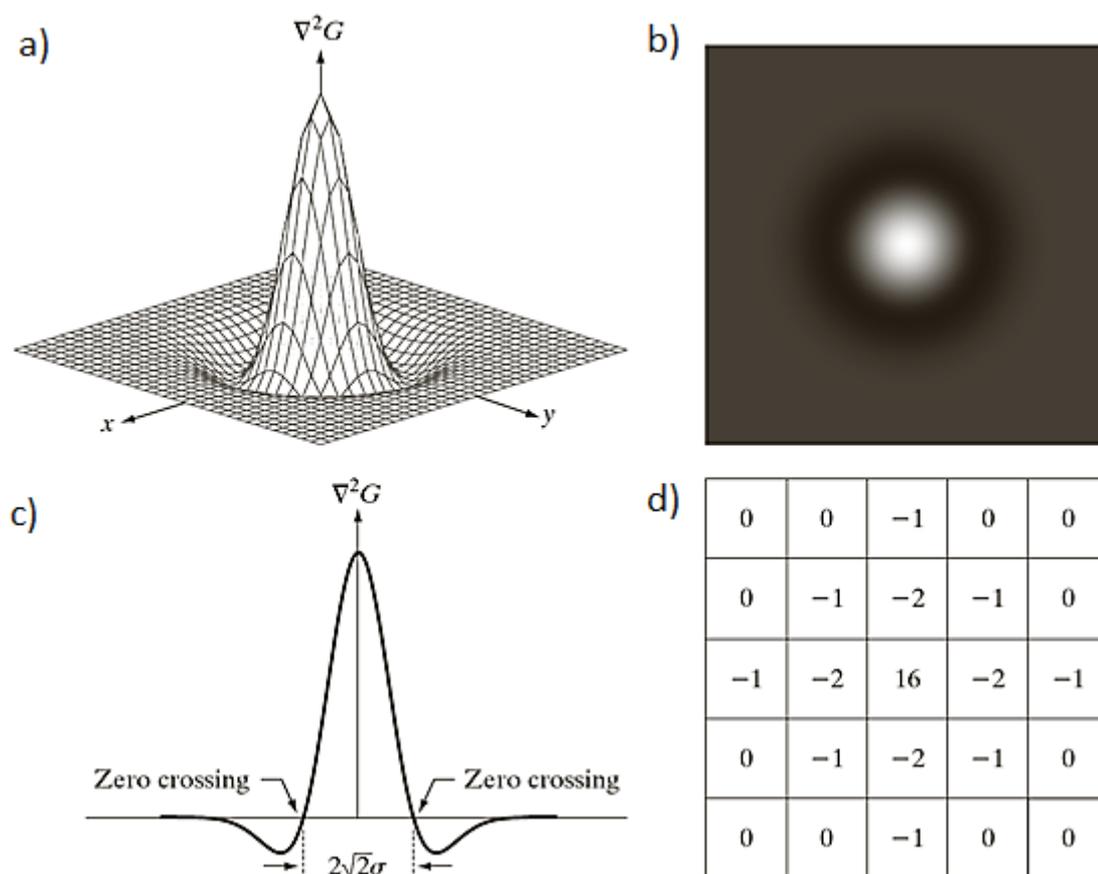


Figura 3.4: *Rappresentazione del negativo del LoG.*

1. *Basso tasso di errore.* Tutti gli edge devono essere trovati e non esistono risposte spurie. Ovvero gli edge individuati devono essere più vicini possibile ai veri edge.
2. *I punti di edge devono essere ben localizzati.* La distanza tra un punto segnato come edge e il centro del vero edge deve essere minima.
3. *Risposta puntuale per edge singolo.* Per ogni punto di un edge reale, l'algoritmo deve restituire un solo punto. Questo vuol dire che non si dovrebbero individuare edge multipli dove esiste un singolo punto di edge.

L'uso dell'ottimizzazione numerica con edge a gradino monodimensionale corrotti da rumore bianco gaussiano additivo porta alla conclusione che una buona approssimazione per individuare questi edge è la *derivata prima del gaussiano*:

$$\frac{d}{dx} e^{-\frac{x^2}{2\sigma^2}} = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}.$$

Per generalizzare il risultato in 2-D bisogna ammettere che il metodo 1-D continui a valere nella direzione normale dell'edge. Per fare ciò, essendo sconosciuta a priori la direzione normale, è necessario applicare l'individuatore di edge monodimensionale in tutte le direzioni. Questo può essere approssimato applicando prima un processo di smoothing con una funzione gaussiana circolare. Indicando con  $f(x, y)$  l'immagine in ingresso e con  $G(x, y)$  la funzione gaussiana

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}},$$

produciamo un'immagine sfocata  $f_S$  operando la convoluzione di  $G$  con  $f$

$$f_S(x, y) = G(x, y) * f(x, y).$$

Successivamente calcoliamo la magnitudo e la direzione (l'angolo) del gradiente

$$M(x, y) = \sqrt{g_x^2 + g_y^2}, \quad \alpha(x, y) = \tan^{-1} \left[ \frac{g_y}{g_x} \right],$$

dove  $g_x$  e  $g_y$  sono le derivate parziali di  $f_S$  rispetto a  $x$  e  $y$ . Dato che  $M(x, y)$  è stata generata con il gradiente, contiene delle "creste" attorno ai massimi locali: per eliminarle bisogna effettuare un assottigliamento di tali valori facendo uso del metodo *non maxima suppression*.

Questo metodo può essere così riassunto:

1. Trovare la direzione  $d_k$  che è più vicina ad  $\alpha(x, y)$ .
2. Se il valore di  $M(x, y)$  è minore di almeno uno dei suoi due vicini lungo  $d_k$ , poniamo  $g_N(x, y) = 0$  (soppressione), altrimenti  $g_N(x, y) = M(x, y)$ , dove  $g_N(x, y)$  è l'immagine finale.

L'operazione finale è sogliare  $g_N(x, y)$  per ridurre i falsi punti di edge. Se scegliamo una soglia troppo bassa ci saranno alcuni falsi edge (falsi positivi), se la soglia è troppo alta allora anche alcuni punti di edge validi saranno eliminati (falsi negativi). L'algoritmo di Canny tenta di migliorare questa situazione utilizzando la *thresholding mediante isteresi*, che usa due soglie, una bassa  $T_L$  e una alta  $T_H$ . Canny ha suggerito che il rapporto tra le due soglie dovrebbe essere di circa 2 o 3 a uno.

Possiamo visualizzare il thresholding creando due immagini aggiuntive

$$g_{NH}(x, y) = g_N(x, y) \geq T_H, \quad g_{NL}(x, y) = g_N(x, y) \geq T_L;$$

dove inizialmente entrambe sono settate a zero. Adesso  $g_{NH}$  avrà meno pixel non nulli di  $g_{NL}$ , ma tutti i pixel non nulli in  $g_{NH}$  sono contenuti in  $g_{NL}$ . Bisogna quindi eliminare da  $g_{NL}$  tutti i pixel non nulli di  $g_{NH}$  ponendo

$$g_{NL}(x, y) = g_{NL}(x, y) - g_{NH}(x, y).$$

I pixel diversi da zero in  $g_{NH}$  e  $g_{NL}$  rappresentano rispettivamente i pixel di edge forti e deboli. Dopo vengono considerati i pixel di edge validi e vengono etichettati come tali. In base al valore di  $T_H$ , gli edge in  $g_{NH}$  saranno tipicamente dei vuoti.

Riassumendo, l'algoritmo di individuazione degli edge di Canny consiste nei seguenti passi fondamentali.

1. Sottoporre a smoothing l'immagine di input con un filtro gaussiano.
2. Calcolare la magnitudo e l'angolo del gradiente.
3. Applicare la non maxima suppression all'immagine della magnitudo.
4. Utilizzare la doppia soglia e la connettività per individuare e collegare gli edge.

Di solito si fa seguire al passo 4 un algoritmo di assottigliamento degli edge, dal momento che alcuni degli edge spessi più di un pixel potrebbero ancora essere presenti.

### **3.2.3 Segmentazione basata sulle regioni: region growing**

Il region growing è una procedura che raggruppa i pixel o le sottoregioni in regioni via via più grandi basandosi su criteri predefiniti. Tutto inizia da un insieme di punti detti *seed* (seme) dai quali le regioni crescono aggiungendo ad ogni seme i pixel del suo contorno che possiedono delle proprietà simili a quelle del seme stesso. La selezione dei criteri di similarità dipende dalla tipologia di dati a disposizione e dal problema in questione. Ad esempio nelle immagini in scala di grigi si esegue l'analisi della regione su un insieme di descrittori basati sui livelli di intensità e sulle proprietà spaziali come i momenti o la tessitura. I descrittori devono essere usati insieme alle proprietà di connettività nel processo di crescita, altrimenti si potrebbero generare delle segmentazioni poco sensate. Un altro punto importante è il criterio di arresto, la crescita dovrebbe cessare quando non ci sono più pixel che soddisfano i requisiti per l'inclusione nella regione stessa. Supponiamo che  $f(x, y)$  sia la matrice dell'immagine di partenza,  $S(x, y)$  una matrice seme in cui nella posizione dei punti seme si ha il valore 1 e 0

nelle restanti,  $Q$  un predicato da verificare su ogni  $(x, y)$ . Siano  $f$  ed  $S$  della stessa dimensione, allora un algoritmo di region growing basato sulla  $8$ -connettività<sup>3.1</sup> può essere così descritto:

1. Trovare tutte le componenti connesse in  $S(x, y)$  ed erodere<sup>3.2</sup> ogni componente connessa ad un pixel, assegnare 1 a tali pixel e 0 ai rimanenti.
2. Formare un'immagine  $f_Q$  in modo che nel punto  $(x, y)$ , se l'immagine di input soddisfa il predicato  $Q$ , si abbia  $f_Q(x, y) = 1$ , altrimenti  $f_Q(x, y) = 0$ .
3. Sia  $g$  l'immagine che si è formata dall'aggiunta a ogni punto seme in  $S$  di tutti i punti di valore 1 in  $f_Q$  che sono 8-connessi al suddetto seed.
4. Marcare ogni componente connessa in  $g$  con un'etichetta diversa per ogni regione.

### 3.2.4 Segmentazione morfologica watershed

Il concetto del metodo watershed si basa sull'interpretazione di un'immagine come tridimensionale, ovvero le due coordinate spaziali e l'intensità che va a formare la terza dimensione, come illustrato in Figura 3.5 dove a sinistra si ha un'immagine generata artificialmente e a destra la sua visualizzazione in 3D secondo il watershed. In questa interpretazione tridimensionale si possono distinguere tre tipi di punti:

---

<sup>3.1</sup>Connettività: due pixel si dicono connessi, se sono adiacenti e se soddisfano un criterio di similarità del livello di grigio.

Intorno di un pixel: 4-intorno di  $p(x, y)$  ( $N_4(p)$ ) con coordinate  $(x, y)$  sono i pixel di coordinate  $(x + 1, y)$ ,  $(x, y + 1)$ ,  $(x - 1, y)$ ,  $(x, y - 1)$ ; 8-intorno di  $p(x, y)$  ( $N_8(p)$ ) include anche i 4 pixel diagonali.

<sup>3.2</sup>Per approfondimenti sull'erosione consultare il Paragrafo 9.2.1 di [2]

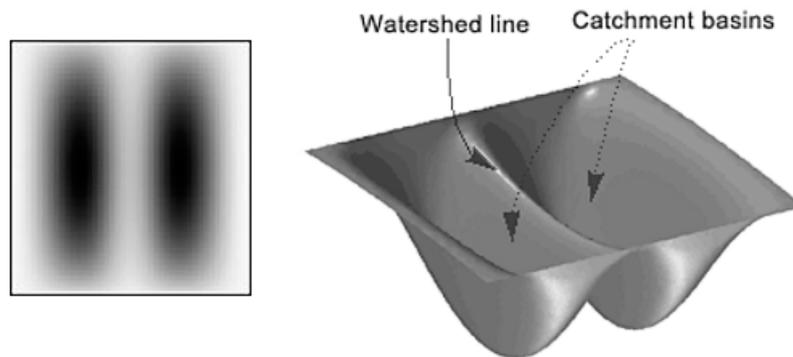


Figura 3.5: *Significato Watershed.*

1. punti che appartengono ad un minimo locale,
2. punti in cui una goccia d'acqua scorrerebbe verso uno dei minimi,
3. punti in cui l'acqua potrebbe egualmente cadere in più di un punto di minimo.

Per un dato minimo locale, l'insieme dei punti che soddisfano la condizione 2. è detto *bacino di raccolta*. Invece l'insieme dei punti che soddisfano la condizione 3. formano delle creste sulla superficie e vengono dette *linee di separazione* o *linee di watershed*. L'obiettivo principale è quello di identificare queste linee di watershed. Supponiamo che ogni minimo locale venga perforato e si faccia entrare dell'acqua, facendola risalire fino ai contorni. Quando l'acqua rischia di strabordare dai bacini rischiando di invadere gli altri viene costruita una diga al fine di prevenire tale evento. L'acqua continua a salire finché non rimangono visibili soltanto le cime di tali dighe. I contorni delle dighe costituiscono le linee di divisione watershed, si tratta di contorni connessi estratti dall'algoritmo di segmentazione. Un esempio[2] di questo procedimento è illustrato in Figura 3.6, dove partendo dalla prima figura in alto a sinistra si ha: *a)* l'immagine originale, *b)* visualizzazione

