



UNIVERSITÀ DEGLI STUDI DI PADOVA

CORSO DI LAUREA IN BIOINGEGNERIA

TESI DI LAUREA MAGISTRALE

COMPRESSED SENSING IN DSC-MRI

PARTE I: METODI

Relatore: **Prof. Enrico Grisan**

Correlatore: **Dott. Ing. Marco Castellaro**

Laureando: **Davide Boschetto**

Anno Accademico 2011-2012

A Te

Indice

Sommario	1
1 Introduzione	3
2 Teoria	7
2.1 NMR Imaging	7
2.1.1 Principi fisici della risonanza magnetica nucleare	8
2.1.2 k-Spazio	10
2.1.3 Sequenza Echo Planar Imaging	12
2.1.4 Dynamic Susceptibility Contrast - MRI	16
2.2 Compressed Sensing	17
2.2.1 Acquisizione	18
2.2.2 Sottocampionamento e Sparse Signal Recovery	20
2.2.3 Robustezza	22
2.2.4 Casualità	25
2.2.5 Osservazioni ed Applicazioni	27
2.3 L'utilizzo del CS in Risonanza Magnetica	28
2.3.1 Il dominio Wavelet	29
2.3.2 Acquisizione	31
2.3.3 Applicazioni del CS in MRI	36
3 Materiali e metodi	39
3.1 Algoritmo veloce per la minimizzazione TV-L1-L2	39
3.1.1 Introduzione del problema e trattazione matematica	39
3.1.2 Algoritmo risolutivo	42
3.1.3 Criteri di stop e condizioni di ottimalità	45

INDICE

3.2	Miglioramento del metodo	47
3.2.1	Aggiunta dell'informazione spaziale a priori	48
3.2.2	Modello matematico modificato	49
3.2.3	Aggiornamento dei criteri di stop	51
3.3	Dati a disposizione	53
3.4	Metodo di campionamento del k-space	54
3.5	Ambiente software	55
4	Risultati	59
4.1	Analisi statistica	59
4.1.1	Definizione degli errori	59
4.1.2	Analisi temporale su ROI	60
4.1.3	Scelta dei parametri	60
4.2	Risultati delle simulazioni	61
5	Discussione	77
5.1	Analisi visiva spaziale dei risultati	77
5.1.1	Ricostruzione a T_1	77
5.1.2	Ricostruzione a T_2	78
5.2	Andamento temporale della ricostruzione	79
5.2.1	ROI non appartenenti alla maschera	79
5.2.2	ROI appartenenti alla maschera	81
5.3	Riflessioni generali	82
6	Conclusioni e Sviluppi Futuri	85
A	Definizioni	A1
B	Elenco degli acronimi	A5
	Bibliografia	A7

Sommario

In questa tesi abbiamo sfruttato uno dei possibili approcci relativi all'utilizzo del compressed sensing in ambito di imaging a risonanza magnetica, cercando di migliorarlo ed adattarlo ad un utilizzo in esami dinamici con mezzo di contrasto, dove la presenza nei dati acquisiti dello stato basale garantisce un'informazione a priori assente negli esami MRI generali. Il metodo su cui ci siamo basati è quello di Yang (denominato *RecPF* [1]), che per migliorare l'acquisizione sfrutta un processo di minimizzazione di una funzione obiettivo composta da tre termini (TV e norme ℓ_1 ed ℓ_2) nel dominio wavelet. Creando una maschera spaziale, contenente le zone vascolari in cui vi è transito diretto di mezzo di contrasto, abbiamo introdotto un quarto termine alla minimizzazione, che varia di significato a seconda dei due metodi da noi derivati, chiamati *RecPFmod* e *RecPFalt*. Il termine aggiuntivo, in *RecPFmod*, minimizza la differenza dell'immagine ricostruita rispetto alla baseline nei punti non appartenenti alla maschera, mentre pesa l'informazione della baseline in maniera minore rispetto alla ricostruzione nei punti appartenenti alla maschera, visto il cambiamento di informazione temporale che ci si aspetta in queste regioni dell'immagine. In *RecPFalt*, invece, si introduce in entrambi i casi anche un contributo pesato derivante dalla ricostruzione al passo precedente. Questi tre metodi (l'originale e i due derivati dal nostro lavoro), dopo essere stati affinati e tarati mediante benchmarking, sono stati applicati su immagini simulate di DSC-MRI (Dynamic Susceptibility Contrast MRI) ottenute grazie all'utilizzo della funzione *phantom* di MATLAB e ad un fit di una gamma monovariata sui punti dell'immagine in comune con la maschera spaziale sopra definita, per simulare il passaggio del Gadolinio (mezzo di contrasto). I risultati ottenuti sono molto incoraggianti, suggeriscono la necessità di un test su dati reali

INDICE

ed evidenziano l'utilità del nostro metodo in applicazione a qualsiasi esame DSC-MRI previa definizione univoca ed arbitraria di una maschera spaziale che funga da prior. La definizione di questa maschera ricopre uno dei punti fondamentali del nostro algoritmo: in questa direzione dovranno concentrarsi gli sviluppi futuri, per riuscire ad implementare un metodo automatico per l'aggiornamento della maschera anche durante la fase di picco del mezzo di contrasto utilizzato.

Capitolo 1

Introduzione

Questa tesi rappresenta la prima parte di un più ampio progetto volto a verificare la robustezza e la solidità di un nuovo approccio alla risonanza magnetica con mezzo di contrasto sfruttando il metodo del Compressed Sensing.

La risonanza magnetica nucleare (NMR) è una tecnica di indagine basata sulla misura della precessione dello spin di protoni dotati di momento magnetico quando sono sottoposti ad un campo magnetico esterno. Essa viene sfruttata in medicina a scopi diagnostici nell'imaging a risonanza magnetica (MRI). Tale tecnica è considerata non dannosa nei confronti del paziente in quanto non prevede l'utilizzo di radiazioni ionizzanti, come invece accade nel caso delle tecniche facenti uso di raggi X o isotopi radioattivi (PET). Le informazioni ricavate da immagini RM sono di natura diversa rispetto a quelle degli altri metodi di imaging, in quanto questo tipo di esame permette la visualizzazione esclusiva dei tessuti molli, rendendo possibile una discriminazione fra di essi sulla base della loro composizione biochimica. L'acquisizione viene effettuata attraverso una successione di impulsi a radiofrequenza e di applicazione di gradienti, fasi che costituiscono le cosiddette sequenze di eccitazione. Gli impulsi a radiofrequenza servono a perturbare l'equilibrio degli spin (generando il segnale misurato), mentre i gradienti influenzano frequenza e fase del segnale ricevuto permettendo una correlazione fra misure ottenute e ricostruzione dell'immagine nel dominio spaziale. Al termine del processo di acquisizione, si ottiene uno spazio contenente coefficienti di Fourier denominato *k-space*. Tale spazio rappresenta la trasformata di Fourier bidimensionale dell'immagi-

Introduzione

ne MR misurata. Per riottenere quindi l'immagine di partenza è sufficiente applicare l'antitrasformata di Fourier al k -space.

Il Compressed Sensing (CS), invece, è una tecnica di acquisizione dei segnali alternativa a quella basata sul teorema di Nyquist-Shannon. Secondo la teoria del CS, sfruttando le proprietà di sparsità e incoerenza di alcuni segnali rappresentabili in una base opportunamente scelta, è possibile effettuare una ricostruzione ottima di tali segnali partendo da un set di campioni molto meno numerosi rispetto a quanto richiesto dal teorema di Nyquist-Shannon. Tale tecnica si basa sulla ridondanza tipica di molti segnali naturali: la maggior parte dei dati acquisiti, infatti, può essere scartata senza andare incontro ad una perdita di informazione significativa. Questo concetto sta alla base dei metodi di compressione *lossy*, fra cui JPEG, MP3 e MPEG. Applicando quindi l'approccio CS all'acquisizione di segnali generali che esibiscono una bassa coerenza con la base di sensing se rappresentati in una base opportuna, è possibile ricevere la gran parte dell'informazione utile del segnale originale utilizzando un esiguo numero di campioni. Secondo la teoria di Nyquist, però, è impossibile ricostruire accuratamente un segnale partendo da un set di misure altamente sottocampionato, dato che ci si troverebbe a dover risolvere un sistema di equazioni sotto-determinato e, quindi, a soluzioni infinite. Sfruttando alcune proprietà del segnale, però, una ricostruzione accurata è possibile attraverso un processo di minimizzazione di una funzione obiettivo, dipendente dal tipo di ricostruzione che si vuole ottenere.

A causa del fatto che le immagini MR sono ridondanti e acquisite in frequenza (nel k -space), il CS risulta una tecnica perfetta per il miglioramento dell'imaging a risonanza magnetica. La sua applicazione in questo campo potrebbe far risparmiare ai pazienti il fastidio relativo a sostenere esami lunghi e provanti, e ai sistemi sanitari ingenti quantità di denaro.