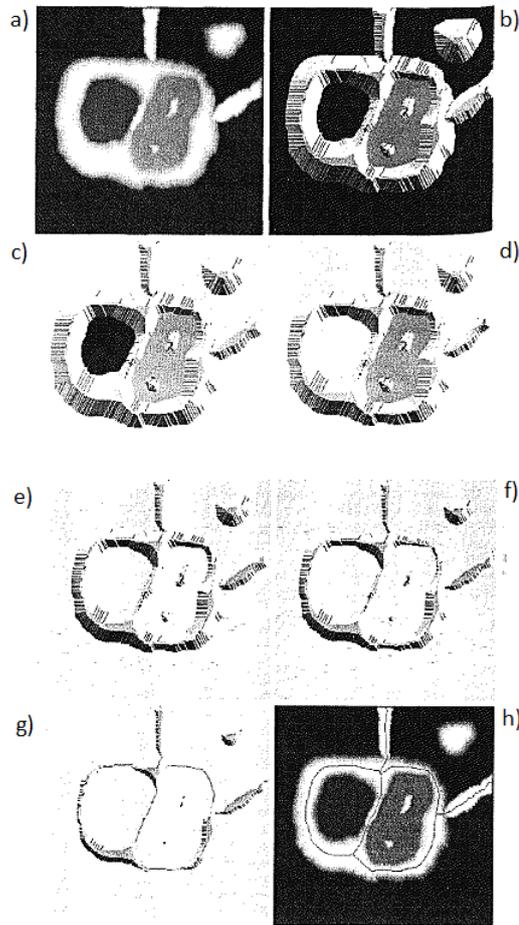


Figura 3.6: *Esempio di procedimento watershed.*

topografica, poi *c*) e *d*) due fasi dell'allagamento, *e*) ulteriore allagamento, *f*) principio di fusione dell'acqua dei due bacini e costruzione di una diga molto corta, *g*) costruzione di dighe più lunghe, *h*) linee di watershed finale. Indichiamo con  $M_1, M_2, \dots, M_R$  gli insiemi che indicano le *coordinate* dei punti nei minimi locali di un'immagine  $g(x, y)$ . Di solito  $g(x, y)$  è il gradiente dell'immagine di input, non viene utilizzata direttamente l'immagine di partenza. Indichiamo con  $C(M_i)$  l'insieme delle coordinate dei punti nei bacini di raccolta associati all' $i$ -esimo minimo  $M_i$ . Da notare che i punti nel bacino di raccolta formano una componente connessa. Supponiamo in-

fine che  $T(n)$  sia l'insieme delle coordinate  $(s, t)$  tali per cui  $g(s, t) < n$ , ovvero

$$T(n) = \{(s, t) | g(s, t) < n\}.$$

Geometricamente è l'insieme dei punti in  $g(x, y)$  che si trovano al disotto del piano individuato da  $g(x, y) = n$ . Indichiamo con  $C_n(M_i)$  l'insieme delle coordinate dei punti nel bacino di raccolta associati al minimo  $M_i$ , che vengono allagati al passo  $n$ -esimo.  $C_n(M_i)$  può essere considerata come un'immagine binaria data da

$$C_n(M_i) = C(M_i) \cap T(n).$$

Ovvero  $C_n(M_i) = 1$  nelle coordinate  $(x, y)$  se  $(x, y) \in C(M_i)$  AND  $(x, y) \in T(n)$ , mentre negli altri casi  $C_n(M_i) = 0$ . Praticamente al passo  $n$  del processo di inondazione si usa l'operatore AND per isolare la parte dell'immagine binaria in  $T(n)$  associata al minimo locale  $M_i$ . Successivamente indichiamo con  $C(n)$  l'unione dei bacini di raccolta inondati al passo  $n$

$$C(n) = \bigcup_{i=1}^R C_n(M_i),$$

e con  $C(\max\{g(x, y)\} + 1)$  l'unione di tutti i bacini di raccolta

$$C(\max\{g(x, y)\} + 1) = \bigcup_{i=1}^R C(M_i).$$

L'algoritmo per trovare le linee di watershed viene inizializzato con

$$C(\min\{g(x, y)\} + 1) = T(\min\{g(x, y)\} + 1),$$

poi procede calcolando  $C(n)$  a partire da  $C(n-1)$ . Indichiamo con  $Q$  l'insieme delle componenti connesse in  $T(n)$ , per ogni componente connessa  $q \in Q(n)$  si hanno tre possibilità

1.  $q \cap C(n-1)$  è vuoto,
2.  $q \cap C(n-1)$  contiene una componente connessa di  $C(n-1)$ ,
3.  $q \cap C(n-1)$  contiene più di una componente connessa di  $C(n-1)$ .

La condizione 1. si verifica quando ci imbattiamo in un nuovo minimo, ovvero nel caso in cui la componente connessa  $q$  viene incorporata in  $C(n-1)$  per formare  $C(n)$ . La condizione 2. si verifica quando  $q$  appartiene al bacino di raccolta di qualche minimo regionale, anche in questo caso la componente connessa di  $q$  viene incorporata in  $C(n-1)$  per formare  $C(n)$ . Infine la condizione 3. si verifica quando si ha una cresta che separa due bacini di raccolta. Una successiva inondazione provocherebbe lo straripamento dell'acqua da un bacino all'altro e quindi è necessaria la costruzione di una diga all'interno di  $q$  per prevenire lo straripamento. Si può costruire una diga di un pixel attraverso la dilatazione di  $q \cap C(n-1)$  con un elemento strutturante  $3 \times 3$  e limitare la dilatazione a  $q$ . L'efficienza dell'algoritmo può essere migliorata utilizzando solo valori di  $n$  che corrispondono ai valori di intensità esistenti in  $g(x, y)$ , così come accade per il minimo e il massimo.

### 3.2.5 Segmentazione basata sui contorni attivi

L'idea di base di questi metodi è di far evolvere una curva, soggetta a vincoli dettati dalle caratteristiche degli oggetti da individuare, deformandola da una configurazione iniziale fino al raggiungimento dei contorni degli oggetti. I contorni attivi possono essere rappresentati in maniera esplicita attraverso curve parametriche, noti come snakes, oppure come livelli di una funzione la cui evoluzione ne descrive implicitamente il movimento. Quest'ultimo è un approccio *level-set* che è stato introdotto da degli studi[5] di Osher e Sethian e poi sviluppato da altri (e.g. Chan-Vese). Questo metodo ha diversi campi di applicazione che vanno dalla fisica, alla grafica, alla computer vision. Il principio fondamentale dei modelli level set, mostrato

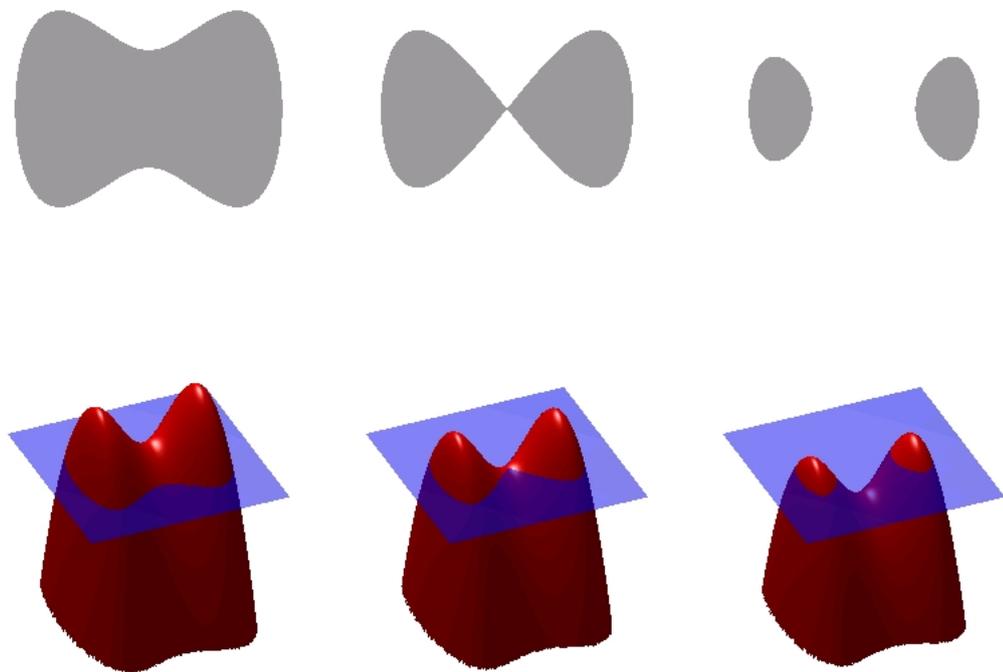


Figura 3.7: *Illustrazione del metodo Level-Set.*

in Figura 3.7, è quello di rappresentare un contorno come intersezione tra una superficie e il piano passante per il suo *livello zero*, quindi se la superficie varia allora cambierà anche il contorno intercettato da suddetto piano, per cui è possibile descrivere il movimento del contorno valutando l'evoluzione della superficie la cui sezione rappresenta il contorno. La tecnica level-set dell'equazione di evoluzione di una curva (2D) o di una superficie (3D) consente di risolvere alcuni problemi tipici dell'evoluzione di fronti, in particolare il problema dei cambiamenti di topologia. Sia  $\Gamma$  una curva chiusa (contorno nell'immagine di input) in uno spazio a due dimensioni e  $\Omega_{in}, \Omega_{out}$  le regioni interne ed esterne alla curva. Il contorno  $\Gamma$  può essere rappresentato attraverso una funzione, la cosiddetta *funzione level set*, in particolare in corrispondenza dello *zero level set*, questa funzione  $\phi$  assume valori negativi all'interno della curva (in  $\Omega_{in}$ ) e valori positivi all'esterno (in  $\Omega_{out}$ ) ed è data da

$$\begin{cases} \phi < 0, & \text{se } (x, y) \in \Omega_{in} \\ \phi = 0, & \text{se } (x, y) \in \Gamma \\ \phi > 0, & \text{se } (x, y) \in \Omega_{out} \end{cases}$$

In questo modo il livello zero della funzione,  $\phi = 0$ , costituisce l'insieme dei punto del contorno  $\Gamma$ . In pratica possiamo scrivere che  $\Gamma = \phi^{-1}(0)$ , l'idea è di sostituire l'evoluzione nel tempo di  $\Gamma$  con l'evoluzione nel tempo di  $\phi$ , che in generale è funzione anche del tempo  $\phi(x, y, t)$ . L'evoluzione della funzione level set può essere espressa attraverso la cosiddetta *equazione level set del moto* data da

$$\frac{\partial \phi}{\partial t} + \vec{V} \cdot \nabla \phi = 0,$$

dove  $\vec{V}$  è il vettore velocità costituito da una componente tangenziale e una normale  $\vec{V} = V_n \hat{n} + V_t \hat{t}$ . Il parametro su cui è necessario intervenire per controllare l'evoluzione della curva è la componente normale della velocità  $V_n$ . Possiamo avere tre modalità di evoluzione in base ad altrettanti tipi di velocità per la componente normale, la prima evoluzione a velocità costante ( $V_n = \pm 1$ ), poi evoluzione per curvatura  $V_n = -k$ , con  $k$  termine medio di curvatura. Infine si ha l'evoluzione per trasporto, caratterizzata dalla presenza di un campo vettoriale a velocità  $\vec{U}$ .

### 3.2.5.1 Chan-Vese

I modelli region based sono fondati sulla minimizzazione di un funzionale di energia. Ovvero consistono nel trovare un contorno ottimo  $C$  che divide il dominio  $\Omega$  dell'immagine input  $I$  in sottoregioni separate, e trovare una funzione ottima  $u$  che approssimi l'immagine  $I$ .

La soluzione si ottiene minimizzando rispetto a  $C$  e  $u$  il seguente funzionale

$$F(C, u) = \int_{\Omega} (I(x, y) - u(x, y))^2 dx dy + \int_{\Omega \setminus C} |\nabla u(x, y)|^2 dx dy + \mu |C|.$$

Il modello proposto da *Chan-Vese*[6] è uno dei region based e suppone che la funzione  $u$  che approssima l'immagine di partenza sia costante a tratti e che assuma soltanto due valori  $c_1, c_2$  che corrispondono ai valori medi dell'immagine  $I$  nelle due rispettive regioni separate dalla curva  $C$  nel dominio  $\Omega$ , indicando con  $C_{in}$  e  $C_{out}$  i domini interno ed esterno a  $C$  possiamo scrivere:

$$u(x, y) = \begin{cases} c_1, & \text{se } (x, y) \in C_{in} \\ c_2, & \text{se } (x, y) \in C_{out} \end{cases}$$

Quindi il funzionale proposto da Chan-Vese è

$$F(C, c_1, c_2) = \lambda_1 \int_{C_{in}} |I(x, y) - c_1|^2 dx dy + \\ + \lambda_2 \int_{C_{out}} |I(x, y) - c_2|^2 dx dy + \mu \cdot L_C + v \cdot A_{C_{in}},$$

dove con  $L_C$  si è indicata la lunghezza di  $C$  e con  $A_{C_{in}}$  l'area di  $C_{in}$ , mentre i parametri  $\mu, v, \lambda_1$  e  $\lambda_2$  sono fissati, in particolare  $\mu \geq 0, v \geq 0$  e  $\lambda_1 > 0, \lambda_2 > 0$ .

Gli ultimi due termini additivi servono a regolarizzare la lunghezza di  $C$  e l'area di  $C_{in}$ .

Secondo i metodi level set si può rappresentare il contorno  $C$  attraverso lo zero level della cosiddetta *funzione di Lipschitz*  $\phi$

$$\begin{cases} C = \partial\omega = \{(x, y) \in \Omega | \phi(x, y) = 0\} \\ C_{in} = \omega = \{(x, y) \in \Omega | \phi(x, y) > 0\} \\ C_{out} = \Omega \setminus \bar{\omega} = \{(x, y) \in \Omega | \phi(x, y) < 0\} \end{cases}$$

Sostituendo a  $C$  la funzione  $\phi$  e facendo uso del gradino di Heaviside e della delta di Dirac

$$H(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}, \quad \delta(x) = \frac{d}{dx} H(x) = \begin{cases} 1, & \text{se } x = 0 \\ 0, & \text{altrove} \end{cases},$$

il funzionale di energia  $F(\phi, c_1, c_2)$  secondo Chan-Vese si può esprimere

secondo il level set nel modo seguente

$$\begin{aligned}
 F(\phi, c_1, c_2) = & \lambda_1 \int_{\Omega} |I(x, y) - c_1|^2 H(\phi(x, y)) dx dy + \\
 & + \lambda_2 \int_{\Omega} |I(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy + \\
 & \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + v \int_{\Omega} H(\phi(x, y)) dx dy.
 \end{aligned}$$

Per motivi di fattibilità bisogna considerare delle versioni più regolari per Heaviside e delta di Dirac

$$H_{\varepsilon}(x) = \frac{1}{2} \left( 1 + \frac{2}{\pi} \arctan \left( \frac{x}{\varepsilon} \right) \right), \quad \delta_{\varepsilon} = \frac{d}{dx} H_{\varepsilon}(x) = \frac{1}{\pi} \frac{\varepsilon}{\varepsilon^2 + x^2};$$

per le quali vale  $H_{\varepsilon} \rightarrow H$  e  $\delta_{\varepsilon} \rightarrow \delta$  per  $\varepsilon \rightarrow 0$ ,  $\varepsilon \ll 1$  è detta costante di regolarizzazione. Andando a minimizzare il funzionale  $F_{\varepsilon}(\phi, c_1, c_2)$ , al cui interno sono state sostituite  $H$  e  $\delta$  con le rispettive regolarizzate, rispetto a  $c_1, c_2$  e  $\phi$ , siano  $\phi_0(x, y)$  una funzione level set iniziale e  $\bar{n}$  il versore normale uscente alla frontiera  $\partial\Omega$ , si ricavano le *equazioni di Eulero-Lagrange* associate al funzionale

$$\left\{ \begin{array}{l} \phi(0, x, y) = \phi_0(x, y) \text{ in } \Omega, \\ c_1(t) = \frac{\int_{\Omega} I(x, y) H_{\varepsilon}(\phi(x, y, t)) dx dy}{\int_{\Omega} H_{\varepsilon}(\phi(x, y, t)) dx dy}, \\ c_2(t) = \frac{\int_{\Omega} I(x, y) (1 - H_{\varepsilon}(\phi(x, y, t))) dx dy}{\int_{\Omega} (1 - H_{\varepsilon}(\phi(x, y, t))) dx dy} \\ \frac{\partial \phi}{\partial t} = \delta_{\varepsilon}(\phi) \left( -\lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2 + \mu \cdot \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - v \right) \text{ in } \Omega, \\ \frac{\delta_{\varepsilon}(\phi) \partial \phi}{|\nabla \phi| \partial n} = 0 \text{ su } \partial \Omega, \end{array} \right.$$

Le equazioni di solito sono approssimate grazie alle differenze finite, vengono risolte iterativamente fino al raggiungimento di uno stato stazionario nel quale si ottiene la funzione level set  $\phi$  che minimizza il funzionale energetico  $F$ .