Questa tesi applica il CS a immagini simulate di risonanza magnetica dinamica con mezzo di contrasto. L'esame diagnostico da cui le immagini reali sono ricavate (DSC-MRI, cioè Dynamic Susceptibility Contrast MRI) viene solitamente applicato al cervello e sfrutta l'utilizzo di un agente paramagnetico (Gadolinio, Gd^{3+}) per distinguere le zone vascolarizzate da quelle non vascolarizzate. Il tracciante infatti attenua il segnale NMR, rendendo le zone vascolarizzate più scure. La particolarità di questo esame diagnostico consiste

nel fatto che l'acquisizione inizia prima dell'iniezione del mezzo di contrasto, in quanto è necessario acquisire anche lo stato basale in cui versa il paziente. La presenza del livello basale (baseline) fra i dati a disposizione ci ha fornito l'idea che sta alla base di questa tesi: creare una maschera che comprenda le zone vascolarizzate (quelle che saranno cioè interessate dal passaggio diretto del mezzo di contrasto) fornirebbe un'informazione spaziale a priori sulle zone dell'immagine contenenti l'informazione fondamentale da acquisire. L'acquisizione attraverso CS quindi risulterebbe ancor più ottimizzata, in quanto conoscere a priori le aree informativamente più importanti permetterebbe in teoria di redistribuire il range di frequenze fondamentali da acquisire per fare in modo che questo sia più concentrato sulle armoniche presenti nelle aree appartenenti alla maschera, tralasciando maggiormente le zone in cui l'informazione a priori ottenuta dalla baseline rimane teoricamente valida per tutto il corso dell'esame. Le aree non interessate dal passaggio del mezzo di contrasto, infatti, non dovrebbero idealmente esibire cambiamenti significativi e quindi, per esse, si potrebbe sfruttare l'informazione proveniente dalla baseline. La maschera, visto l'ambito simulativo in cui è inserita questa tesi, è stata stabilita a priori in maniera fissa e non adattativa. Sono state infatti decise a priori regioni esatte e con bordi ben definiti all'interno delle quali vi fosse passaggio simulato di mezzo di contrasto. Tale condizione è stata ottenuta fittando una gamma monovariata nelle regioni dell'immagine originale in comune con la maschera. In questa maniera si è creato un vettore di immagini che simulasse l'andamento temporale della luminosità dell'immagine tipico di un esame DSC-MRI. Infatti, l'intensità del segnale temporale è strettamente correlata alla luminosità dell'immagine: il passaggio di Gadolinio, ad esempio, causa un abbassamento dell'intensità del segnale misurato e, di conseguenza, una diminuzione della luminosità dell'immagine nei punti in cui tale segnale viene rilevato.

Il nostro lavoro pratico per implementare tale approccio si è basato sulla modifica di un algoritmo già esistente relativo all'applicazione generale del CS alla MRI basato sulle trasformate discrete Wavelet e di Fourier. Tale algoritmo prevede per la ricostruzione dell'immagine originale una minimizzazione di una funzione obiettivo formata da tre termini (il cui significato viene trattato dettagliatamente nel capitolo 3 e in appendice A), che rappresentano

Introduzione

rispettivamente:

1. la Total Variation (TV) del segnale: la presenza di questo termine nella minimizzazione permette alla ricostruzione di preservare l'informazione relativa ai bordi delle strutture in essa contenute;
2. il termine tipico della minimizzazione nella teoria del CS, riguardante la norma ℓ_1 del segnale;
3. la norma ℓ_2 della ricostruzione rispetto ai coefficienti di Fourier originali.

A questo algoritmo abbiamo aggiunto un quarto termine, contenente l'informazione relativa del nostro prior. A seconda del significato dato a questo termine aggiuntivo abbiamo ricavato due metodi che, nella ricostruzione dell'immagine, implementassero in maniera differente la nostra informazione a priori. In uno (*RecPFmod*) la ricostruzione sfrutta il solo prior ricavato dalla baseline per le zone non interessate direttamente dal mezzo di contrasto, in maniera pesata. Nell'altro metodo (*RecPFalt*), invece, si considera nella ricostruzione anche il contributo pesato dell'immagine ricostruita al tempo precedente.

Tali algoritmi, assieme a quello originale, sono stati applicati ad immagini simulate di sezioni di cervello ottenute con la funzione *phantom* di MATLAB, ricavata dagli studi di Shepp-Logan e largamente utilizzata dai ricercatori nel campo della tomografia. Lo scopo di tale applicazione è quello di verificare che tali approcci siano efficienti e funzionali prima della loro applicazione a dati reali. Sono stati utilizzati tre diversi sottocampionamenti del *k-space* (di fattori 4, 8 e 16) per valutare la fedeltà delle ricostruzioni ai dati originali. Dalle immagini ricostruite, inoltre, è stato misurato il segnale temporale relativo all'intensità dei pixel di quattro regioni d'interesse, due a due appartenenti allo stesso tessuto simulato. Le ROI appartenenti allo stesso tessuto sono state scelte in modo tale che una facesse parte della maschera mentre l'altra non fosse direttamente interessata dal passaggio del mezzo di contrasto. Tale confronto temporale ha permesso quindi un'analisi più dettagliata delle prestazioni dei tre algoritmi sotto inchiesta, che andasse oltre alla semplice analisi visiva delle ricostruzioni.

Capitolo 2

Teoria

2.1 NMR Imaging

Le tecniche di imaging a Risonanza Magnetica (MRI) si basano sul principio della Risonanza Magnetica Nucleare (NMR), una tecnica spettroscopica di imaging che consente di ricavare informazioni sulla natura chimico-fisica della materia grazie all'interazione a livello nucleare con i campi magnetici [2]. Questo fenomeno fu osservato per la prima volta nel 1946 da E. Purcell e F. Bloch [3] [4], valse loro il premio Nobel per la fisica nel 1952 e venne utilizzato nell'imaging in medicina a partire dagli anni Settanta. I due fisici osservarono che alcuni elementi, fra cui l'idrogeno, hanno deboli proprietà magnetiche quando si trovano in un forte campo magnetico e, in tali circostanze, possono scambiare energia con campi elettromagnetici di frequenze opportune: questa interazione con le onde radio viene chiamata *risonanza*. Essendo il corpo umano costituito prevalentemente da acqua e grasso, è possibile sfruttare questo fenomeno per ottenere immagini in cui si riescano a distinguere i tessuti in base alla loro composizione in idrogeno. Un sistema MRI acquisisce i segnali NMR provenienti dai singoli nuclei e combina le informazioni così ottenute fino a formare immagini relative al volume d'interesse. Quest'ultimo viene suddiviso in elementi tridimensionali fondamentali (cioè di dimensione minima), denominati *voxel*. Per ricostruire l'immagine di una determinata sezione si associa l'intensità del segnale NMR con l'intensità del pixel corrispondente, per ogni voxel. Di seguito si riportano in maniera generale i principi fisici che

stanno alla base del segnale NMR e le modalità di formazione dell'immagine finale.

2.1.1 Principi fisici della risonanza magnetica nucleare

Ogni nucleo atomico di idrogeno possiede un protone: la rotazione di questa particella attorno al proprio asse produce un vettore microscopico parallelo all'asse di rotazione che viene chiamato momento magnetico (μ). Questa proprietà magnetica viene definita *spin*: essa rappresenta il motivo per cui il protone tende ad allineare il suo momento magnetico (che, altrimenti, avrebbe direzione tridimensionale casuale) ad un campo esterno B_0 . Tale allineamento (che non è mai perfetto) può avvenire in due modi differenti, ognuno dei quali determina una configurazione energetica: ad alta energia (chiamata *spin-down*) o a bassa energia (definita *spin-up*). Il vettore μ può assumere quindi una posizione nello spazio tale da formare un angolo con B_0 : le due possibilità di allineamento, quindi, vengono identificate da due possibili angoli fra loro supplementari con cui si identificano i due stati energetici (effetto Zeeman). Il campo esterno esercita sugli spin anche un momento meccanico che produce in questi ultimi un moto di precessione intorno alle linee di flusso di B_0 , con una frequenza descritta dalla legge di Larmor:

$$\Omega_0 = \gamma B_0, \quad (2.1)$$

dove Ω_0 è la pulsazione di Larmor, γ è il rapporto giromagnetico e B_0 l'intensità del campo magnetico esterno.

Se si considerano tutti i nuclei immersi nel campo B_0 si osserva che il numero di spin a bassa energia (spin-up) è di poco superiore a quello degli spin ad alta energia (spin-down) ed il loro rapporto $\frac{N_{\text{up}}}{N_{\text{down}}}$ risulta definito dalla statistica di Boltzmann:

$$\frac{N_{\text{up}}}{N_{\text{down}}} = \exp\left(\frac{\Delta E}{kT}\right). \quad (2.2)$$

Il segnale proveniente da un voxel rappresenta quindi il comportamento della popolazione costituita dagli spin presenti in quella regione di spazio: l'NMR sfrutta quindi l'informazione del comportamento magnetico della popolazione di spin riportandola in scala macroscopica. La somma vettoriale

di tutti i singoli momenti magnetici è definita *magnetizzazione macroscopica netta* e, in assenza di campi esterni, è ovviamente un vettore nullo a causa dell'orientazione casuale della popolazione. Prendendo come riferimento la direzione del vettore B_0 (e facendola coincidere con l'asse Z), notiamo che nel momento in cui viene applicato tale campo gli spin subiscono un momento meccanico che determina sia un allineamento verso la direzione di B_0 sia un moto di precessione. Questi due effetti rendono il vettore di magnetizzazione risultante non nullo nella sua componente lungo l'asse Z (magnetizzazione longitudinale, M_z), a causa della prevalenza di nuclei con spin-up. La componente sul piano XY invece, dato che gli spin precessono alla stessa frequenza ma con fase casuale (come detto, infatti, B_0 agisce lungo l'asse Z), è nulla. Applicando un altro campo magnetico di stimolazione B_1 perpendicolare a B_0 accoppiato ad una radiofrequenza con pulsazione identica alla pulsazione di Larmor si ottiene l'effetto di risonanza, che causa una cessione di energia dal campo di stimolazione agli spin che, saturando il sistema, portano all'annullamento della componente di magnetizzazione longitudinale. La saturazione del sistema è causata dal passaggio degli spin da spin-up a spin-down grazie alla cessione di energia, fino ad arrivare alla condizione in cui $N_{\text{up}} = N_{\text{down}}$, con un conseguente annullamento di M_z . Inoltre si ha che l'azione di B_1 mette gli spin in concordanza di fase: dato che la precessione di tali spin continua attorno all'asse Z , si verifica la comparsa di una componente del vettore magnetizzazione risultante sul piano XY (magnetizzazione trasversale, M_{xy}). All'interruzione dell'applicazione del campo di stimolazione, il sistema tende a tornare alla configurazione di equilibrio attraverso due fenomeni: il rilassamento longitudinale (che comporta una graduale ricomparsa della magnetizzazione longitudinale) e il rilassamento trasverso (che, invece, causa una progressiva scomparsa della magnetizzazione trasversale). Entrambi questi processi hanno un andamento esponenziale con costanti di tempo che dipendono dal tipo di tessuto in esame: la costante T_1 regola il ritorno al valore originale di M_z ed è chiamata *tempo di rilassamento longitudinale* (o spin-reticolo), mentre la costante T_2 descrive l'abbattimento di M_{xy} e viene detta *tempo di rilassamento trasversale* (o spin-spin).