Il modello Chan-Vese descritto è particolarmente adatto nella segmentazione di immagini caratterizzate da regioni statisticamente omogenee. In molti casi reali però (e.g. bio-immagini derivanti da MRI) tale assunzione non è plausibile in quanto vi è una disomogeneità dovuta ad artefatti o ad imperfezioni degli strumenti, per cui il modello risulta inutilizzabile.

### 3.2.5.2 Geodesic

Questa tecnica è edge based, ovvero basta sull'analisi dei contorni degli oggetti presenti nell'immagine, solitamente individuati attraverso il gradiente dell'immagine di input. L'algoritmo denominato *Geodesic active contour*[7] consiste in due passi consecutivi, il primo è definire una funzione  $g$  detta *edge indicator* basata sul gradiente dell'immagine di input  $I$ , allo scopo di riconoscere i contorni locali. Questa funzione blocca il processo

di evoluzione sul contorno dell'oggetto desiderato. La funzione  $g$  è definita a partire dalla convoluzione tra l'immagine di partenza e la distribuzione gaussiana, ottenendo uno smoothing dell'immagine di partenza

$$g(|\nabla I_0|) = \frac{1}{1 + |\nabla G_\sigma * I|^2}.$$

Il moto per la curvatura geodetica è definito da

$$\begin{cases} \phi_t = gk |\nabla \phi| + \nabla g \cdot \nabla \phi, & \text{in } \Omega \times (0, \infty) \\ \phi(x, y, t) = \min(\phi_0), & \text{in } \partial\Omega \times (0, \infty) \\ \phi(x, y, 0) = \phi_o(x, y), & \text{in } \Omega \end{cases}$$

dove il primo termine permette alla curva di muoversi con velocità pari alla curvatura  $k$  pesata, lungo la direzione normale. Mentre il secondo termine consiste nel moto di trasporto verso i contorni individuati da  $g$ .

### 3.2.5.3 Snakes

I metodi discussi fin ora erano di tipo level set, per quanto riguarda invece i metodi parametrici, detti appunto *snakes*, si basano sulla formulazione di una curva parametrica  $c$  contenuta nel piano  $(x, y)$  dell'immagine di partenza, definita come

$$c(s) = (x(s), y(s)),$$

dove  $x(s)$  e  $y(s)$  sono le coordinate lungo il contorno ed  $s = [0, 1]$  è il dominio parametrico. Associamo il funzionale energetico  $E_{Snake}$  alla curva

parametrica  $c$

$$E_{Snake}(c(s)) = \int_0^1 (E_{int} + E_{ext}) ds.$$

Dove  $E_{int}$  è l'energia di deformazione interna dello snake che caratterizza la sua estensibilità e flessibilità, possiamo esprimerla come

$$E_{int}(c(s)) = \alpha(s) \left| \frac{\partial c(s)}{\partial s} \right|^2 + \beta(s) \left| \frac{\partial^2 c(s)}{\partial s^2} \right|^2,$$

in cui il primo termine è dovuto a forze interne di tensione e definisce il comportamento elastico dello snake, mentre il secondo esprime la sua rigidità, ovvero la resistenza ai piegamenti. I parametri  $\alpha(s)$  e  $\beta(s)$  intervengono sull'allungarsi e il flettersi della curva.

Invece  $E_{ext}$  è una funzione potenziale esterna  $P(x, y)$  che dipende dall'immagine e dai vincoli esterni che guida lo snake verso determinate caratteristiche, ad esempio per attirarlo verso i contorni degli oggetti, dove il gradiente è molto alto, si può usare

$$P(x, y) = -\lambda |\nabla (G_\sigma(x, y) * I(x, y))|,$$

ovvero il gradiente della convoluzione tra un filtro gaussiana e l'immagine, uno smoothing, regolato in ampiezza attraverso il termine  $\lambda$ . Quindi una segmentazione di questo tipo può essere ricondotta al processo di minimizzazione dell'energia. Ad esempio una curva  $c(s)$  che minimizza il funzionale  $E_{Snake}(c(s))$  deve soddisfare l'equazione di *Eulero-Lagrange*

$$-\frac{\partial}{\partial s} \left( \alpha(s) \frac{\partial c}{\partial s} \right) + \frac{\partial^2}{\partial s^2} \left( \beta(s) \frac{\partial^2 c}{\partial s^2} \right) + \nabla E_{ext}(c(s)) = 0,$$

la quale descrive il bilancio di forze interne ed esterne dello snake quando il contorno raggiunge uno stato di equilibrio. In generale questa equazione differenziale alle derivate parziali si risolve con metodi numerici fino al contorno desiderato. Il metodo snakes deve essere inizializzato vicino all'oggetto di interesse e presenta il difetto di entrare in difficoltà con strutture di forma complessa. Questo limite può essere aggirato rappresentando il contorno attivo tramite level set.

#### 3.2.5.4 Fast Marching

Fa sempre parte dei metodi basati sul level set e ha delle somiglianze con lo snake classico, ovvero consiste in un contorno dinamico e può essere reso dipendente da dei criteri. I vantaggi rispetto a snake sono che è privo di marker, non ha bisogno che si specifichi un numero di step iterativi, gestisce gli angoli netti e i cambiamenti di tipologia come split e merge. L'idea di base è di discretizzare il dominio dell'immagine (lo zero level set) in una griglia di punti. Questo punti vengono classificati, come si vede in Figura 3.8, in base alla loro posizione rispetto al contorno in nodi *far*, nodi *accepted* e nodi *narrow band*. Un'idea del funzionamento dell'algoritmo è illustrato in Figura 3.9 dove è rappresentato l'avanzamento del fronte del contorno. I nodi *far* rappresentano i punti non ancora raggiunti dal fronte, i nodi *accepted* rappresentano la regione già raggiunta, mentre i nodi *narrow band* sono i prossimi candidati all'inclusione alla prossima iterazione.

Questo algoritmo si basa sulla risoluzione dell'equazione *Eikonale*

$$\begin{cases} |\nabla u(x)| = \frac{1}{f(x)}, & x \in \Omega \\ u(x) = 0, & x \in \partial\Omega \end{cases},$$

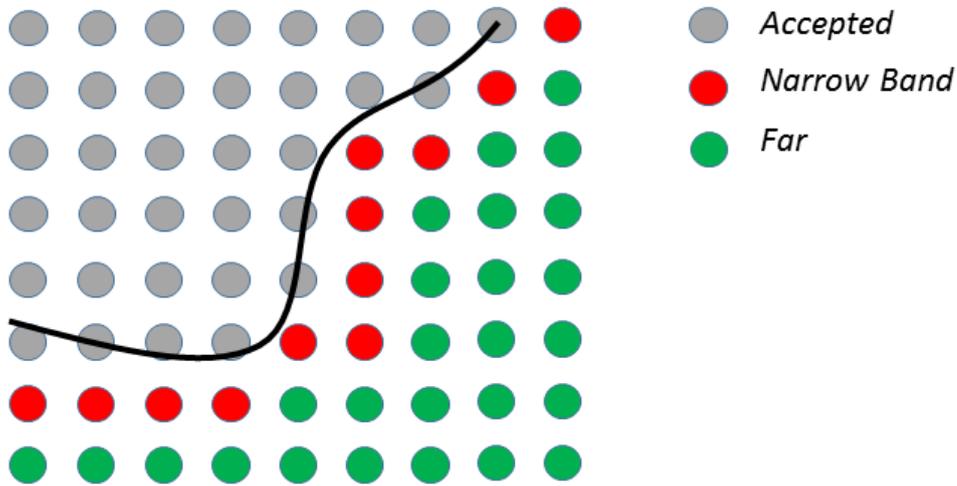


Figura 3.8: *Punti della griglia in Fast Marching.*

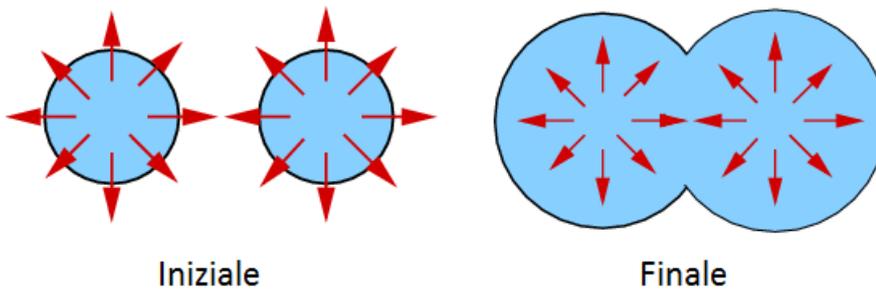


Figura 3.9: *Evoluzione del Fast Marching.*

che descrive l'evoluzione di una curva chiusa in funzione del tempo  $u$  con velocità  $f(x)$  in direzione normale al punto  $x$  nella superficie di propagazione.  $u(x)$  può essere pensata come la quantità di tempo minima impiegata per raggiungere  $\partial\Omega$  partendo dal punto  $x$ . Assumiamo che il dominio sia discretizzato e che ogni nodo  $x_i$  identifichi un valore  $U_i = U(x_i) \approx u(x_i)$ . Possiamo riassumere l'algoritmo in questo modo:

1. Ad ogni nodo  $x_i$  viene assegnato il valore  $U_i = +\infty$  ed etichettato come *far*, mentre pre tutti i nodi  $x_i \in \partial\Omega$  viene settato  $U_i = 0$  ed etichettati come *accepted*.

2. Per ogni nodo *far* si usa l'equazione *Eikonale* per ricalcolare il nuovo valore  $\tilde{U}$ , se  $\tilde{U} < U_i$  allora si pone  $U_i = \tilde{U}$  ed etichettato come *narrow band*.
3. Sia  $\tilde{x}$  il nodo *narrow band* con il valore minore di  $U$ ,  $\tilde{x}$  viene etichettato come *accepted*.
4. Per ogni  $x_i$  contiguo a  $\tilde{x}$  che non sia *accepted* si calcola  $\tilde{U}$ .
5. Se  $\tilde{U} < U_i$  allora  $U_i = \tilde{U}$ . Se  $x_i$  era etichettato come *far* allora verrà aggiornato a *narrow band*.
6. Se esistono ancora dei nodi *narrow band* si ritorna al punto 3. altrimenti finisce il loop.

Per calcolare  $\tilde{U}$  bisogna risolvere l'equazione *Eikonale* e una corretta approssimazione della soluzione di viscosità può essere data da

$$\left( \max \left\{ \max \left\{ \frac{U_{i,j} - U_{i-1,j}}{\Delta x}, 0 \right\}, -\min \left\{ \frac{U_{i+1,j} - U_{i,j}}{\Delta x}, 0 \right\} \right\} \right)^2 + \left( \max \left\{ \max \left\{ \frac{U_{i,j} - U_{i,j-1}}{\Delta x}, 0 \right\}, -\min \left\{ \frac{U_{i,j+1} - U_{i,j}}{\Delta x}, 0 \right\} \right\} \right)^2 = \frac{1}{f_{i,j}^2}.$$

I metodi Fast Marching sono molto efficienti anche in 3D, hanno molteplici applicazioni come la segmentazione, robotica, propagazione di onde sismiche, ricerca del percorso più breve. In sostanza è un modo intelligente di risolvere l'equazione *Eikonale*.



## Capitolo 4

### Segmentatore automatico

**L**'obiettivo generale di questo lavoro è quello di realizzare un software con interfaccia grafica per la segmentazione di immagini biomediche (CT, MR, PET, RX etc...) bidimensionali. In particolare il software deve possedere un'interfaccia grafica *user friendly* in grado di aprire immagini di qualsiasi tipo (comprese DICOM), scegliere tra i diversi algoritmi implementati, settare i parametri rilevanti per ogni algoritmo, visualizzare e salvare la segmentazione finale. Oltre ad algoritmi automatici, il software, permette anche di segmentare manualmente l'immagine. In più è richiesta l'implementazione di uno strumento per il confronto diretto tra diversi metodi sulla stessa immagine per avere un riscontro sia visivo che parametrico, attraverso dei coefficienti solitamente usati in letteratura (i.e. differenza di area, coefficiente di Dice, distanza quadratica media, accuratezza, etc...). Bisogna sottolineare che, essendo la segmentazione una tra le elaborazioni più complesse e richiedendo spesso l'interazione con un operatore, il software di segmentazione non può essere totalmente automatico in quanto prevede una scelta degli algoritmi e dei parametri adeguata all'immagine da segmentare. Gli algoritmi da implementare, già illustrati nel capitolo precedente, sono stati suddivisi concettualmente in tre gruppi in

base al loro grado di complessità. Tecniche di base: Otsu, logic operations (Gradient, Dilate, Fill,...), Watershed, Region Growing. Tecniche avanzate: active contours without edges Chan-Vese, Geodesic active contours, Fast Marching Level-Set Method, Snakes. Tecniche manuali: freehand, polygonal. Il software è stato realizzato con *Matlab* in versione *R2016a* e l'ausilio del tool *GUIDE* (*Graphical User Interface Development Environment*) per la creazione dell'interfaccia grafica. Di seguito verrà esposta inizialmente un guida al primo utilizzo e alla comprensione dell'interfaccia, successivamente si passerà alla descrizione in dettaglio dell'implementazione e infine verranno illustrati i parametri di confronto presi in considerazione.

## **4.1 Guida all'utilizzo, interfaccia**

Al primo avvio l'interfaccia è molto semplice, Figura 4.1, l'idea è quella di costruirla dinamicamente in base agli algoritmi selezionati in maniera da non creare confusione con strumenti non utili. La prima schermata non permette di fare nulla se non di aprire un'immagine di partenza cliccando l'apposita icona in alto a sinistra, infatti premendo su qualsiasi altra cosa il programma restituisce un avviso invitando l'utente ad aprire un'immagine.

### **4.1.1 Tecniche di base**

Dopo aver aperto l'immagine si può selezionare una delle categorie di tecniche disponibili (*Basic, Advanced e Manual Segmentation*), immediatamente dopo aver selezionato una delle tre comparirà un menù a tendina nel quale sono elencate le tecniche disponibili per quella categoria, in Figura 4.2 sono mostrate le tecniche di base.

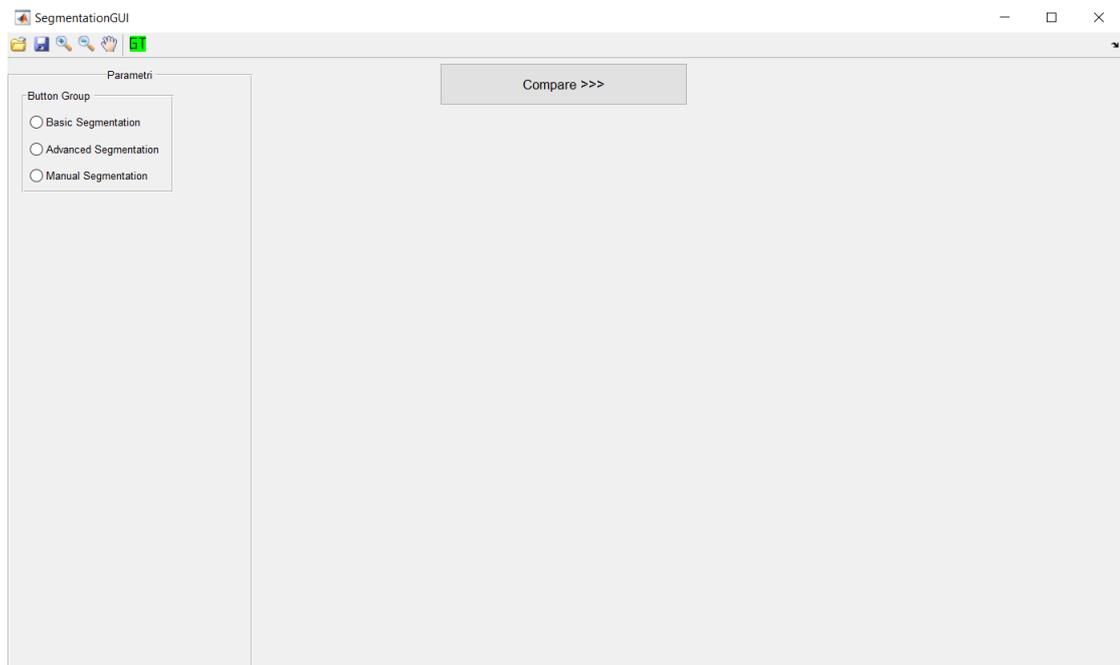


Figura 4.1: Schermata di apertura.

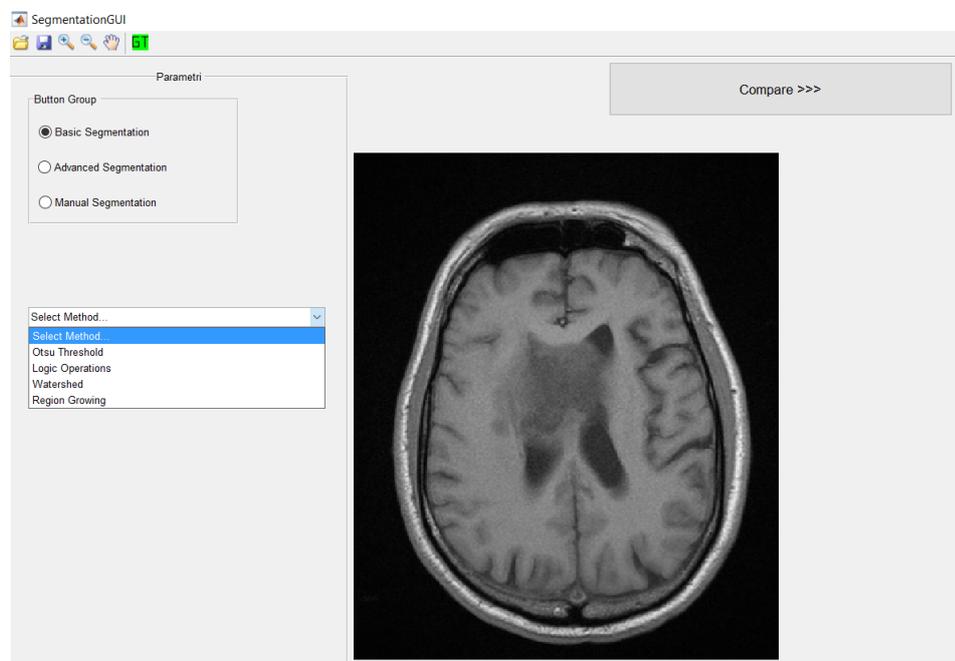


Figura 4.2: Tecniche di base.

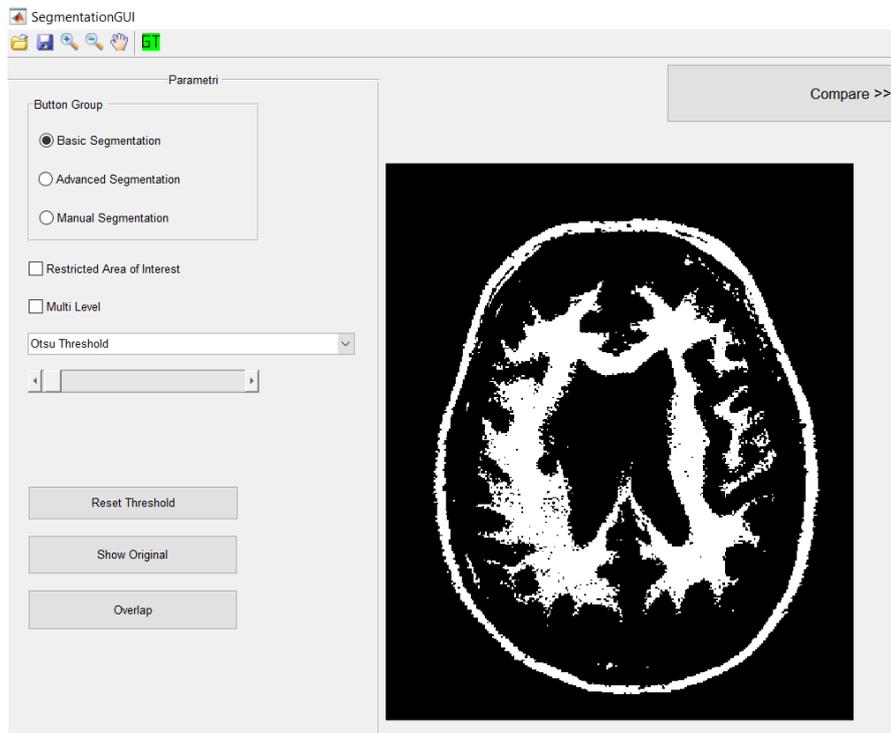


Figura 4.3: *Interfaccia Otsu.*

Dipendentemente dalla tecnica selezionata compariranno altri strumenti quali pulsanti, campi di testo e/o slider. Ad esempio in Figura 4.3 sono mostrati gli strumenti disponibili per il metodo Otsu, in particolare si hanno i pulsanti *Reset Threshold*, *Show Original* e *Overlap*, poi si ha uno *slider* e le caselle *Restricted Area of Interest* e *Multi Level* che saranno illustrati a breve. Bisogna ricordare che questa tecnica calcola automaticamente la soglia ottima per l'immagine considerata, infatti lo *slider* viene settato su tale soglia ed è possibile intervenire sullo stesso per cambiare la soglia. Il pulsante *Reset Threshold*, come si può intuire, resetta il valore della soglia a quello ottimale riportando lo slider al punto di partenza. Il pulsante *Show Original* mostra l'immagine originale non segmentata che è stata aperta inizialmente, mentre il pulsante *Overlap* mostra una sovrapposizione tra

l'immagine segmentata e l'originale in falsi colori, ne vedremo un esempio a breve. Questi ultimi due pulsanti sono ricorrenti in tutte le tecniche ed hanno sempre la stessa funzione, per cui d'ora in poi ne verrà sottintesa la funzionalità. Il flag *Restricted Area of Interest* permette di selezionare una *ROI, Region of Interest*, sulla quale effettuare la segmentazione. Dopo aver selezionato una ROI, in questo caso di tipo quadrilatera, bisogna fare un doppio click all'interno della ROI stessa per far partire la segmentazione. Il flag *Multi Level* estende il metodo Otsu ad un'elaborazione multi soglia



Figura 4.4: *Interfaccia Otsu multilivello/multisoglia.*

e multilivello, non prevista negli obiettivi iniziali di questa tesi, mostrata in Figura 4.4. Selezionando la funzionalità aggiuntiva multilivello compaiono degli strumenti aggiuntivi che permettono di modificare il numero di livelli (da 2 a 20) e di selezionare ogni singolo livello e visualizzarne l'immagine

binaria corrispondente. L'algoritmo elabora automaticamente la segmentazione con il numero di livelli ottimale (che non è detto sia 20) e ne mostra il risultato in falsi colori.



Figura 4.5: *Immagine di partenza della segmentazione di Fig.4.6.*

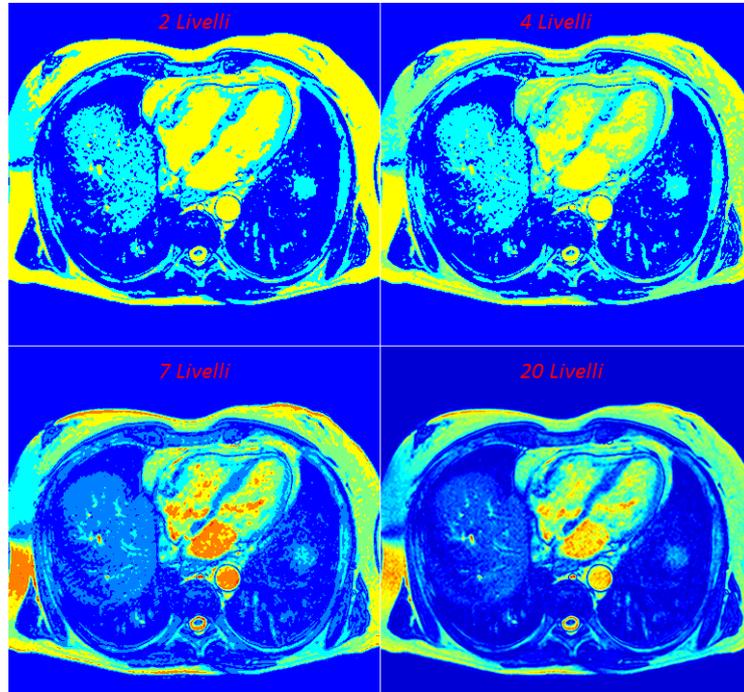


Figura 4.6: *Otsu con diversi livelli di segmentazione, rispettivamente 2, 4, 7 e 20.*

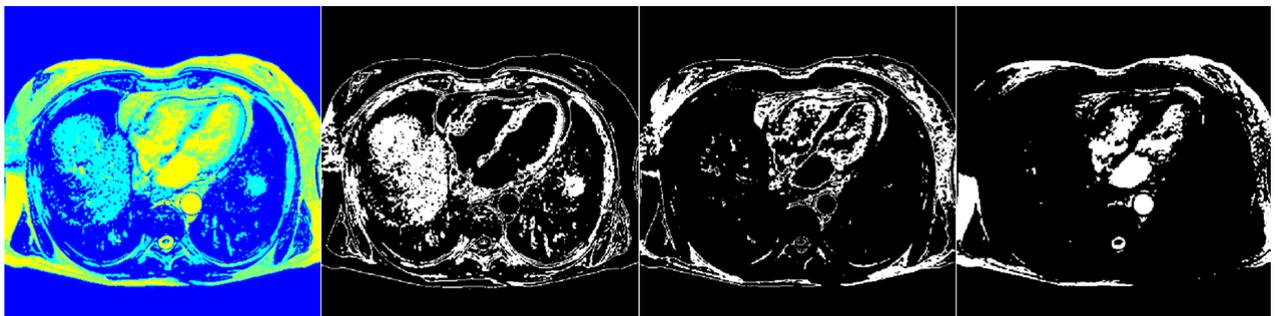


Figura 4.7: *Selezione dei diversi livelli con la tecnica di Otsu multisoglia.*

In Figura 4.6 viene illustrata una diversa scelta del numero di livelli applicata alla stessa immagine di partenza mostrata in Figura 4.5. Mentre in Figura 4.7 si può apprezzare la selezione dei singoli livelli, dove ognuno dei quali rappresenta un diverso colore della prima immagine.

Selezionando invece la tecnica *Logic Operations*, come mostrato in Fi-

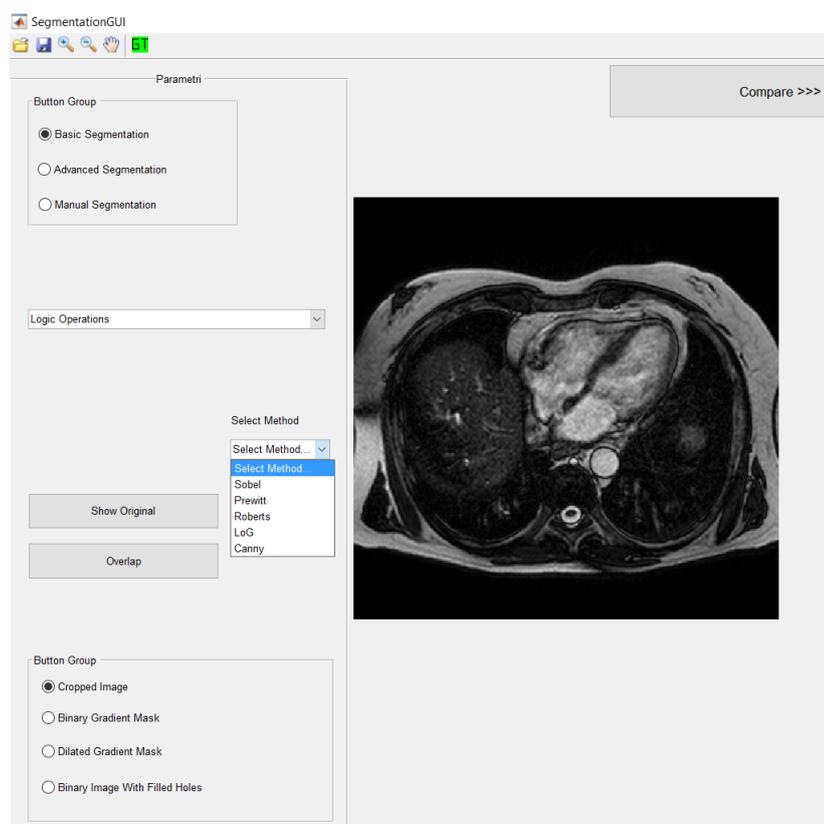


Figura 4.8: *Interfaccia Logic Operations.*

gura 4.8, compariranno: un menù a tendina dal quale si può selezionare il tipo di operatore, tra cui Sobel, Prewitt, Roberts; e una serie di pulsanti per scegliere cosa visualizzare tra la ROI (*Cropped*), il gradiente binario, il gradiente dilatato e l'immagine di segmentazione finale. Questa tecnica richiede di selezionare inizialmente una ROI con le stesse modalità del caso

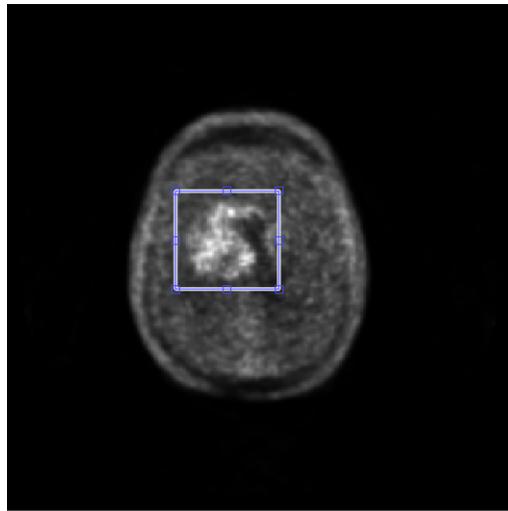


Figura 4.9: *Immagine con ROI segmentata in Figura 4.10.*

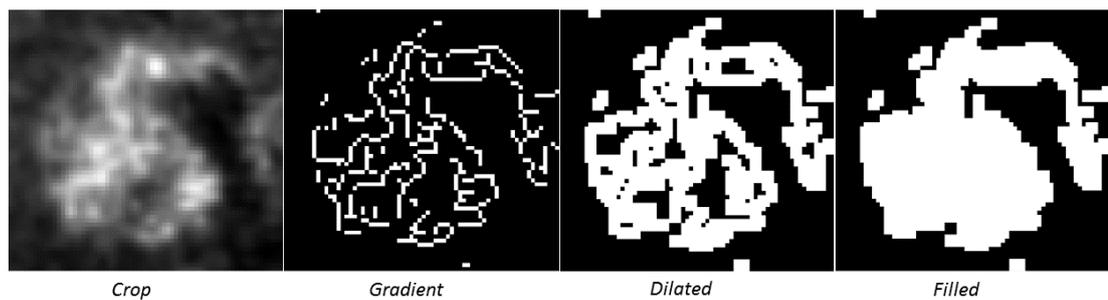


Figura 4.10: *Segmentazione con Logic Operations - Sobel.*

precedente. Ad esempio l'elaborazione dell'immagine di Figura 4.9 attraverso questa tecnica si può apprezzare in Figura 4.10.