$$M_z = M_0(1 - e^{-\frac{t}{T_1}}), \quad (2.3)$$

$$M_{xy} = M_{xy0} e^{-\frac{t}{T_2}}. \quad (2.4)$$

Di fatto, però, sono due i fattori che causano il decadimento della magnetizzazione trasversale: le interazioni molecolari (che definiscono il T_2 puro) e la disomogeneità di B_0 (piccole variazioni del campo magnetico). Per tenere conto di entrambi gli effetti viene pertanto definita una nuova costante di tempo, T_2^* (*T2 star*):

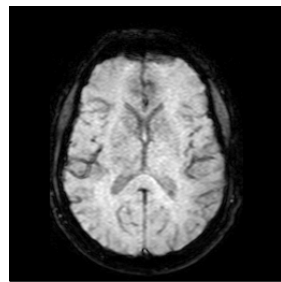
$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2^{(\text{disomog})}} = \frac{1}{T_2} + \frac{\gamma \Delta B_0}{2}. \quad (2.5)$$

Come si può intuire, la nuova costante di tempo è più bassa del T_2 puro: ciò significa che il fenomeno del rilassamento trasversale è accelerato dalla disomogeneità del campo magnetico. Questi tempi di rilassamento sono sperimentalmente utili: pesando infatti una scansione a T_1 si ottengono risultati diversi rispetto a quelli ottenuti con una pesatura in T_2 o in T_2^* . Nelle scansioni T_2 , ad esempio, tessuti ricchi di acqua o fluidi sono più luminosi rispetto a tessuti contenenti grasso, che sono principalmente scuri. Per le scansioni pesate a T_1 , invece, vale esattamente il contrario. Dato che i tessuti danneggiati tipicamente sviluppano edemi, la sequenza pesata a T_2 è maggiormente indicata per distinguere i tessuti patologici da quelli normali. Tali differenze si ottengono modulando i valori del tempo di eco (TE) e del tempo di ripetizione (TR), due parametri fondamentali nel processo di settaggio del macchinario.

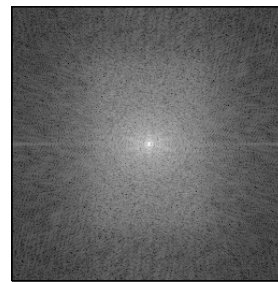
2.1.2 k-Spazio

Il segnale NMR rappresenta il segnale sinusoidale emesso dal sistema di spin dopo la stimolazione con campo B_1 e la formulazione che lo identifica contiene contributi dipendenti da T_1 , da T_2 e da ρ (densità protonica). Per la ricezione del segnale NMR si sfrutta una bobina per radiofrequenze (RF) posta nel piano XY : essendo M_{xy} tempo variante, per la legge di Faraday-Lenz (definizione 16 in appendice A) la bobina misura un segnale indotto (chiamato FID, *Free Induction Decay*) sinusoidale con frequenza $f = \frac{\Omega_0}{2\pi}$ e ampiezza esponenziale decrescente con $\tau=T_2$. Fino ad ora si è considerato il comportamento di una piccola popolazione di spin contenuta in un voxel, la quale produce complessivamente un singolo segnale FID. Le tecniche di imaging a risonanza magnetica si basano però sull'acquisizione di piani (fette

o *slice*) e di interi volumi. Il segnale FID, però, non contiene informazione sulla localizzazione spaziale del voxel che l'ha originato. Per risalire dai dati grezzi ad una mappa ordinata associata all'immagine si utilizza una tecnica chiamata *metodo di Fourier*: dopo aver demodulato il FID proveniente da un voxel (eliminando così le oscillazioni veloci dovute al campo di stimolazione) vengono combinate le informazioni sulla frequenza e la fase di ogni singolo FID (codifica in fase e in frequenza, quest'ultima ottenuta applicando dei gradienti di campo lungo gli assi). Così facendo si ottengono delle coppie di valori che identificano univocamente la posizione nel piano del voxel che ha generato quel particolare segnale. Al termine del processo, sommando i contributi di tutti i voxel, si ottiene uno spazio espresso in coefficienti di Fourier denominato *k-space*. Tale spazio, quindi, rappresenta i coefficienti della trasformata di Fourier 2D dell'immagine MR misurata. Per riottenere la nostra immagine è sufficiente quindi applicare la trasformata di Fourier inversa al *k-space*.



(a) Immagine MR



(b) *k-space*

Figura 2.1: Immagine MR (a) e relativo *k-space* (b)

A livello fisico esistono diverse modalità di acquisizione del *k-space*, che vengono chiamate *traiettorie*. Le più diffuse sono la traiettoria cartesiana, l'Echo Planar Imaging (*EPI*), la traiettoria radiale e quella a spirale. La più utilizzata è la modalità cartesiana, che acquisisce linee parallele ed equispaziate del *k-space* procedendo in maniera sequenziale. Il suo ampio utilizzo è dovuto alla buona qualità delle ricostruzioni che non soffrono di artefatti come nel caso delle altre traiettorie. L'utilizzo di tale traiettoria, però, implica tempi di acquisizione molto lunghi a causa della necessità di ripetere lo stimolo RF

per ogni linea acquisita. Un'altra tecnica molto utilizzata è l'EPI, sviluppata sulla base della traiettoria cartesiana, proprio allo scopo di ridurne i tempi di acquisizione. Essa permette infatti di coprire tutte le linee del k -space con pochi impulsi di eccitazione (*shots*) o utilizzando addirittura un singolo impulso (EPI single shot). Ad una maggior velocità, però, corrisponde lo svantaggio di avere una maggiore distorsione nell'immagine ricostruita. Tale tecnica può risultare molto utile in ambiti in cui è necessaria un'acquisizione molto veloce, come nel caso delle tecniche di imaging dinamico in cui il mezzo di contrasto è visibile solo per un paio di secondi.

Si consideri la traiettoria di acquisizione cartesiana: ogni linea di acquisizione del k -space corrisponde ad una lettura della codifica in frequenza per ogni valore del gradiente di codifica in fase. Tutte le linee sono parallele ed equispaziate. Nonostante ogni linea definisca una sorta di cammino continuo, l'acquisizione viene fisicamente realizzata campionando il segnale ad intervalli discreti lungo tale linea. Tradizionalmente questo campionamento deve rispettare il criterio di Nyquist, in modo che gli intervalli fra i campioni del k -space siano congrui alla banda del segnale. Alcune tecniche di imaging rapido (parallel-imaging) sfruttano bobine multiple di ricezione per sottocampionare il k -space e ricostruire rimuovendo gli artefatti di aliasing [5] [6].

Andando ad osservare le variazioni del k -space nel tempo, si ricade nel dominio definito k - t . Osservando invece la variazione temporale dell'immagine nel dominio spaziale ci si trova nello spazio x - t . Come si può vedere in Fig. 2.2, la rappresentazione nel dominio x - t può essere trasformata nello spazio x - f attraverso trasformazione di Fourier lungo la direzione temporale. Questi concetti risulteranno utili nei prossimi capitoli, per comprendere meglio il dominio a cui potranno appartenere eventuali segnali d'interesse.

2.1.3 Sequenza Echo Planar Imaging

Con lo sviluppo di nuove applicazioni in risonanza magnetica, come ad esempio l'imaging a perfusione e diffusione, l'imaging cardiaco e l'imaging real-time, è diventata una necessità disporre di tecniche di acquisizione abbastanza rapide da catturare le transizioni veloci del segnale analizzato. Una delle tecniche più rapide è, come già accennato, l'EPI: presentata da Peter Mansfield nel

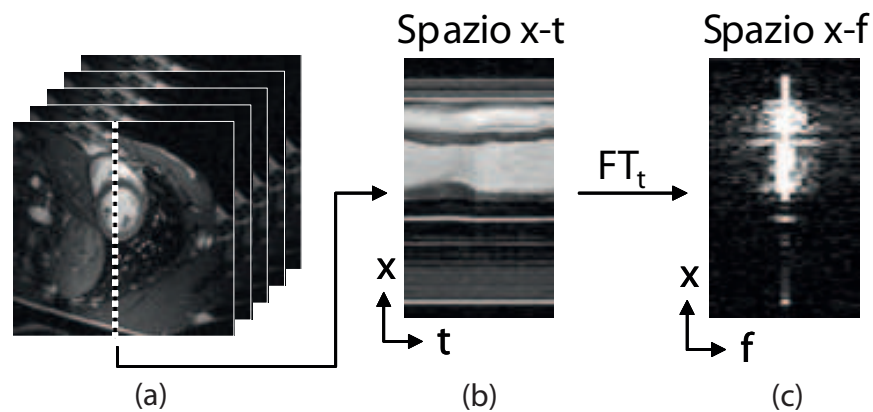


Figura 2.2: (a) Una sequenza temporale di immagini. (b) Rappresentazione nello spazio x-t di una colonna di pixel a tempi differenti. (c) Rappresentazione nello spazio x-f della stessa colonna a diversi tempi.