Nella tecnica *Watershed* bisogna soltanto selezionare una ROI iniziale. Un esempio di segmentazione attraverso *Watershed* è mostrato in Figura 4.11. Selezionando invece la tecnica *Region Growing* (Figura 4.12) compare un campo di testo nel quale inserire il valore<sup>4.1</sup> di *Maximum Intensity Distance*, inizialmente settato automaticamente 0.2. Per far partire l'algoritmo bisogna selezionare un punto di partenza e verrà mostrata la segmentazione.

<sup>4.1</sup>In certi casi per avere dei risultati apprezzabili è necessario inserire valori di 100 o 200.

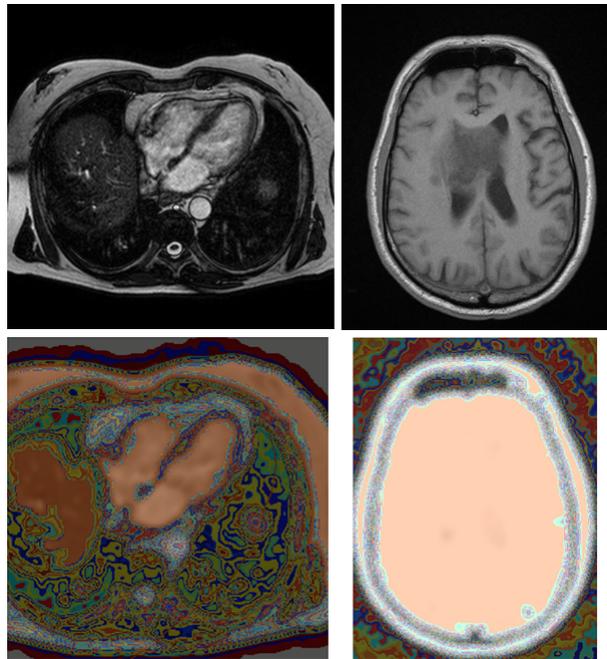


Figura 4.11: *Segmentazione Watershed.*

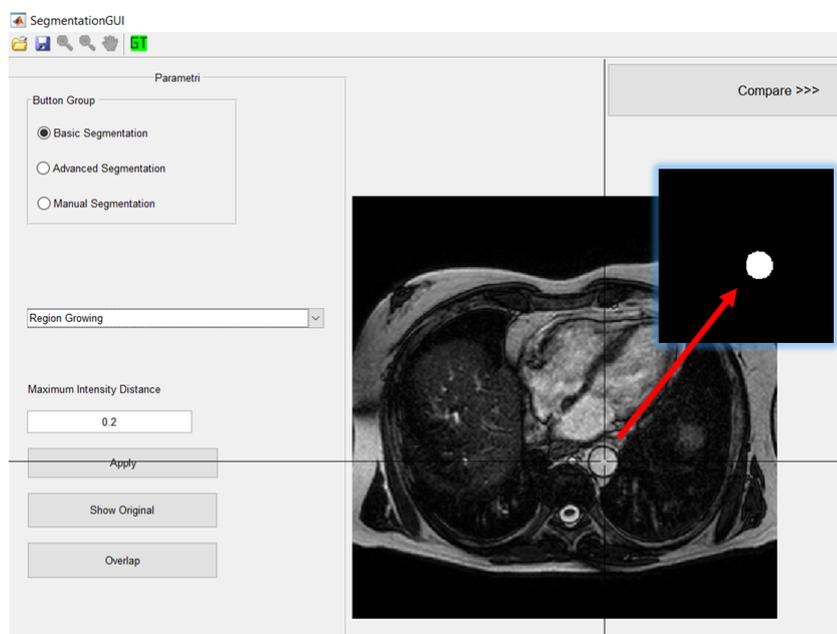


Figura 4.12: *Interfaccia Region Growing.*

### 4.1.2 Tecniche avanzate

Per le tutte tecniche avanzate l'interfaccia con cui interagire è quasi la stessa, selezionando una di suddette tecniche apparirà una situazione come in Figura 4.13, dove si avrà un campo di testo in cui inserire il numero massimo di iterazioni o un valore di soglia, dipendentemente dal tipo di algorit-

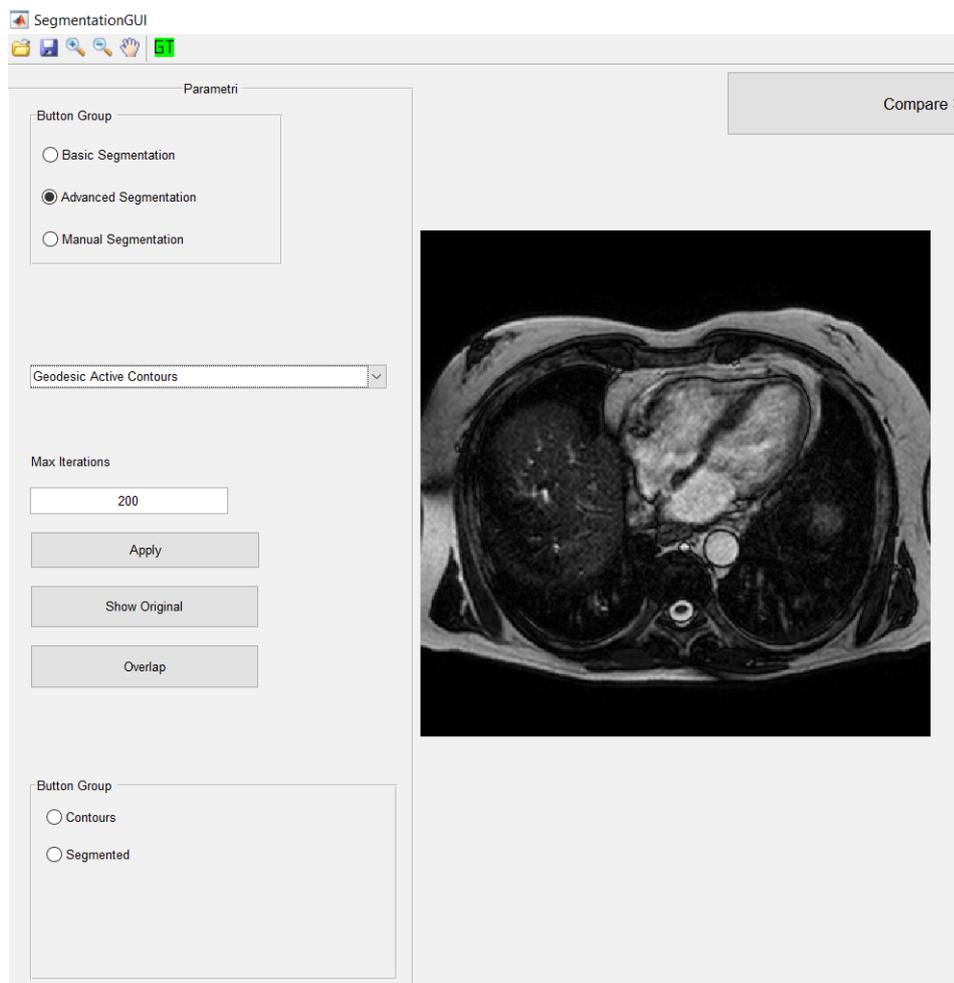


Figura 4.13: *Interfaccia tecniche avanzate.*

mo selezionato, oltre ai soliti due tasti sia avrà anche il tasto *Apply* il quale permette di applicare le eventuali modifiche apportate al valore nel campo

di testo, andando a ricalcolare e mostrare la segmentazione. In alternativa al tasto *Apply* si può direttamente premere invio dopo aver digitato il valore desiderato. In basso invece compariranno dei cosiddetti *button radio*, similmente al caso di *Logic Operations* visto sopra, che danno la possibilità visualizzare il contorno individuato piuttosto che l'immagine segmentata.

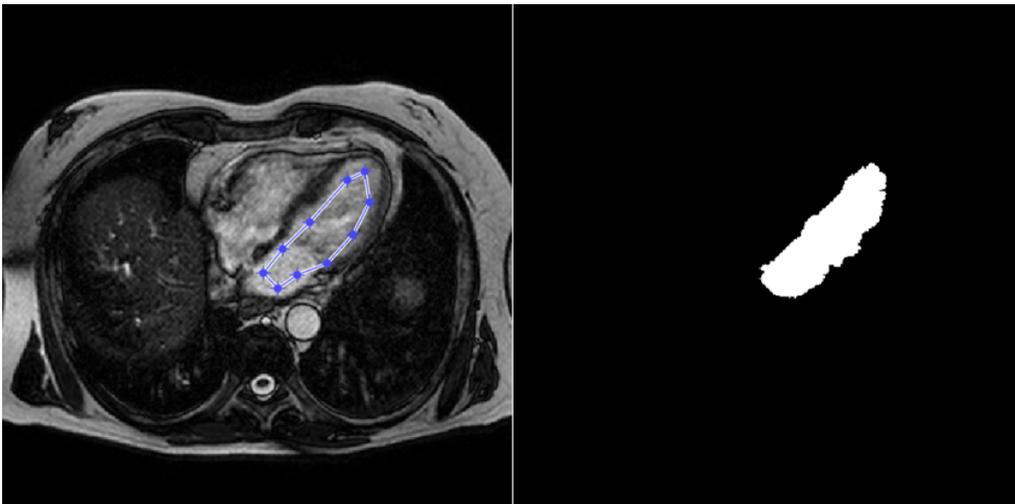


Figura 4.14: *Segmentazione con Chan-Vese.*

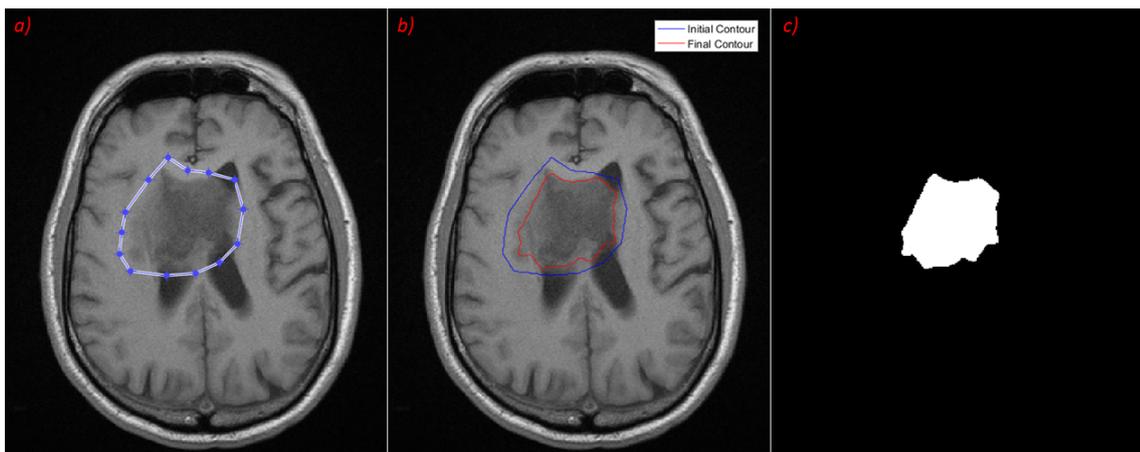


Figura 4.15: *Segmentazione con Geodesic.*

Selezionando l'algoritmo *Chan-Vese* bisognerà selezionare una ROI di tipo poligonale internamente alla regione interessata, come mostrato in Figura 4.14, e successivamente fare un doppio click con il pulsante sinistro del mouse all'interno della ROI per eseguirne la segmentazione. In questo caso viene visualizzata direttamente l'immagine segmentata.

Selezionando invece *Geodesic*, situazione in Figura 4.13, si ha la possibilità di modificare il numero di iterazioni da effettuare ma a differenza di *Chan-Vese* si può scegliere se visualizzare il contorno iniziale-finale o la segmentazione. In Figura 4.15 si può vedere un esempio di segmentazione con l'algoritmo *Geodesic* dove si ha rispettivamente *a)* selezione della ROI, *b)* contorno iniziale in blu e finale in rosso e *c)* l'immagine segmentata. Come si può intuire aumentando il numero di iterazioni diminuirà la regione segmentata e quindi il contorno in rosso si farà sempre più piccolo. In *Geodesic* va invece indicata una ROI poligonale leggermente esterna alla regione interessata.

Nella tecnica *Fast Marching* invece non si ha un numero di iterazioni bensì un valore di soglia che è inizializzato al valore 0.01, è stato verificato sperimentalmente che per la maggior parte dei casi si ottengono risultati soddisfacenti con un range di valori compreso tra 0.001 e 0.03. In questo caso è possibile visualizzare il contorno iniziale-finale, il solo contorno finale enfatizzato oppure l'immagine segmentata. In Figura 4.16 è possibile osservare un esempio di segmentazione con il metodo *Fast Marching* dove si ha rispettivamente *a)* la ROI selezionata, *b)* i contorni iniziale-finale, *c)* il solo contorno finale enfatizzato e *d)* l'immagine segmentata.

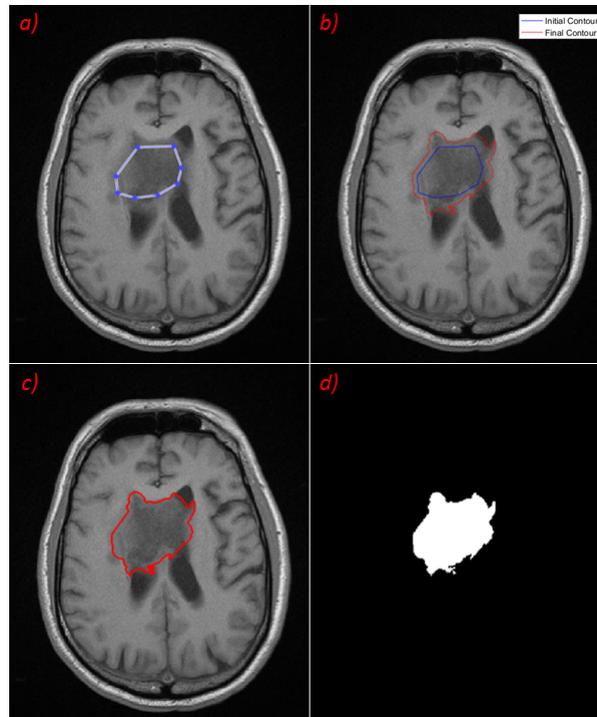


Figura 4.16: *Segmentazione con Fast Marching.*

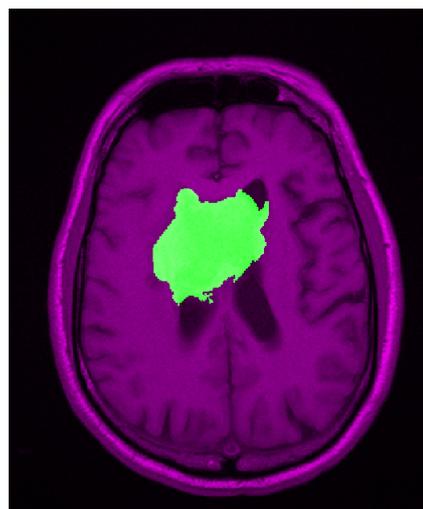


Figura 4.17: *Sovrapposizione della segmentazione di Fig. 4.16 con l'immagine originale.*

Inoltre in Figura 4.17 si mostra a titolo illustrativo la sovrapposizione dell'immagine segmentata con l'originale attraverso la pressione del tasto dedicato *Overlap*. Nel *Fast Marching* va indicata una ROI poligonale internamente alla regione interessata

Per la tecnica cosiddetta *Snakes* si ha un'interazione leggermente maggiore, infatti oltre a selezionare il numero di iterazioni e a poter visualizzare il contorno finale o l'immagine segmentata, bisogna inserire dei punti che circondino l'area di interesse (quindi esternamente). Questi punti saranno

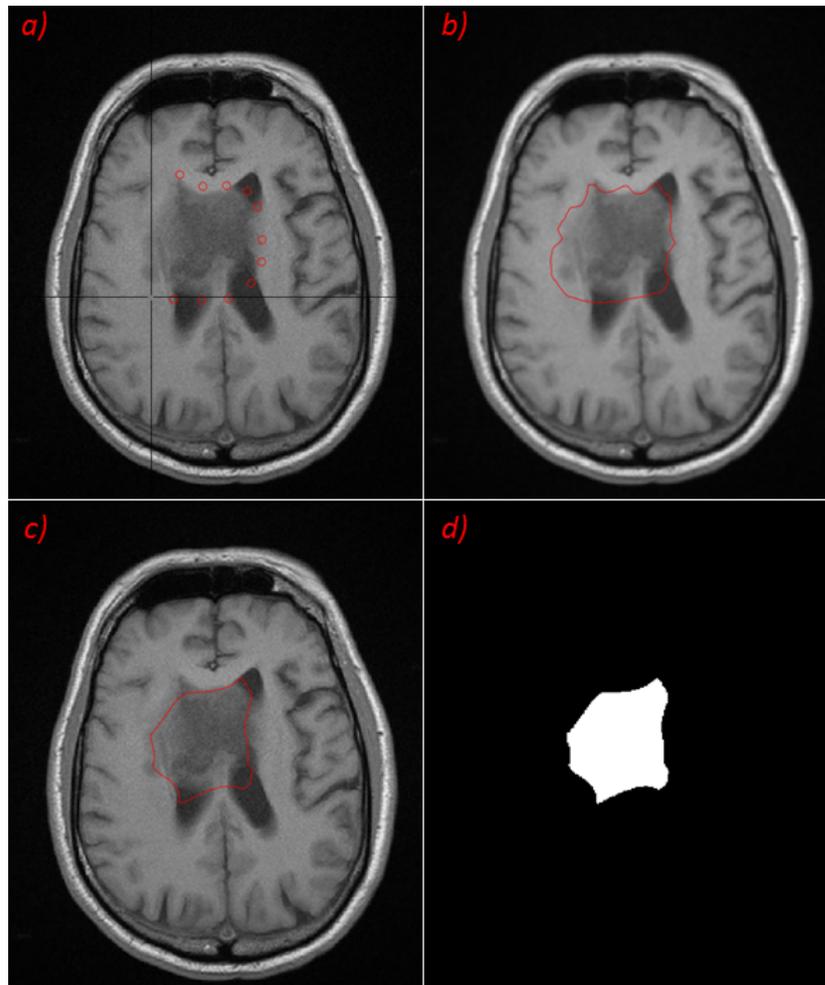


Figura 4.18: *Segmentazione con Snakes*.

la base sui cui verrà interpolato uno snake che iterazione dopo iterazione andrà a stringersi fino ad aderire all'area interessata. I punti vanno inseriti mediante il tasto sinistro del mouse, ogni click inserisce un punto, per inserire l'ultimo punto e dare inizio all'elaborazione si usa il tasto destro del mouse. Inoltre si può eliminare l'ultimo punto inserito (nel caso in cui si sia commesso un errore) premendo il terzo tasto del mouse (click sulla rotella) oppure la barra spaziatrice sulla tastiera. Una volta cominciata l'elaborazione si vedrà un'animazione in tempo reale dello snake mentre lavora. Un esempio di segmentazione attraverso il metodo *Snakes* è mostrato in Figura 4.18, dove rispettivamente viene mostrato *a)* l'inserimento dei punti, *b)* l'elaborazione intermedia dello *Snake*, *c)* il contorno finale individuato e *d)* l'immagine segmentata.

### 4.1.3 Tecniche manuali

Per quanto riguarda le tecniche manuali non c'è molto da aggiungere. Si hanno due tipi di segmentazioni, attraverso la costruzione di una poligonale e la segmentazione a mano libera. L'unica novità riguardante l'interfaccia è la comparsa di un nuovo tasto *Clear*, il quale consente di eliminare l'attuale segmentazione in corso e ricominciare da zero.

### 4.1.4 Strumento di confronto

Premendo il grande tasto *Compare* posto in alto (osservabile in una qualsiasi delle figure raffiguranti l'interfaccia) si dà vita ad una sezione speculare che apparirà sul lato destro, questa nuova sezione è esattamente identica alla principale in cui viene rappresentata la stessa immagine aperta in precedenza. Questa situazione si può osservare in Figura 4.19 dove vengono

confrontati i metodi *Geodesic* e *Fast Marching*. Come si può notare premendo il grande tasto in basso *Get Comparision Parameters* comparirà una tabella contenente dei parametri di confronti tra le due segmentazioni. Nel momento in cui si preme suddetto tasto verrà chiesto quale delle due immagini prendere come riferimento, questo è necessario per il calcolo di alcuni parametri come *Accuracy* e *TRN/TPR*. Inoltre è stata aggiunta la possibilità di aprire un'immagine diversa in una delle due sezioni, questo per dare la possibilità di caricare un'eventuale *Ground Truth* nota a priori (magari segmentata manualmente da medici esperti) e confrontarla con l'algoritmo utilizzato, avendo cura di selezionarla come riferimento quando si andranno a richiedere i parametri di confronto. A questa funzionalità si può accedere premendo il tasto nella toolbar in alto raffigurante la dicitura *GT* su sfondo verde.

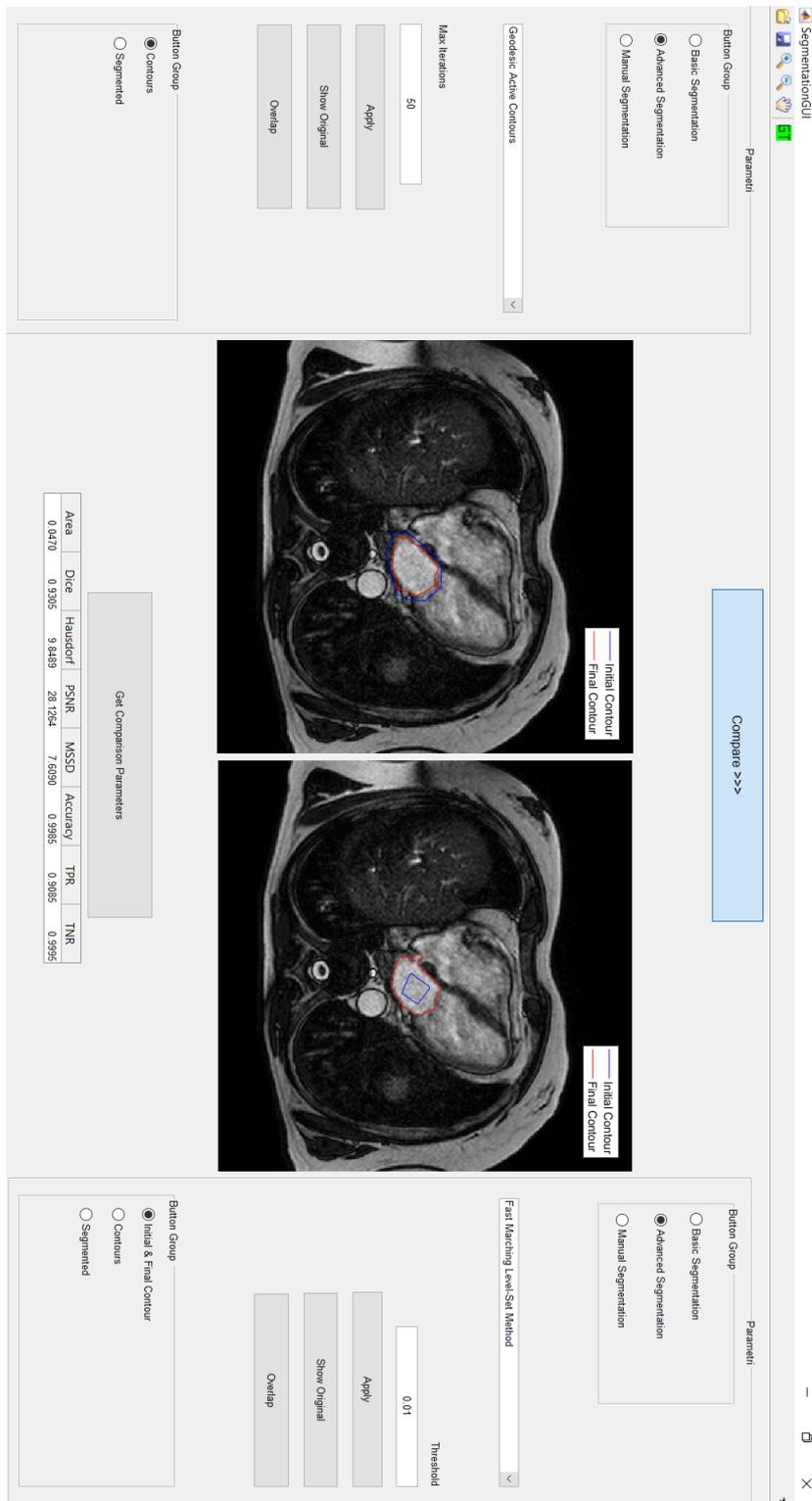


Figura 4.19: *Interfaccia per il confronto.*

## 4.2 Il software

### 4.2.1 Generalità

L'approccio alla progettazione ed implementazione è stato guidato dalla modalità di costruzione delle interfacce grafiche in Matlab, ovvero in una forma simile alla programmazione ad oggetti. Ogni algoritmo è stato implementato in un file separato, quindi una funzione separata. L'interfaccia grafica è invece contenuta in un unico grande file composto da tante funzioni, una per ogni componente dell'interfaccia (pulsanti, toolbar, menu a tendina etc...). Nel file contenente l'interfaccia è stata implementata un'ulteriore funzione dedicata alla gestione dell'immagine visualizzata, che viene invocata alla pressione di alcuni tasti chiave<sup>4.2</sup> e richiama le funzioni degli algoritmi richiesti dall'utente. Inoltre il software può essere esportato in un file eseguibile (.exe) che può funzionare su qualsiasi pc *Windows*.

### 4.2.2 Implementazione algoritmi

Come accennato gli algoritmi sono stati implementati tutti in delle *function* separate che vengono chiamate quando necessario da una funzione contenuta nel file principale dell'interfaccia.

#### 4.2.2.1 Tecniche di base

Nel Codice 4.1 si può vedere la funzione[12] che implementa l'algoritmo di Otsu. Nella quale alla terza riga viene prima calcolata la soglia ideale con `graythresh()` e successivamente viene convertita l'immagine di partenza in una binaria in base alla soglia calcolata. La funzione quindi prende in ingresso soltanto l'immagine da segmentare e restituisce la segmentata e il

---

<sup>4.2</sup>e.g. il tasto *Apply* o la selezione di un algoritmo

```
1 % Otsu method, single threshold, single level
2 function [BW, level1] = Botsu(I)
3     level1 = graythresh(I); %crea l'immagine in scala di grigi
4     BW = im2bw(I,level1);
5 end
```

Codice 4.1: Funzione per *Otsu*.

livello calcolato.

Per quanto riguarda invece Otsu multilivello si ha la funzione[13] mostrata nel Codice 4.2, che prende in ingresso l'immagine originale ed eventualmente il numero di livelli desiderato e restituisce l'immagine in falsi colori (*RGB*) il numero di livelli *m* e la soglia. In questa function dopo aver inizializzato i parametri `bestmetric` e `m` a zero viene prima calcolato il numero di livelli ottimale se non specificato tra i parametri in ingresso. Per quanto riguarda invece il metodo[14] *Logic Operations* è stato implementato in quattro step, come si vede nel Codice 4.3, dove è stato omesso il primo step che prevedeva il crop dell'immagine in quanto alla funzione viene fornita già una ROI. La *function* prende in ingresso la ROI dell'immagine originale, `cropped`, e una stringa `metName` indicante il tipo di metodo da utilizzare (ad esempio `sobel` `prewitt` etc...) e restituisce le tre immagini delle linee individuate, dilatate e riempite. Nel secondo step vengono individuati i contorni dell'immagine attraverso il metodo indicato in `metName`, successivamente nello step 3 vengono dilatate le linee individuate e infine nel quarto passo vengono riempiti i gap.

```

1 % Otsu method, multilevel
2 function [RGB, m, bestthresh] = BotsuML(I,Nlevels)
3 bestmetric = 0;
4 m=0;
5 if nargin <2
6     for n=2:20
7         [thresh, metric] = multithresh(I,n);
8         if metric > bestmetric
9             bestmetric = metric;
10            bestthresh = thresh;
11            m=n;
12        end
13    end
14 else
15     bestthresh = multithresh(I,Nlevels);
16 end
17 seg_I2 = imquantize(I,bestthresh);
18 RGB = label2rgb(seg_I2);
19 end

```

Codice 4.2: Funzione per *Otsu Multilivello*.

```

1 function [binGrad, dilatedGrad, binFill] = BLogicOp(cropped
    ,metName)
2 %Step 2: Detect
3 [~, threshold] = edge(cropped, metName);
4 fudgeFactor = .5;
5 binGrad = edge(cropped,metName, threshold * fudgeFactor);
6 %Step 3: Dilate the Image
7 se90 = strel('line', 3, 90);
8 se0 = strel('line', 3, 0);
9 dilatedGrad = imdilate(binGrad, [se90 se0]);
10 %Step 4: Fill Interior Gaps
11 binFill = imfill(dilatedGrad, 'holes');

```

Codice 4.3: Funzione per *Logic Operations*.