1977, questa scoperta lo portò a vincere il premio Nobel per la medicina nel 2003. Si tratta di una sequenza che utilizza in maniera alternativa i gradienti di codifica di fase, facendo ricorso per la lettura del segnale a serie di impulsi gradient-echo invece che spin-echo. Mentre quest'ultima sequenza utilizza due impulsi di eccitazione RF prima della lettura di un eco del FID (Fig. 2.3), la sequenza gradient-echo effettua la lettura dopo un singolo impulso di sfasamento (Fig. 2.4).

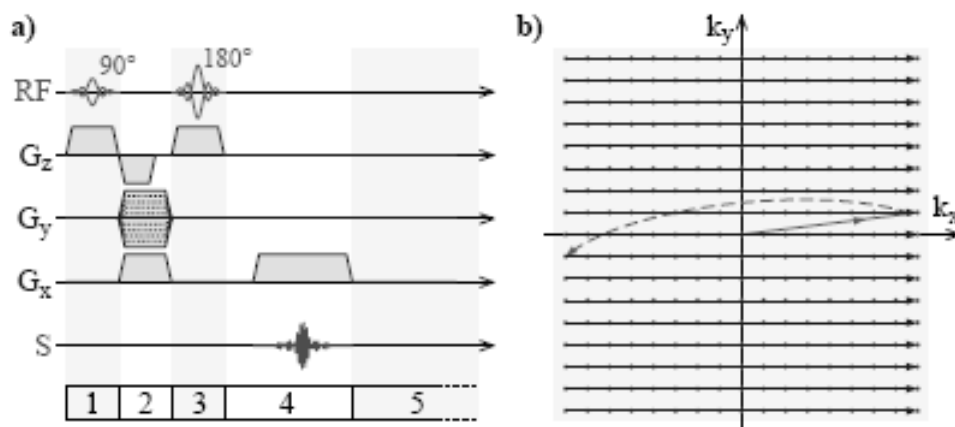


Figura 2.3: La serie di impulsi spin-echo.

La peculiarità della EPI (visibile in Fig. 2.5), sta nel fatto di utilizzare

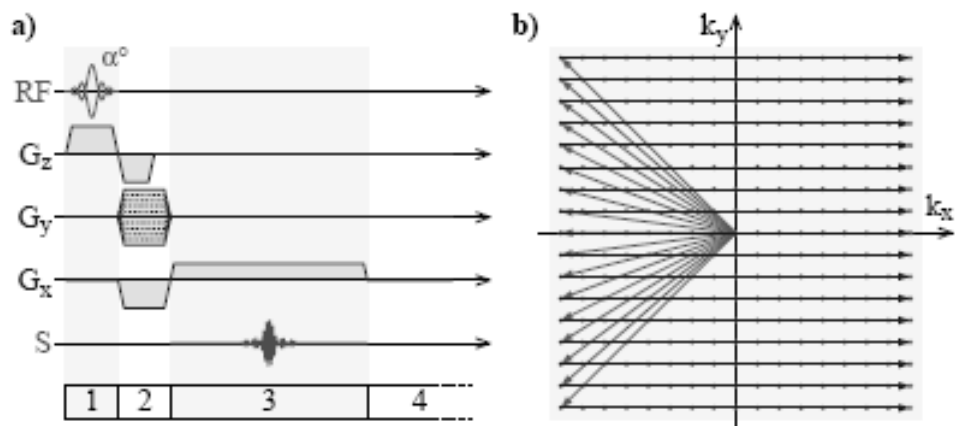


Figura 2.4: La serie di impulsi gradient-echo.

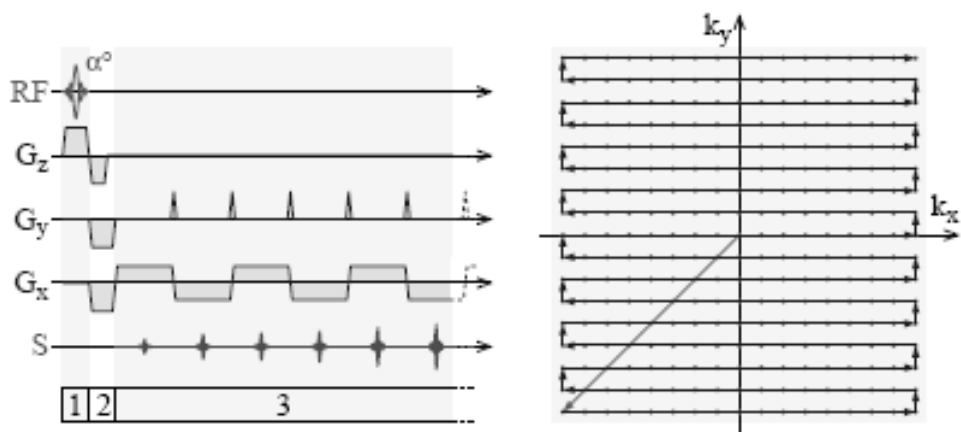


Figura 2.5: La sequenza di acquisizione Echo Planar Imaging.

dei gradienti di lettura bipolari che generano un treno di gradienti di eco: accoppiando un gradiente di codifica di fase è possibile acquisire molteplici linee del k -space con un singolo FID. Poiché per generare il FID si fa ricorso ad impulsi gradient-echo, la EPI accelera notevolmente il processo di acquisizione. Tipicamente un singolo impulso EPI genera circa 100 gradienti di eco: così è possibile ricostruire immagini a bassa risoluzione da una singola stimolazione RF (single shot-EPI). Questo però comporta anche dei problemi, legati alla presenza di artefatti nella ricostruzione delle immagini. Tali distorsioni, dette *ghost*, sono causate da diversi fattori fisici come ad esempio correnti parassite, risposte asimmetriche del filtro di anti-aliasing, ritardi nei gradienti non sincronizzati e isteresi. Inoltre, il decadimento di T_2^* durante la formazione del gradiente di eco provoca un *blurring* dell'immagine. È molto importante quindi ridurre gli artefatti tramite calibrazione della macchina e tecniche apposite di ricostruzione e post-processing.

Nel dettaglio, una normale sequenza gradient-echo genera con la stimolazione RF una magnetizzazione trasversale M_{xy} che decade con legge:

$$M_{xy}(t) = M_{xy}(0)e^{-t/T_2^*}, \quad (2.6)$$

dove T_2^* è già stata definita come la costante di tempo combinata degli effetti di T_2 e di disomogeneità di campo. L'emivita della magnetizzazione trasversale è data da $T_2^* \cdot \ln(2)$. Si suppone che tale M_{xy} possa essere utilizzata per generare segnale durante due emivite. Quindi, nel caso di un tessuto con T_2^* di circa 60 ms, la finestra di acquisizione utile è di $2 \cdot T_2^* \cdot \ln(2) \approx 83$ ms. Però, una tipica sequenza gradient-echo che acquisisce 256 punti per linea del k -space con ampiezza di banda del ricevitore pari a 64 kHz, impiega soltanto 2 ms di intervallo fra una stimolazione e l'altra (TR). Questo significa che solo una frazione dell'emivita di M_{xy} viene effettivamente utilizzata per l'acquisizione. La caratteristica della sequenza EPI è proprio quella sfruttare al massimo una singola magnetizzazione trasversale senza altre stimolazioni RF: a questo scopo un gradiente oscillante bipolare genera dei gradienti di eco prima che M_{xy} decada. Ogni gradiente di eco nella sequenza viene codificato in fase singolarmente ed è quindi possibile acquisire diverse linee del k -space con una singola stimolazione RF.

2.1.4 Dynamic Susceptibility Contrast - MRI

Le tecniche di imaging a risonanza magnetica, oltre al classico utilizzo nelle immagini anatomiche statiche, vengono applicate con successo a sequenze dinamiche per valutare e quantificare ad esempio il flusso ematico cerebrale ed altri parametri emodinamici. In queste applicazioni rientra la Dynamic Susceptibility Contrast (DSC-MRI), una tecnica che sfrutta un mezzo di contrasto paramagnetico (Gadolinio Gd-DTPA) per distinguere il tessuto vascolarizzato da quello non vascolarizzato. Una volta iniettato il tracciante nelle arterie che irrorano il cervello, infatti, il segnale NMR viene attenuato da questo permettendo di osservare le zone vascolarizzate come aree più scure. Tale processo viene registrato da un'acquisizione multipla in tempi consecutivi, ed il set di immagini che ne risulta consente la stima di alcuni parametri emodinamici. Una particolarità di questa tipologia di MRI consiste nel fatto che le acquisizioni non cominciano nel momento in cui viene iniettato il mezzo di contrasto. L'apparecchio infatti comincia a scansionare già qualche decina di secondi prima che ciò accada, per includere nell'esame diagnostico anche il livello basale (definito *baseline*). Tale caratteristica si rivelerà molto importante per questa tesi, come verrà approfondito a partire dai prossimi paragrafi e, in particolare, nel paragrafo 3.3.

Come già accennato nel paragrafo precedente, le sequenze utilizzate per la DSC-MRI sono delle EPI (Echo Planar Imaging), che grazie alla loro velocità nell'acquisire il *k-space* permettono di avere una risoluzione temporale accettabile (circa 1.5 secondi). L'utilizzo dell'EPI, però, non basta a rendere l'acquisizione sufficientemente ricca. A causa della cinetica veloce del Gadolinio, infatti, pur utilizzando l'EPI non si ottiene che qualche campione durante la fase informativa della cinetica del mezzo di contrasto, cioè la fase che permette di caratterizzare la curva temporale, ricavandone i parametri significativi. Per questo, il ricorso alla tecnica del Compressed Sensing, che verrà presentata nei paragrafi successivi, può permettere l'attuazione di un campionamento molto rado del *k-space*, aumentando sensibilmente la velocità di acquisizione.