L'algoritmo Watershed è stato implementato seguendo le linee guida di [15] dove viene prima applicata una maschera di *Sobel* e similmente a quanto fatto per *Logic Operations* viene calcolato la magnitudine del gradiente per rilevare gli edge. Prima di applicare la funzione predisposta *Watershed* di Matlab bisogna marcare gli oggetti sullo sfondo avendo poi cura di eliminare le macchie lasciate dalla procedura di marking. Poi viene binarizzata l'immagine ed elaborata con la function di Matlab `watershed()`. Infine si visualizza il risultato. L'algoritmo *Region Growing* è stato implementato a partire da [16], il quale ha inizio da un singolo seme inserito dall'utente. La regione cresce iterativamente comparando i pixel contigui non allocati alla regione. In questa implementazione la differenza tra il valore dell'intensità dei pixel e la media della regione, è usata come misura di similarità. Il pixel con la minor differenza misurata viene allocato alla regione. Il processo si interrompe quando la suddetta differenza diventa maggiore della soglia specificata.

#### 4.2.2.2 Tecniche avanzate

Gli algoritmi Chan-Vese e Geodesic, facendo parte degli Active Contour, vengono implementati in maniera molto simile. Inizialmente si specifica il contorno iniziale (che deve essere interna all'area interessata i per Chan-Vese e esterna per Geodesic) attraverso una poligonale grazie al comando: `mask = roipoly;` dove in `mask` sarà contenuta la maschera poligonale. Successivamente viene elaborata la segmentazione tramite la comoda funzione *Matlab* `BW = activecontour(I, mask, maxIterations, 'Chan-Vese');`, la quale segmenta l'immagine usando l'active contour, in questo caso è specificato l'utilizzo del Chan-Vese, altrimenti per Geodesic bisogna inserire come ultimo argomento della *function* la dicitura *ed-*

ge. La funzione `BW = activecontour(I, mask, n, method)`; segmenta l'immagine bidimensionale in scala di grigi distinguendo tra sfondo e oggetti in primo piano, in cui lo sfondo è in nero. *Mask* come già detto è l'immagine binaria che specifica il contorno iniziale. Per dei risultati ottimali è necessario individuare un contorno iniziale il più vicino possibile alla regione di interesse. In `BW` è contenuta la segmentazione finale elaborata per un numero massimo di `n` iterazioni. In `method` invece è specificato il metodo di active contour da utilizzare, i.e. 'Chan-Vese' o 'edge'. Inoltre si possono specificare altri parametri nella *function* come ad esempio `BW = activecontour(I, mask, maxIterations, 'edge', 'SmoothFactor', 1.5, 'ContractionBias', 0.6)`; , dove il numero che segue la stringa 'SmoothFactor' indica la 'morbidezza' o regolarità dei contorni della regione segmentata, di default viene impostato 0 per *Chan-Vese* e 1 per *Geodesic*. Il valore associato a 'ContractionBias' è invece la tendenza alla crescita del contorno verso l'esterno o verso l'interno, dipendentemente dal segno. Un valore positivo polarizza i contorni a restringersi, negativo ad espandersi. L'algoritmo *Fast Marching* è stato implementato attraverso [18] la *function* mostrata nel Codice 4.4.

```

1 % FAST MARCHING LEVEL SET
2 function [BW, D] = AFMLM(I, mask, thresh)
3 W = graydiffweight(I, mask, 'GrayDifferenceCutoff', 25);
4 [BW, D] = imseggmm(W, mask, thresh);

```

Codice 4.4: Funzione per *Fast Marching*.

La quale prende in ingresso l'immagine di partenza *I*, il contorno iniziale *mask* e la soglia *thresh* inizialmente settata a 0.01. Il contorno iniziale viene creato interattivamente grazie a `roipoly` prima di chiamare tale funzione, il

quale deve essere individuato internamente alla regione interessata. Per prima cosa, alla terza riga, viene elaborato un array contenente dei pesi basati sulla variazione dell'intensità in scala di grigi. Poi viene segmentata l'immagine utilizzando questi pesi attraverso la *function* di Matlab `imsegfmm`, nella quale vengono specificati appunto i pesi  $W$ , la maschera iniziale e la soglia. La funzione restituisce l'immagine segmentata  $BW$  e  $D$  che rappresenta la matrice della distanza geodetica normalizzata, la quale può essere sogliata utilizzando soglie differenti per avere diversi risultati sulla segmentazione. L'implementazione dell'algoritmo *Snakes* è stata divisa in due step, il primo in cui viene costruito lo snake attraverso l'interazione con il mouse e il secondo che compie l'elaborazione vera e propria. Il primo step viene compiuto da una funzione appositamente costruita, `[smth, xs, ys] = Snakes(I)` che prende in ingresso soltanto l'immagine originale e restituisce le coordinate dei punti della curva che interpola lo snake sulla base dei punti inseriti e l'immagine originale a cui è stato applicato uno smoothing. All'interno della funzione vengono gestiti gli input forniti attraverso il mouse e vengono segnati i punti a video, successivamente viene interpolata una curva attraverso lo *spline*. Poi grazie alla funzione customizzata `[xs, ys] = interate(I, xs, ys, iterations)` viene elaborata la progressione dello snake, dove bisogna specificare, oltre alle coordinate della curva di partenza, i parametri: *alpha*, *beta*, *gamma*, *kappa*,  $w_l$ ,  $w_e$ ,  $w_t$  e *iterations*. Dove l'ultimo è il numero massimo di iterazioni, *alpha* e *beta* regolano il controllo sulla tensione e rigidità dello snake, *gamma* indica la dimensione del passo,  $w_l$ ,  $w_e$  e  $w_t$  sono i pesi rispettivamente dell'energia della linea, edge e terminale. La funzione restituisce il contorno finale individuato.

### 4.2.2.3 Tecniche manuali

La segmentazione manuale viene distinta in due modalità, Freehand e Polygonal. Dove in entrambe viene prima tracciata un contorno interattivamente e poi viene calcolata la corrispondente immagine binaria e quindi visualizzata la segmentazione. Per effettuare una segmentazione cosiddetta a *mano libera*, viene usato il comando `M=imfreehand(gca)` in ingresso viene passato `gca` che rappresenta il *Current axes handle* ovvero in questo caso l'immagine che si sta visualizzando. Allo stesso modo per un contorno poligonale si usa `M=impoly(gca)`; . Entrambe restituiscono il contorno individuato, il quale viene utilizzato per ricavare una maschera, o immagine binaria, attraverso la funzione `createMask(M, I)`; , dove bisogna specificare il contorno  $M$  e l'immagine originale  $I$ , e restituisce l'immagine segmentata.

### 4.2.3 Implementazione dell'interfaccia

L'interfaccia è stata realizzata con l'ausilio del tool *GUIDE* disponibile in Matlab, il quale permette di creare una struttura per l'interfaccia grafica e di associare ad ogni elemento una determinata funzione nello script generale del file `.m` creato. L'impostazione del codice è tutta basata sulle function, ogni elemento dell'interfaccia possiede una function, la quale non restituisce nessun output e prende in ingresso due variabili cosiddette di *handle*. Attraverso queste ultime è possibile creare delle interazioni tra le function ed eventualmente condividere alcune variabili. Solitamente<sup>4.3</sup> le variabili di input di ogni callback sono tre: *hObject*, *eventdata* e *handles*. Il primo rappresenta l'elemento dell'interfaccia grafica che ha innescato la callback,

---

<sup>4.3</sup>Ci sarebbe anche un altro parametro, *varargin*, che permette di passare degli argomenti dalla *command line* alla *function*, ma nel nostro caso non è stato necessario usarlo.

quindi ad esempio con il comando `get(hObject, 'Value')`; otteniamo il valore di tale oggetto, ad esempio 1 o 0 nel caso di un radio button, oppure la posizione dell'elemento selezionato in un menù a tendina. Il parametro *eventdata* contiene informazioni specifiche sugli eventi che occorrono sulla tastiera o click del mouse. Infine *handles* è una struttura contenente tutti gli oggetti della UI, bisogna usare la funzione `guidata` per aggiornare eventuali valori modificati durante una callback e per mantenere la sua struttura. Quindi la struttura *handles* viene usata per creare interazioni tra le diverse function. Oltre a tutte le funzioni rappresentanti ogni elemento della GUI, è stata creata una function totalmente custom, la quale non è associata ad un determinato elemento dell'interfaccia ma interagisce con tutti essendo richiamata quasi ad ogni interazione con la UI. Questa funzione, `function DisplayImage1(hObject, handles)`, gestisce tutti gli algoritmi e decide i metodi e i parametri con cui verrà fatta la segmentazione e si occupa di visualizzare la corretta immagine a video, tutto in base all'interazione dell'utente. Questa è molto articolata e lunga (circa 350 righe di codice) ma non per questo complicata, è ben strutturata e commentata, è stata implementata cercando di renderla più comprensibile possibile. Sono presenti due funzioni indipendenti di questo tipo una per la sezione di sinistra e una per quella di destra, necessarie per il confronto. Di seguito ne viene descritto sommariamente il funzionamento. Inizialmente viene controllata se l'immagine è realmente in scala di grigi ed eventualmente convertita, questo per non creare problemi ed errori con gli algoritmi. Attraverso uno *switch case* viene eseguito un blocco di istruzioni in base all'algoritmo selezionato. In tutto il percorso viene tenuta traccia dell'eventuale ROI o contorno selezionato, in modo da tenerli in memoria per ricalcolare la segmentazione con

parametri diversi ma mantenendo lo stesso contorno iniziale, oppure confrontando diversi metodi bisogna elaborare la stessa ROI. Il tasto *compare* è collegato a tutta l'interfaccia rappresentata sul lato destro della finestra, infatti alla sua pressione viene visualizzata o nascosta. Mentre il tasto del calcolo dei parametri di confronto permette di calcolarli e visualizzarli in una tabella. Quando invece l'utente sceglie di aprire un'immagine di riferimento (o *ground truth*, GT) succede che viene bypassata la funzione generale che gestisce gli algoritmi e viene visualizzata la GT nella posizione scelta.

### 4.3 Parametri di confronto

I parametri di confronto sono stati implementati in funzioni separate, come fatto per gli algoritmi, di seguito vengono introdotti. Come già detto alla richiesta del calcolo dei parametri viene chiesto all'utente, tramite una finestra popup, di scegliere quale tra le due segmentazioni visualizzate è quella di riferimento. Alcuni parametri possono subire variazioni in base alla scelta fatta perchè per loro natura vengono calcolati sulla base di un'immagine di riferimento, come ad esempio avviene nel calcolo dei *veri positivi*, *accuratezza* etc...

I parametri di confronto implementati sono:

- differenza di area,
- coefficiente di *Dice*,
- distanza di *Hausdorff*,
- *Peak Signal-to-Noise Ratio* (PSNR),
- *Mean Sum of Square Distance* (MSSD),

- *Accuracy* (accuratezza),
- *True Positive Ratio* TPR,
- *True Negative Ratio* TNR,

La differenza di area è semplicemente il valore assoluto della differenza delle due aree normalizzata rispetto all'area massima, ovvero indicando con  $D_a$  la differenza di area risulta:

$$D_a = \frac{|Area_1 - Area_2|}{\max(Area_1, Area_2)},$$

è quindi un valore percentuale che vale zero nel caso in cui non ci sia alcuna differenza.

In generale la formula di Dice si applica alla presenza o assenza di dati, nel nostro caso al valore 0 o 1 nell'immagine segmentata, il coefficiente è calcolato come

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|},$$

dove  $|X|$  e  $|Y|$  sono il numero di 1 nelle due immagini confrontate.

La distanza di *Hausdorff* è una particolare definizione di distanza tra due sotto insiemi. Indicando la distanza di *Hausdorff* tra  $A$  e  $B$  come  $d(A, B)$  viene calcolata come

$$d(A, B) = \max \{e(A, B), e(B, A)\},$$

dove  $e(A, B)$  è l'eccedenza di  $A$  su  $B$ . In altre parole è la maggior distanza che intercorre tra un punto qualsiasi di  $A$  e più vicino punto di  $B$ . Cioè per ogni punto di  $A$  viene calcolata la distanza con il punto più vicino di  $B$  e ne viene trovato il massimo.

Il *PSNR* è utilizzato per valutare la qualità di un'immagine (compressa) rispetto all'originale, quindi nel nostro caso il risultato sarà relativo alla scelta dell'immagine di riferimento, ed è stato calcolato come

$$PSNR = -10 \log_{10} \left( \frac{I_1 + I_2}{N \cdot M} \right),$$

dove  $N \cdot M$  è la dimensione dell'immagine di riferimento,  $I_1$  è il numero di pixel in cui si verifica che nella prima immagine segmentata si ha un valore 1 e nella seconda 0,  $I_2$  viceversa quando si ha 0 nella prima immagine e 1 nella seconda. Quindi indicano quando le due immagini hanno dei valori discordanti in corrispondenza dello stesso pixel.

Il MSSD viene calcolato in questo modo: viene prima effettuata una trasformata distanza su entrambe le immagini, poi vengono estrapolati i contorni attraverso *imerode*, e successivamente calcolate due distanze come somme di quadrati e ne viene preso in considerazione il massimo.

I parametri *Accuracy*, *TPR* e *TNR* sono stati raggruppati in un'unica funzione, **function** [Accuracy, tpr, tnr] = valuationSeg(ref, img), dove vengono calcolati nel modo seguente

$$TPR = \frac{V^+}{V^+ + F^-},$$

o anche detta *sensibilità*, corrisponde ai veri positivi diviso la somma tra i veri positivi e i falsi negativi, cioè diviso il totale dei malati<sup>4.4</sup>, mentre per la *specificità*

$$TNR = \frac{V^-}{F^+ + V^-},$$

ovvero i veri negativi diviso la somma dei falsi positivi e dei veri negativi,

<sup>4.4</sup>Nel nostro caso per positivi e negativi intendiamo 1 e 0.

cioè diviso la somma dei sani. Per quanto riguarda l'accuratezza abbiamo

$$Accuracy = \frac{V^+ + V^-}{V^+ + F^+ + F^- + V^-}.$$

Questi parametri vengono poi inseriti in una tabella e visualizzati attraverso l'interfaccia.

## Capitolo 5

### *Risultati*

**P**er verificare la corretta implementazione degli algoritmi e lo strumento di comparazione, sono stati effettuati svariati test su diversi tipi di immagini. A titolo dimostrativo verranno illustrati tre confronti, di cui due rispetto a delle ground truth e uno tra diversi metodi. Prendendo in considerazione l'immagine in Figura 5.1 di tipo *CT* dove è presente del tessuto adiposo cardiaco (dati provenienti da [20, 21]), a destra abbiamo la ground truth rappresentante in rosso il tessuto adiposo nell'epicardio, in verde nel mediastino e in blu il gap tra i due. Nel mezzo si ha la ground truth binarizzata e a sinistra la segmentazione ottenuta con il software progettato utilizzando il metodo di *Otsu multilivello*, nel quale sono stati scelti opportunamente il numero di livelli e selezionato quello di interesse. In particolare è stata fatta una segmentazione basata su 7 livelli ed è stato selezionato il quarto. In Figura 5.2 possiamo invece osservare il confronto tra *Geodesic* a destra e la ground truth a sinistra, l'algoritmo è stato applicato utilizzando 30 iterazioni. L'immagine è una MRI raffigurante un tumore al cervello. Vediamo come anche i parametri rispecchiano le aspettative, infatti confermano l'impressione visiva di una soddisfacente segmentazione. Invece in Figura 5.3 vediamo un confronto tra due diversi metodi, *Fast Mar-*

ching a sinistra, con una soglia di 0.03, e *Geodesic* a destra, con un numero di iterazioni pari a 100. L'immagine è una CT che rappresenta un ictus al cervello, anche qui si vede la bontà dei risultati evidenziati dai parametri.