2.2 Compressed Sensing

L'approccio tradizionale nel campo del campionamento dei segnali e delle immagini segue i dettami del Teorema di Nyquist-Shannon, il quale afferma che la minima frequenza necessaria per campionare un segnale analogico a banda finita e ricostruirlo correttamente è pari al doppio della sua frequenza massima. Soddisfacendo questa condizione infatti, si evita il fenomeno dell'aliasing e la conseguente distorsione del segnale causata appunto dalla sovrapposizione delle armoniche dello spettro del segnale campionato. Questo principio, in pratica, regola quasi tutti i protocolli di acquisizione del segnale in ogni campo (elettronica di consumo, strumenti di imaging medicale, ricevitori radio, e molti altri). Nella conversione di dati, ad esempio, la tecnologia dei convertitori standard analogico-digitali (ADC) implementa una rappresentazione quantizzata del teorema di Shannon: il segnale è campionato uniformemente alla frequenza di Nyquist (o a una maggiore).

La teoria del compressed sensing (CS) (emersa nei lavori [7], [8] e [9]) invece fornisce un approccio differente, che va contro le conoscenze comuni nel campo dell'acquisizione dei dati. Tale teoria infatti afferma che alcuni segnali o immagini possono essere ricostruiti completamente partendo da misure sottocampionate, le quali secondo il teorema di Nyquist-Shannon sarebbero altamente incomplete. Perché questo sia possibile, però, devono verificarsi due condizioni:

- Sparsità del segnale d'interesse: il concetto di sparsità esprime l'idea che le informazioni effettivamente contenute in un segnale a tempo continuo siano molto minori di quanto suggerito dalla sua lunghezza di banda, o che un segnale a tempo discreto dipenda da un numero di gradi di libertà molto minore del suo numero di campioni. Più precisamente, il CS sfrutta il fatto che molti segnali naturali sono sparsi o comprimibili, nel senso che sono rappresentabili in maniera compressa quando espressi in una base di sparsificazione adeguata Ψ .
- Incoerenza del campionamento: l'approccio CS richiede che l'acquisizione venga fatta seguendo l'incoerenza del segnale fra il dominio di acquisizione e la trasformata di sparsificazione. Un protocollo di *sensing* è quindi

progettato per catturare l'informazione più utile del segnale grazie a forme d'onda fissate, correlate al segnale stesso ma incoerenti con la base di sparsificazione.

L'osservazione cruciale consiste nel rendersi conto che per raggiungere protocolli di scansione e campionamento efficienti basterebbe teoricamente prendere l'informazione utile contenuta in un segnale sparso e condensarla in piccole quantità di dati. Questi protocolli non sono adattativi e richiedono semplicemente che il segnale correli con un piccolo numero di forme d'onda prestabilite che siano incoerenti con la base di sparsificazione.

Questo aspetto è fondamentale poiché consente di acquisire secondo un criterio prestabilito, quindi indipendentemente dall'andamento del segnale. In altre parole il CS può essere visto come un protocollo di acquisizione di segnali semplice ed efficiente che sfrutta la natura sparsa del segnale da acquisire per ricostruirlo correttamente.

2.2.1 Acquisizione

In generale, l'informazione relativa ad un segnale $f(t)$ è ottenuta da funzionali lineari φ_k che registrano i valori

$$y_k = \langle f, \varphi_k \rangle, \quad k = 1, \dots, m, \quad (2.7)$$

dove con la notazione $\langle a, b \rangle$ si indica l'applicazione di un funzionale b ad un oggetto a d'interesse. Così facendo, correlando l'oggetto che vogliamo acquisire con le forme d'onda $\varphi_k(t)$ otteniamo l'acquisizione discreta di nostro interesse. Se, ad esempio, le forme d'onda fossero impulsi (delta di Dirac) allora y sarebbe un vettore di valori campionati di f nel tempo o nello spazio. Se le forme d'onda invece fossero funzioni che rappresentano il pixel da campionare, allora y sarebbe l'immagine che tipicamente viene immagazzinata da un sensore in una fotocamera. Se si utilizzassero delle sinusoidi allora y rappresenterebbe un vettore dei coefficienti di Fourier (come avviene nel caso dell'imaging a risonanza magnetica). L'interesse del CS è indirizzato verso situazioni di sottocampionamento in cui il numero m dei campioni disponibili è molto minore della dimensione n del segnale f . Questo problema è molto comune nell'ambito

pratico, perché o vi è un numero limitato di sensori o le misure sono troppo dispendiose nel tempo, non permettendo di avere un numero soddisfacente di dati a disposizione. Secondo la teoria classica sarebbe impossibile ricostruire accuratamente f partendo da $m \ll n$ misure, dato che ci si troverebbe a dover risolvere un sistema di equazioni sotto-determinato (cioè con un numero di incognite maggiore del numero di equazioni).

Indicando con A la matrice $m \times n$ di acquisizione che ha come righe i vettori $\varphi_1^*, \dots, \varphi_m^*$ (definendo φ_i^* il complesso coniugato di φ_i), il processo di ricostruzione di f da $y = Af \in \mathbb{R}^m$ è malposto nel caso in cui $m < n$: ci sono infatti infinite soluzioni \hat{f} che soddisfano $A\hat{f} = y$.

Sparsità

Molti dei segnali naturali hanno una rappresentazione compatta se espressi in una base accuratamente scelta. Ciò significa che i coefficienti dei funzionali in tale base sono per la maggior parte uguali o prossimi allo zero. Questa caratteristica ben nota sta alla base di tutti i paradigmi di compressione dei segnali. Matematicamente, un vettore $f \in \mathbb{R}^n$, dove n è il numero degli elementi del vettore, si può espandere in una base ortonormale Ψ come segue:

$$f(t) = \sum_{i=1}^n x_i \psi_i(t), \quad (2.8)$$

con x rappresentante la sequenza di coefficienti di f , $x_i = \langle f, \psi_i \rangle$. Sarà quindi conveniente esprimere f come Ψx , dove Ψ è la matrice $n \times n$ avente come colonne i vettori ψ_1, \dots, ψ_n . L'ipotesi di sparsità è ora chiara: quando un segnale può essere espanso in rappresentazione sparsa, si possono scartare i coefficienti "piccoli" senza troppa perdita di informazione. Formalmente, considerando l'espansione 2.8, f_S viene ottenuta mantenendo solo gli S coefficienti più grandi di x_i . Pertanto $f_S := \Psi x_S$, dove x_S è il vettore dei coefficienti di partenza x_i , in cui però tutti tranne gli S coefficienti più grandi sono stati posti a zero. Questo vettore è sparso in senso stretto, dato che quasi tutti i suoi coefficienti sono nulli: chiameremo S -sparso un oggetto con al più S valori diversi da zero. Dato che Ψ è una base ortonormale, abbiamo che $\|f - f_S\|_2 = \|x - x_S\|_2$ e se x è sparso o comprimibile allora esso è ben approssimato da x_S , quindi l'errore $f - f_S$ è piccolo. Quindi, si può non considerare buona parte dei

Teoria

coefficienti senza perdere troppa informazione. La sparsità, quindi, è una caratteristica fondamentale che permette una stima statistica accurata e un'efficiente compressione dei dati.

Incoerenza

Supponiamo di avere una coppia (Φ, Ψ) di basi ortonormali in \mathbb{R}^n . La prima base Φ è usata per acquisire l'oggetto f come visto in 2.7, e la seconda è usata per rappresentare f .

Definizione 1. *La coerenza fra la base d'acquisizione Φ e quella di rappresentazione Ψ è*

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \leq k, j \leq n} |\langle \varphi_k, \psi_j \rangle|. \quad (2.9)$$

La coerenza misura la massima correlazione fra due elementi qualsiasi di Φ e Ψ (si rimanda a [10] per approfondimenti). Se i due vettori contengono elementi correlati, la coerenza è alta. In caso contrario, è bassa. Quindi, in base alla formula 2.9, la coerenza può assumere valori da 1 a \sqrt{n} . Il CS utilizza soprattutto coppie di basi con bassa coerenza. Matrici casuali sono molto incoerenti con qualsiasi base Ψ fissata. Scegliendo casualmente una base Φ uniforme, con alta probabilità la coerenza fra Φ e Ψ sarà di $\sqrt{2 \log n}$. Quindi, anche forme d'onda casuali ($\varphi_k(t)$) a valori indipendenti e identicamente distribuiti (*iid*, come ad esempio Gaussiane o ingressi binari) esibiranno una coerenza molto bassa con qualsiasi Ψ . Quindi, se il processo di acquisizione utilizzando sistemi incoerenti è buono, un meccanismo efficiente è quello che punta ad ottenere alte correlazioni con forme d'onda casuali come, ad esempio, rumore bianco.

2.2.2 Sottocampionamento e Sparse Signal Recovery

Anche se idealmente vorremmo acquisire tutti gli n coefficienti di f , in realtà riusciamo a osservarne solo un sottoinsieme e ad acquisire le misure

$$y_k = \langle f, \varphi_k \rangle, \quad k \in M, \quad (2.10)$$

dove $M \subset 1, \dots, n$ è un sottoinsieme di cardinalità $m < n$. Il paradigma fondamentale del CS afferma che si può ricostruire il segnale dalle sue M

2.2 Compressed Sensing

misure attraverso una minimizzazione di norma ℓ_1 : la ricostruzione proposta f^* è data da $f^* = \Psi x^*$, dove x^* è la soluzione dell'ottimizzazione convessa

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1 \quad \text{soggetta a} \quad y_k = \langle \varphi_k, \Psi \tilde{x} \rangle, \quad \forall k \in M. \quad (2.11)$$

Fra tutte le \tilde{f} consistenti coi dati, scegliamo quelle le cui sequenze di coefficienti della base sparsa hanno la norma ℓ_1 minima. La minimizzazione di una norma ℓ_1 soggetta a vincoli di uguaglianza lineari può essere vista come un caso di programmazione lineare. Per esso, quindi, si possono stabilire algoritmi risolutivi molto efficienti. Il metodo della minimizzazione della norma ℓ_1 per promuovere la sparsità viene utilizzato da diversi decenni, ma non è l'unico modo in cui agire per ottenere una sparse recovery: negli anni infatti si sono affinati anche altri metodi, fra cui l'uso degli algoritmi greedy [11]. I risultati ottenuti da [12] e [13] confermano che quando f è abbastanza sparsa, la ricostruzione attraverso norma ℓ_1 è affidabile.

Teorema 1. [14] *Dato $f \in \mathbb{R}^n$ e supposto che la sequenza di coefficienti x di f nella base Ψ sia S -sparsa, scelgo m misure nel dominio Φ in maniera casuale. Se*

$$m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log n \quad (2.12)$$

per una costante positiva C , allora la soluzione a 2.11 è esatta con probabilità quasi assoluta, cioè superiore a $1-\delta$ se $m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log(n/\delta)$, con $0 < \delta \ll 1$.

Dalla 2.12 si evince che a coerenza minore sono necessari meno campioni: questo spiega il perché si cerchi di lavorare con sistemi a bassa coerenza. Viene inoltre dimostrata l'efficacia teorica del sottocampionamento oltre la soglia di Nyquist: basta infatti campionare un qualsiasi set di m componenti per evitare di ottenere perdita di informazione. Se la coerenza si avvicina a 1, allora bastano circa $S \log n$ campioni. La ricostruzione del segnale f a partire dai dati acquisiti non richiede alcuna conoscenza a priori sul numero di campioni di f non nulli, sulla loro posizione o sull'ampiezza. Se il segnale è abbastanza sparso, la ricostruzione sarà esatta.

Il ruolo della probabilità

Per ottenere risultati attendibili è necessario introdurre una descrizione probabilistica della ricostruzione tramite norma ℓ_1 . Infatti i risultati, che si sono introdotti nel paragrafo precedente, non sono comparabili fra diversi data-set, anche se tutti di cardinalità m . Esistono segnali sparsi speciali che nel dominio Φ hanno la maggior parte dei campioni nulli. In altre parole, si possono identificare segnali f sparsi e sottoinsiemi molto grandi di cardinalità circa n (esattamente $n - S$) per cui $y_k = 0$ per tutti i $k \in M$. Come si può intuire, non esiste algoritmo in grado di ricostruire un segnale partendo da tali insiemi, che presentano una percentuale elevata di elementi posti a zero. Dall'altro lato, il Teorema 1 garantisce che la frazione di dataset per cui non si ha una ricostruzione esatta è veramente trascurabile (una potenza di n con modulo grande e segno negativo). Quindi dobbiamo tollerare una possibilità di insuccesso estremamente piccola che nella pratica è approssimabile a zero, se la dimensione del campionamento è sufficientemente grande. Gli studi [7] e [15] hanno dimostrato che si ha bisogno di un numero di campioni almeno dell'ordine di $\mu^2 \cdot S \cdot \log n$. Con meno campioni, infatti, la probabilità che l'informazione possa smarrirsi è troppo elevata e la ricostruzione è non solo improbabile, ma persino impossibile con qualsiasi metodo. Quindi, quando la coerenza è 1, non abbiamo bisogno di un numero maggiore di $S \cdot \log n$ campioni, ma non possiamo contemporaneamente averne di meno.

Portiamo l'esempio di uno schema di campionamento incoerente, considerando un segnale bidimensionale (1MPixel) che ha circa 25.000 campioni wavelet diversi da zero. Acquisendo 96.000 misure incoerenti e risolvendo l'equazione 2.11, la ricostruzione ℓ_1 è perfetta: $f^* = f$. Questo esempio mostra come un numero di campioni pari a circa quattro volte il livello di sparsità sembri bastare: molti altri studi hanno riportato successi analoghi utilizzando questa indicazione puramente sperimentale (fra gli altri [14]).

2.2.3 Robustezza

Per essere sfruttato a dovere, il CS deve saper gestire situazioni in cui ci siano sia segnali quasi-sparsi sia rumore. In ambito pratico, infatti, gli oggetti d'interesse non sono mai esattamente sparsi ma sempre quasi-sparsi. Inoltre,

qualsiasi applicazione pratica soffrirà di un certo livello di rumore, dato che gli strumenti di misura non hanno precisione infinita. È imperativo perciò che il CS sia robusto in presenza di entrambe queste complicazioni: esso infatti deve almeno garantire che piccole perturbazioni dei dati si trasformino al massimo in piccole perturbazioni sulla ricostruzione, riducendo il mal-condizionamento. Il nostro problema di ricostruzione di un vettore $x \in \mathbb{R}^n$ parte da dati:

$$y = Ax + z, \quad (2.13)$$

dove A è una matrice di acquisizione $m \times n$ e z rappresenta l'errore (deterministico oppure stocastico e sconosciuto). Dato che $f = \Psi x$ e $y = R\Phi f$ (R è la matrice $m \times n$ che estrae le coordinate campionate in M), si può indicare A come $R \cdot \Phi \cdot \Psi$. Così facendo, posso utilizzare il modello 2.13 tenendo in mente che x non è altro che la sequenza dei coefficienti dell'oggetto di interesse nella base appropriata.

Isometrie Ristrette

Questa sezione introdurrà dei concetti chiave che sono molto utili per studiare la robustezza generale del CS: la cosiddetta Proprietà di Isometria Ristretta (RIP) [16].

Definizione 2. Per ogni intero $S = 1, 2, \dots$, definiamo la costante isometrica δ_S di una matrice A come il numero più piccolo che permetta a

$$(1 - \delta_S)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_S)\|x\|_2^2 \quad (2.14)$$

di essere vera per ogni vettore x S -sparso.

Diremo che una matrice A rispetta la RIP di ordine S se δ_S non è troppo vicino a 1. Quando questa proprietà viene rispettata, A mantiene la distanza Euclidea di un segnale S -sparso. Questo fatto implica a sua volta che vettori S -sparsi non possono far parte del kernel di A (chiamando cioè x un vettore S -sparso, non può essere che $Ax = 0$). Per vedere ancor meglio la connessione fra RIP e il CS, immaginiamo di dover acquisire segnali S -sparsi con A , supponendo che δ_{2S} sia sufficientemente minore di 1. Questo implica che tutte le

Teoria

distanze fra segnali S-sparsi devono mantenersi nello spazio di misura, cioè deve verificarsi che

$$(1 - \delta_2 S) \|x_1 - x_2\|_2^2 \leq \|Ax_1 - Ax_2\|_2^2 \leq (1 + \delta_2 S) \|x_1 - x_2\|_2^2$$

si mantenga per qualsiasi coppia di vettori S-sparsi x_1, x_2 . Come verrà dimostrato di seguito, questo fatto garantisce l'esistenza di un algoritmo robusto ed efficiente per discriminare segnali S-sparsi a seconda delle loro misurazioni compresse.

Se la RIP è verificata, allora la minimizzazione ℓ_1

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1 \quad \text{soggetta a} \quad A\tilde{x} = y (= Ax) \quad (2.15)$$

fornisce un'accurata ricostruzione.

Teorema 2. [17] *Assumendo che $\delta_{2S} < \sqrt{2} - 1$, la soluzione x^* a 2.15 soddisfa*

$$\|x^* - x\|_2 \leq C_0 \cdot \|x - x_S\|_1 / \sqrt{S} \quad \text{e} \quad \|x^* - x\|_1 \leq C_0 \cdot \|x - x_S\|_1 \quad (2.16)$$

per una costante C_0 , dove x_S è il vettore x con tutte le componenti a zero tranne le S maggiori [17], [18].

Le conclusioni di questo teorema sono più forti di quelle del Teorema 1. Se x è S-sparsa allora $x = x_S$ e, quindi, la ricostruzione è esatta. Il Teorema 2, però, tratta di qualsiasi tipo di segnale. Se x non è S-sparsa, allora esso dice che la qualità del segnale ricostruito è tanto buona quanto lo sarebbe stata se si fosse conosciuta a priori la posizione degli S valori maggiori di x (cioè come se questi si fossero misurati direttamente). In altre parole è come se avessimo una conoscenza perfetta di x , estraendo gli S valori con l'informazione più significativa. Un'altra sostanziale differenza fra Teorema 1 e Teorema 2 è che quest'ultimo è deterministico e non probabilistico. Ciò significa che trovando una matrice A che obbedisca alle ipotesi del teorema e utilizzandola, abbiamo la certezza di ricostruire tutti i vettori S-sparsi esattamente, o comunque gli S campioni più grandi di qualsiasi vettore.

Rumore

Con la presenza di dati rumorosi come in 2.13, dobbiamo rilassare i vincoli lineari della ricostruzione

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1 \quad \text{soggetta a} \quad \|A\tilde{x} - y\|_2 \leq \varepsilon \quad (2.17)$$

utilizzando un $\varepsilon > 0$, indicativo della quantità di rumore nei dati. Tale problema, in letteratura noto come LASSO (Least Absolute Shrinkage and Selection Operator [19]), è un problema convesso e come tale può essere risolto efficientemente.