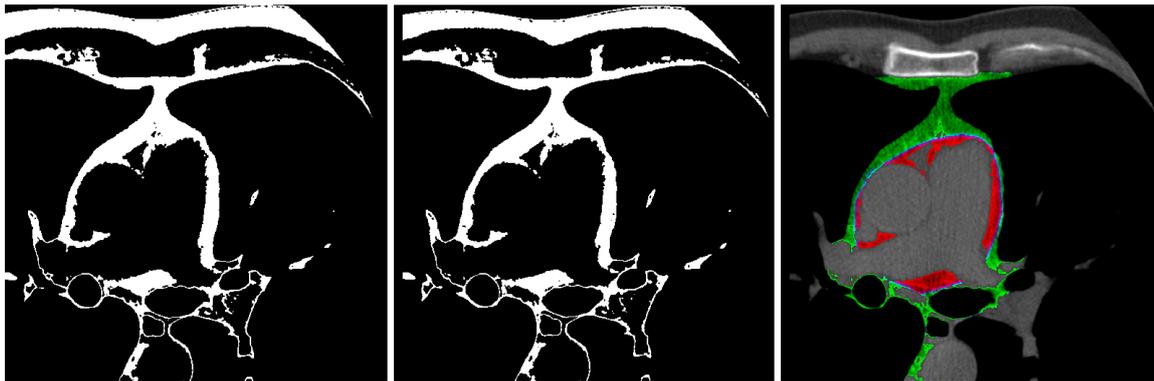
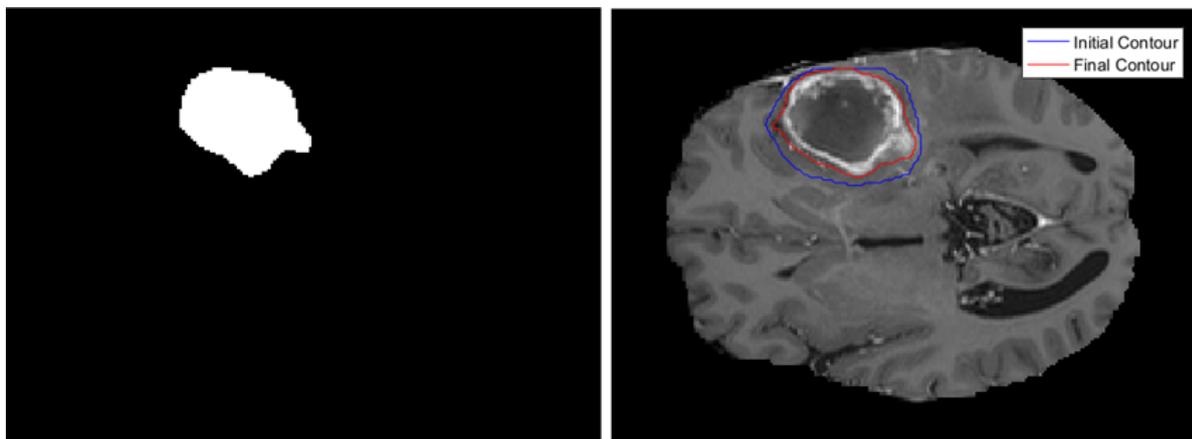


Figura 5.1: Confronto tra *Otsu multilivello* e *Ground Truth* di tessuto adiposo cardiaco su un'immagine CT.



Area	Dice	Hausdorff	PSNR	MSSD	Accuracy	TPR	TNR
0,051	0,972	3	26,465	2,074	0,998	0,998	0,998

Figura 5.2: Confronto tra *Ground Truth* (sinistra) e *Geodesic* (destra) di una MRI rappresentante un Tumore al cervello.

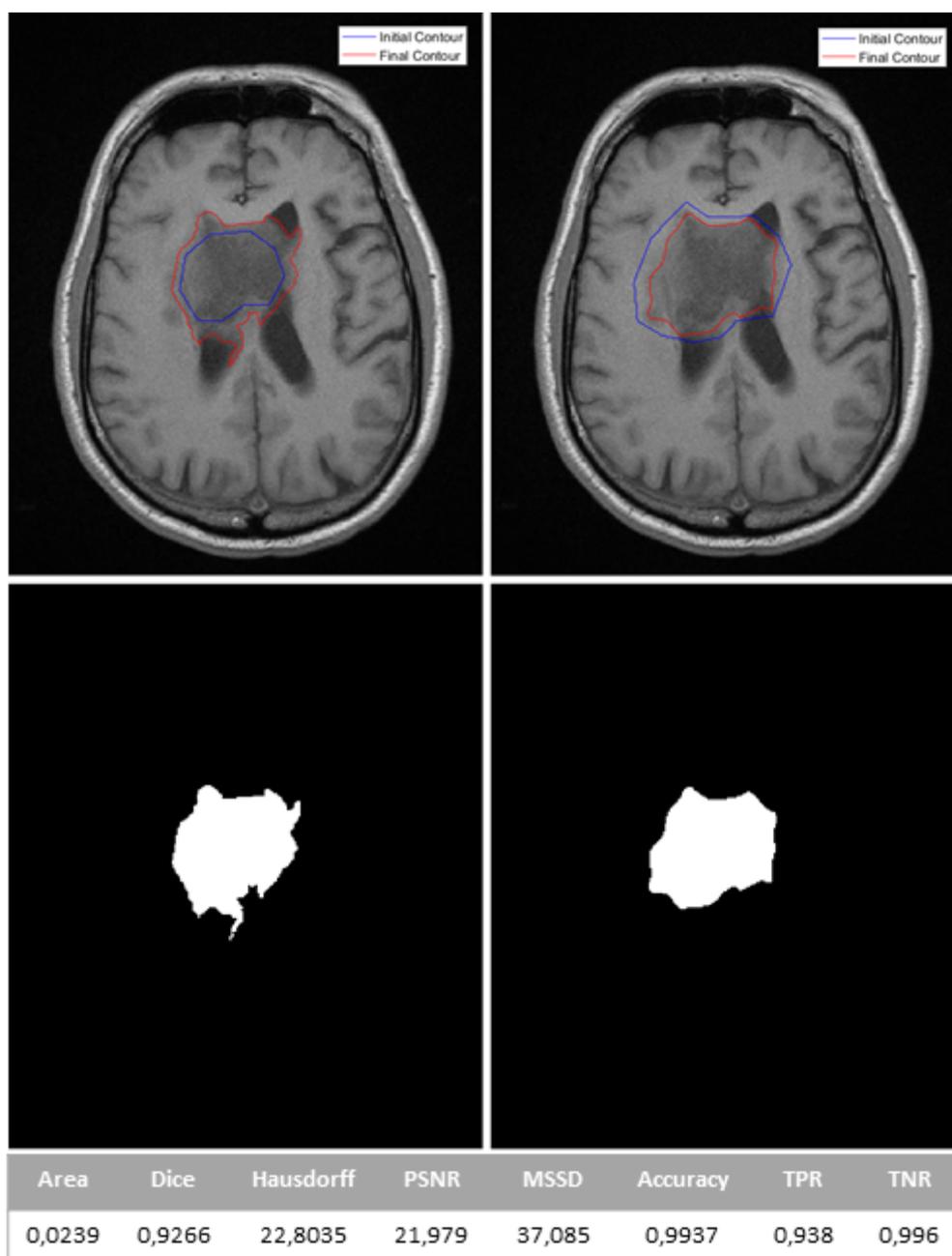


Figura 5.3: Confronto tra *Fast Marching* (sinistra) e *Geodesic* (destra) di un'immagine CT di un ictus al cervello.



## Capitolo 6

### Conclusioni

**L**a segmentazione è tra le elaborazioni delle immagini più articolate ed impegnative, è molto importante in quanto permette di riuscire a discriminare gli oggetti dallo sfondo in maniera automatica, trovando così molteplici campi applicativi. In particolare nella diagnostica per immagini la segmentazione viene utilizzata per evidenziare aree di notevole interesse medico. Normalmente questo lavoro spetta ad un'equipe di medici esperti del settore, impiegando quindi molto tempo e risorse, l'utilizzo di segmentatori automatici potrebbe aiutare il personale ad individuare un'area di interesse in maniera più efficace e rapida, rispetto alla segmentazione manuale. In questa tesi è stato presentato un software con interfaccia grafica in grado di segmentare immagini biomediche (anche se può essere utilizzato con qualsiasi immagine) applicando degli algoritmi noti in letteratura, sia di base che avanzati. Il segmentatore richiede un'interazione continua con l'utente che dovrà scegliere l'immagine da segmentare, l'algoritmo da applicare e i parametri da utilizzare nel metodo scelto. Il software prevede inoltre la possibilità di effettuare dei confronti tra diversi metodi applicati alla stessa immagine, oppure di confrontare i risultati forniti da un certo algoritmo con un'immagine di ground truth; il tutto con la possibilità di

salvare le segmentazioni visualizzate. Il segmentatore permette la messa a punto di alcuni parametri rilevanti attraverso l'interazione tra utente e algoritmo che quindi, avvenendo tramite interfaccia grafica, permette l'utilizzo anche a chi non può o non è in grado di fare uso di linguaggi di programmazione di alto livello. Nel confrontare due segmentazioni vengono forniti dei coefficienti solitamente usati per valutare la qualità della segmentazione, come ad esempio la differenza di area, il coefficiente di *Dice*, piuttosto che parametri come l'accuratezza, la sensibilità o la specificità. Il software sviluppato può inoltre essere impacchettato in un file eseguibile che, previa installazione della runtime *Matlab* fornita insieme al file eseguibile, può essere eseguito su un qualsiasi computer con sistema operativo *Windows*.

## 6.1 Applicazioni

Lo scopo di questo software è di fornire uno strumento di confronto e verifica nella segmentazione di immagini biomedicali, nonché semplificare la procedura di individuazione di aree di interesse e quindi la diagnosi. Perciò offre all'utilizzatore la possibilità di confrontare e valutare diverse tecniche di segmentazione o di misurare la bontà di una tecnica comparandola ad una ground truth. Il segmentatore può essere anche uno strumento per valutare le evoluzioni di talune patologie, ad esempio misurando e confrontando un'emorragia cerebrale o il tessuto adiposo cardiaco prima e dopo un trattamento, riuscendo così ad analizzare qualitativamente e quantitativamente, attraverso i parametri di confronto, eventuali miglioramenti.

## 6.2 Sviluppi futuri

Eventuali miglioramenti e sviluppi potrebbero consistere nell'ampliare il parco degli algoritmi di segmentazione applicabili, aggiungere la modifica per via grafica di ulteriori parametri e quindi dare la possibilità di intervenire più in dettaglio nell'applicazione degli algoritmi. Infine sarebbe interessante introdurre la possibilità di calcolare le misure reali delle aree individuate dalla segmentazione attraverso l'utilizzo di informazioni generate in fase di acquisizione, come ad esempio i dati contenuti nei *Voxel*.



## Bibliografia e Sitografia

- [1] Bioimmagini / Giuseppe Coppini, Stefano Diciotti, Guido Valli. - 3. ed Bologna : Pàtron, 2012.
- [2] Digital image processing / Rafael C. Gonzalez, Richard E. Woods. - 3. ed., inter. ed Upper Saddle River, NJ : Pearson Prentice-Hall, c2008.
- [3] Digital image processing using MATLAB / Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins Upper Saddle River : Pearson, 2004.
- [4] Nobuyuki Otsu, A threshold selection method from gray-level histograms, in IEEE Trans. Sys., Man., Cyber., vol. 9, 1979, pp. 62–66.
- [5] S. Osher, J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulation. Journal of Computational Physics, vol. 79, pp. 12-49, 1988.
- [6] T. F. Chan, L. A. Vese, Active contours without edges. IEEE Transactions on Image Processing, Volume 10, Issue 2, pp. 266-277, 2001.
- [7] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours. International Journal of Computer Vision, Volume 22, Issue 1, pp. 61-79, 1997.
- [8] A. Protiere and G. Sapiro, Interactive Image Segmentation via Adaptive Weighted Distances, IEEE Transactions on Image Processing, Volume 16, Issue 4, 2007.
- [9] Sethian, J. A. Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, Cambridge University Press, 2nd Edition, 1999.

- [10] Kass, M.; Witkin, A.; Terzopoulos, D. (1988). "Snakes: Active contour models". *International Journal of Computer Vision* 1 (4): 321. doi:10.1007/BF00133570.
- [11] Electronic text book developed by students enrolled in MDSC 689.01 at The University of Calgary in Fall Terms over 2013-15: [wiki.imagingkt.ca](http://wiki.imagingkt.ca)
- [12] Otsu - Matlab <http://it.mathworks.com/help/images/ref/graythresh.html>
- [13] Multi-threshold Otsu - Matlab <http://it.mathworks.com/help/images/ref/multithresh.html>
- [14] Logic Operations - Matlab <http://it.mathworks.com/help/images/examples/detecting-a-cell-using-image-segmentation.html>
- [15] Marker-Controlled Watershed Segmentation - Matlab <http://it.mathworks.com/help/images/examples/marker-controlled-watershed-segmentation.html?prodcode=IP&language=en>
- [16] Region Growing - Matlab <http://it.mathworks.com/matlabcentral/fileexchange/19084-region-growing>
- [17] Active Contour, Geodesic - Chan-Vese - Matlab <http://it.mathworks.com/help/images/ref/activecontour.html>
- [18] Fast Marching Level-set Method - Matlab <http://it.mathworks.com/help/images/ref/imsegfmm.html>
- [19] Snakes: Active Contour Models - Matlab <http://it.mathworks.com/matlabcentral/fileexchange/28109-snakes--active-contour-models>
- [20] <http://visual.ic.uff.br/en/cardio/ctfat/>
- [21] RODRIGUES, É.O.; MORAIS, F.F.C.; MORAIS, N.A.O.S.; CONCI, L.S.; NETO, L.V.; CONCI, A.; A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography. *Computer Methods and Programs in Biomedicine (Print)*, v. 123, p. 109-128, 2015.

## Ringraziamenti

Ringrazio il prof. Lamberti per la sua grande disponibilità, per l'aiuto datomi nell'indirizzarmi in questo lavoro e il sostegno morale.

Ringrazio la mia famiglia, Enzo, Luciana, Massimo e Livia. Loro sono le prime persone che hanno creduto in me, incoraggiandomi in questo insidioso percorso, sostenendo e rispettando le mie scelte. Senza di loro non sarebbe stato possibile per me raggiungere questo importante traguardo.

Ringrazio i miei parenti tutti, gli zii e i cugini presenti oggi per condividere questo giorno importante, per me sono stati come una seconda famiglia. Ringrazio mia nonna Maria e miei nonni che non ci sono più per tutto ciò che mi hanno insegnato.

Ringrazio Teresa, compagna da una vita, compagna della mia vita. Un punto di riferimento su cui, in questi anni, ho sempre potuto contare; con cui ho condiviso gioie, dolori, scelte e passioni. Una persona che si incontra una sola volta nella vita. Sempre presente, pronta a consigliarmi e ad ascoltarmi.

Ringrazio i vecchi e i nuovi amici incontrati durante questo percorso che nel bene o nel male mi hanno insegnato qualcosa, con i quali ci siamo so-

stenuti a vicenda anche con ironia e risate, ringrazio chi è rimasto fino alla fine.