Teorema 3. [17] *Assumendo che $\delta_{2S} < \sqrt{2} - 1$, la soluzione x^* a 2.17 soddisfa*

$$\|x^* - x\|_2 \leq C_0 \cdot \|x - x_S\|_1 / \sqrt{S} + C_1 \cdot \varepsilon \quad (2.18)$$

per un certo valore di C_0 e C_1 .

La situazione è di facile comprensione: l'errore di ricostruzione è legato alla somma di due termini. Il primo è l'errore che ci sarebbe se si avessero dati non rumorosi, mentre il secondo è proporzionale al livello di rumore stesso. Inoltre, le costanti C_0 e C_1 sono tipicamente piccole. Da dati sperimentali, l'applicazione di questo modello conferma che il CS è un meccanismo di acquisizione pratico e robusto, che funziona con qualsiasi tipo di segnale (non necessariamente con segnali sparsi) e che gestisce al meglio segnali rumorosi [13] [12].

2.2.4 Casualità

Ciò che deve essere approfondita ora è la relazione fra S (numero di componenti che è possibile ricostruire) e m (numero di righe della matrice A , o numero delle misure): è possibile avere una matrice di acquisizione che soddisfi la RIP e che abbia un valore di S vicino a m ? Ciò è possibile, ma è indispensabile riprendere in considerazione la casualità. Consideriamo a questo scopo le diverse matrici di acquisizione A costruite attraverso le seguenti modalità:

1. campionando n colonne uniformemente e casualmente dalla sfera unitaria di \mathbb{R}^m ;
2. campionando a valori *iid* dalla distribuzione normale a media nulla e varianza $1/m$

Teoria

3. Campionando una proiezione casuale P come in 2.2.1 e normalizzando:

$$A = P\sqrt{n/m};$$

4. Campionando a valori *iid* da una distribuzione simmetrica di Bernoulli o da altre distribuzioni sub-gaussiane.

Con probabilità quasi certa (per i primi tre punti utilizzando la teoria della probabilità generale, per il quarto utilizzando in particolare [20]), tutte queste matrici soddisfano la RIP, dato che

$$m \geq CS \log(n/S), \quad (2.19)$$

dove C è una costante dipendente dalla matrice scelta. In tutti questi esempi, la probabilità di campionare una matrice che non soddisfi la RIP quando 2.19 è soddisfatta è esponenzialmente piccola in m . Inoltre, non esistono matrici di misure o algoritmi di ricostruzione che diano le conclusioni del Teorema 2 utilizzando meno campioni che in 2.19 [9], [8]. Quindi, l'utilizzo della minimizzazione ℓ_1 utilizzando matrici randomizzate è una strategia di acquisizione quasi ottima. Per verificare la RIP per una coppia di basi ortonormali, come il caso in cui $A = R\Phi\Psi$, dove R estrae m coordinate uniformemente e casualmente, è sufficiente avere

$$m \geq CS \log^4 n \quad (2.20)$$

perché la proprietà regga con elevata probabilità [8],[21]. Se si vuole una probabilità di fallimento minore di $O(n^{-\beta})$ con $\beta > 0$, allora il miglior esponente in 2.20 è cinque invece di quattro, anche se la 2.20 reggerebbe anche con esponente pari a uno. Questo prova che è possibile ricostruire stabilmente e accuratamente segnali quasi sparsi da set di dati molto sottocampionati in un dominio incoerente. Inoltre, la RIP regge anche per matrici di sensing $A = \Phi \cdot \Psi$, dove Ψ è una base ortonormale arbitraria e Φ è una matrice di misura $m \times n$ ricavata casualmente da una distribuzione adatta. Se si scelgono Ψ e Φ seguendo uno dei metodi sopra citati, allora con probabilità quasi totale la matrice A soddisfa la RIP se la 2.19 è soddisfatta. Queste matrici di misure casuali Φ sono in un certo senso universali [20]: non è necessario infatti conoscere la base di sparsificazione quando viene scelto il sistema di misura.

2.2.5 Osservazioni ed Applicazioni

L'acquisizione solitamente funziona in questo modo: una quantità massiccia di dati viene acquisita per poi essere, in gran parte, scartata durante la compressione per facilitare trasmissione e conservazione dei dati stessi. Tecnicamente, si acquisisce un array di pixel f ad alta risoluzione, si crea il set completo di coefficienti della trasformata, quindi si codificano i coefficienti più grandi e si scartano tutti gli altri, ottenendo alla fine f_S . Questo processo di acquisizione completa seguito da compressione è estremamente dispendioso. Il CS si comporta in maniera molto differente, come se fosse possibile acquisire direttamente solo l'informazione importante dell'oggetto di interesse. Utilizzando solo $O(S \log(n/S))$ proiezioni casuali si hanno abbastanza informazioni per ricostruire il segnale con un'accuratezza almeno grande quanto quella fornita da f_S (la miglior rappresentazione compressa). In altre parole, i protocolli di misura del CS traducono dati analogici in una forma digitale già compressa, così da poter (in linea di principio) ottenere segnali a risoluzione massima da molti meno sensori: tutto ciò che serve fare dopo il passo di acquisizione è la ricostruzione delle misure (passo equivalente alla decodifica nel paradigma classico della compressione). Dato che segnali comprimibili possono essere acquisiti efficientemente utilizzando un numero di misure proporzionale al livello di informazione del segnale $S \ll n$, la tecnica del CS può essere utilizzata in moltissime applicazioni di uso comune:

- **Compressione dei dati:** In alcune situazioni, Ψ può essere sconosciuta o non pratica da implementare per scopi di compressioni di dati. Come discusso in precedenza, però, una Φ scelta casualmente può essere considerata una tecnica di codifica universale, dato che per essere scelta non ha bisogno di informazioni a priori sulla struttura di Ψ (tale informazione è richiesta solo per la decodifica o la ricostruzione di f).
- **Problemi inversi [16]:** l'unico modo di acquisire f può essere quello di usare un sistema di misura Φ con modalità particolari. Supponendo l'esistenza di una Ψ per f che sia anche incoerente con Φ , un'acquisizione efficiente è possibile. Un esempio di tale applicazione riguarda l'angiografia in risonanza magnetica [7] ed altri tipi di setup MR, dove Φ registra un sottoinsieme nel dominio delle frequenze, e l'immagine desiderata f è

sparsa nel tempo o in wavelet. In tale categoria rientra il lavoro, fra gli altri, di Lustig [12], presentato nel prossimo paragrafo.

- **Acquisizione di dati:** In alcune situazioni l'intera collezione di n campioni a tempo discreto di un segnale analogico potrebbe essere difficile da ottenere e comprimere. Per questo potrebbe essere utile creare degli strumenti che campionino fisicamente e registrino direttamente misure discrete incoerenti del segnale analogico d'interesse.

Quest'ultima applicazione suggerisce che metodi matematici come il CS possano avere un impatto immenso in aree dove l'hardware convenzionale oggi utilizzato ha limiti evidenti. Ad esempio, nel campo di acquisizione di immagini, gli strumenti che usano sensori CCD o CMOS sono limitati dalla lunghezza di banda dello spettro visibile. Una "fotocamera CS" che acquisisse misure non coerenti usando un array digitale (che contenga un solo elemento fotosensibile invece di milioni come accade nei sensori attuali) migliorerebbe le capacità di acquisizione in maniera significativa [22]. In un ambito simile, la ricerca di Candès [13] si è concentrata anche sulla messa a punto di strumenti avanzati per la conversione "Analog to information (A/I)" di segnali a banda larga, per alleviare la pressione sulle tecniche convenzionali di conversione A/D, che sono correntemente limitate da una frequenza di circa 1GHz.

2.3 L'utilizzo del CS in Risonanza Magnetica

Come si è ampiamente discusso, il CS è una tecnica innovativa che ha permesso dei notevoli passi avanti in tutti i campi di elaborazione del segnale. In particolare, l'applicazione del CS alla risonanza magnetica permetterebbe di ridurre i tempi di scansione, con vantaggi diretti sia di tipo economico (per le strutture sanitarie), sia per il paziente che vedrebbe ridotto il tempo necessario per sottoporsi ad un esame. L'MRI soddisfa due requisiti fondamentali per un'applicazione di successo del CS:

- Le immagini medicali sono comprimibili naturalmente attraverso opportune trasformate;

2.3 L'utilizzo del CS in Risonanza Magnetica

- Gli scanner MRI acquisiscono campioni codificati nel dominio della frequenza.

Questo paragrafo mira a spiegare la perfetta compatibilità fra la teoria del CS e il campo della risonanza magnetica, nonché a fornire esempi di promettenti applicazioni in via di sviluppo.

2.3.1 Il dominio Wavelet

Come già discusso nel paragrafo riguardante il CS, la strategia di sparsificazione di un'immagine inizia con l'applicazione di una trasformata opportuna, che sia in grado di mappare il contenuto dell'immagine in un vettore di coefficienti sparsi; poi si codifica il vettore così ottenuto approssimando i coefficienti maggiori e scartando i più piccoli. Un esempio di trasformata di sparsificazione è la trasformata Wavelet discreta (DWT, Discrete Wavelet Transform), già utilizzata per la codifica JPEG-2000 [23], che utilizza le wavelet per trasformare un'immagine densa in una sparsa.

Consideriamo ad esempio l'immagine in Fig. 2.6(a) e la sua trasformata wavelet in Fig. 2.6(b). Anche se quasi tutti i pixel dell'immagine hanno valori diversi da zero, la maggior parte dei coefficienti wavelet sono piccoli: gran parte dell'informazione è infatti raccolta dai coefficienti più grandi, che sono relativamente pochi. La Fig. 2.6(c) è il risultato della compressione effettuata prendendo solo il 5% dei coefficienti originali e ponendo gli altri a zero (Fig. 2.6(d)). Come si vede la perdita di informazione è a malapena percettibile.

Una wavelet è definita come un'oscillazione simile ad un'onda che ad una fase di crescita fa seguire una fase di discesa. Combinando attraverso convoluzione tali funzioni con il segnale da acquisire si estrae informazione sulla localizzazione spaziale delle frequenze associate a tali funzioni. Questa proprietà della wavelet è fondamentale, poichè permette di ottenere rappresentazioni tempo-frequenza o spazio-frequenza dei segnali. Prendiamo per esempio una canzone qualsiasi e scegliamone una nota a caso: convolvendo una funzione wavelet con frequenza pari a quella della nota scelta e durata di un trentaduesimo di nota, possiamo risalire a quando nella canzone viene suonata quella particolare nota. Matematicamente, la wavelet entrerà in risonanza con il segnale sconosciuto quando esso contiene informazioni di frequenza simile, allo

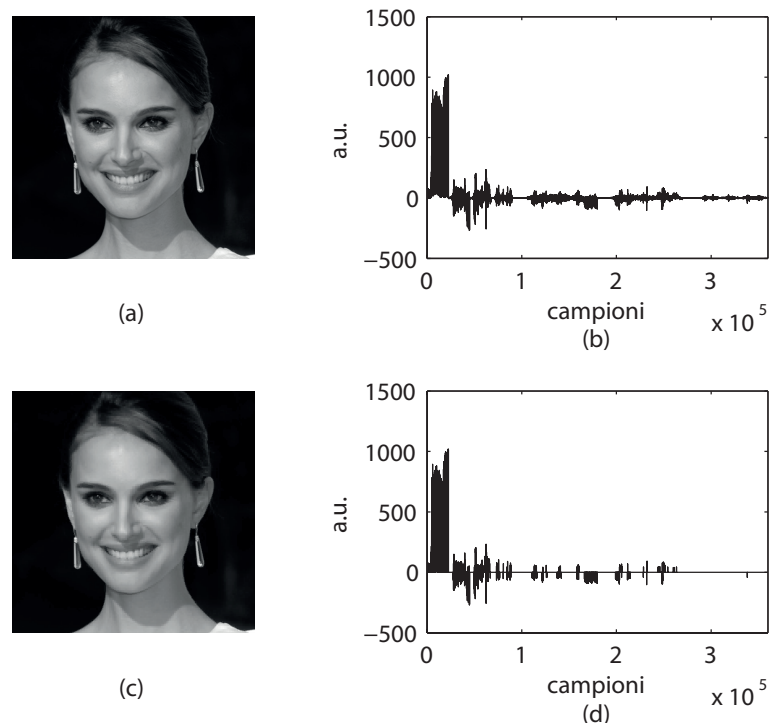


Figura 2.6: (a) Immagine originale da 0.36Mpixel e (b) i coefficienti della sua trasformata wavelet. Pochi coefficienti raccolgono la maggior parte dell'informazione. (c) Immagine ottenuta conservando il 5 % dei coefficienti wavelet, cioè i 18.000 valori maggiori come visibile in (d).

stesso modo in cui un diapason risuona con le onde sonore alla sua frequenza specifica. Famiglie di wavelet, quindi, possono essere utilizzate per algoritmi di compressione/decompressione per ricostruire informazioni con perdite minime.

Le trasformate wavelet possono essere considerate come forme di rappresentazione tempo-frequenza per segnali analogici che utilizzano banchi di filtri a tempo discreto. Questi banchi di filtri sono chiamati funzione wavelet $\varpi(t)$ (o funzione madre) e funzione di scalamento $\phi(t)$ (o funzione padre). Una combinazione di queste due funzioni (la funzione madre scalata in base alla funzione padre) crea la famiglia di filtri utilizzata per identificare i coefficienti wavelet dell'immagine d'interesse. Tale trasformata può essere implementata mediante una serie di filtri passa-alto e passa-basso, effettuando ad ogni passo una decimazione. Questo porta ad una struttura ad albero dei coefficienti wa-

2.3 L'utilizzo del CS in Risonanza Magnetica

velet di un'immagine, in cui ogni coefficiente fa da padre a quattro coefficienti figli (da qui il nome di *quadtree*). Una proprietà molto importante fa sì che se un coefficiente è trascurabile, tutti i suoi coefficienti figli (quindi l'intero sotto-albero da lui generato) è trascurabile. In questo modo, sfruttando tale proprietà (approfondita in seguito), è facile identificare quali siano i coefficienti contenenti l'informazione e quali siano invece trascurabili.

La trasformata wavelet può essere rappresentata in forma matriciale come [24], [25]

$$x = \Psi y, \quad (2.21)$$

dove x è un vettore di dati reali $M \times 1$, Ψ è una matrice $M \times M$ di trasformazione dal dominio wavelet al dominio proprio dei dati e y rappresenta il vettore $M \times 1$ dei coefficienti wavelet. Tali coefficienti raccolti in y possono essere rappresentati mediante una struttura ad albero, come evidenziato in Fig. 2.7. I coefficienti al livello $s = 1$ corrispondono ai nodi radice e i coefficienti al livello maggiore $s = L$ corrispondono ai nodi foglia. Il blocco a $s = 0$ corrisponde ai coefficienti di scalamento, che catturano la rappresentazione al livello più grezzo dell'immagine. Ogni coefficiente wavelet ai livelli $1 \leq s \leq L - 1$ ha quattro coefficienti "figli" al livello $s + 1$. La relazione statistica fra padri e figli è la proprietà che viene sfruttata nel campo del CS.

Un coefficiente wavelet può assumere uno stato "basso" o uno stato "alto". Intuitivamente, se un coefficiente è relativamente piccolo, esso apparterrà allo stato "basso", mentre se è grande c'è una grande probabilità che appartenga allo stato "alto". La probabilità che uno stato fissato sia condizionato dallo stato del padre è rappresentabile attraverso una matrice di probabilità P di dimensioni 2×2 , dove $P(i, j)$ rappresenta la probabilità che una volta assegnato il coefficiente padre allo stato i i coefficienti figli siano allo stato j . È osservabile che $P(1, 1) = 1 - \varepsilon$ e $P(1, 2) = \varepsilon$, dove $0 < \varepsilon \ll 1$. Questo conferma che se il coefficiente padre è piccolo, i suoi figli saranno con elevata probabilità anch'essi piccoli.

2.3.2 Acquisizione

Come osservato nel paragrafo 2.1, i metodi di imaging MRI utilizzano diverse sequenze di acquisizione, e ognuna di esse campiona secondo un certo schema

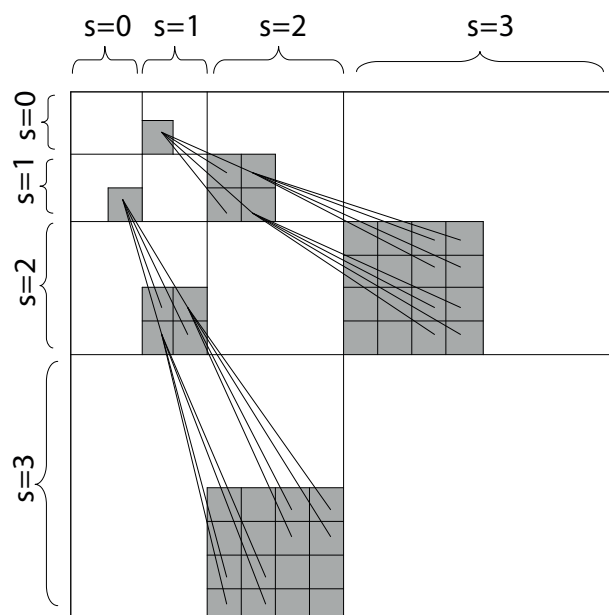


Figura 2.7: Struttura ad albero dei coefficienti wavelet: ogni coefficiente ha quattro figli.

una porzione o l'intero k -space. La traiettoria più utilizzata, come detto, è quella cartesiana. Con tali acquisizioni infatti il processo di ricostruzione è molto semplice: è necessaria infatti una semplice applicazione della trasformata di Fourier inversa. Inoltre, le traiettorie cartesiane sono robuste per quanto riguarda svariate sorgenti di errore o imperfezioni. Dato che la fisiologia umana pone limiti fondamentali alle performance di imaging (metodi basati su gradienti di alta intensità o su impulsi troppo rapidi possono infatti portare a stimolazioni nervose indesiderate nei pazienti [26]), la ricerca di metodi per ridurre la quantità di dati da raccogliere (favorendo così la velocità) è uno dei campi più attivi in ambito biomedico. Tali sforzi si basano sull'idea che i dati raccolti da una MRI sono ridondanti. Inoltre, la maggior parte delle immagini MR è sparsa in un dominio appositamente scelto. Gli angiogrammi, ad esempio, sono sparsi già nella loro rappresentazione nello spazio e possono essere resi ancora più sparsi attraverso differenze finite. Altre immagini più complesse, come quelle di MR al cervello, possono essere sparsificate in domini più sofisticati, come quello Wavelet o quello DCT (Discrete Cosine Transform), impiegati rispettivamente negli standard JPEG-2000 e JPEG. La rappresentazione sparsa non è limitata a segnali mono o bidimensionali: anche i video

2.3 L'utilizzo del CS in Risonanza Magnetica

possono essere compressi in sicurezza e in maniera molto più pesante, e tale fatto è dimostrato dal successo del codec MPEG. Anche le immagini di MR dinamiche sono molto comprimibili: la quasi periodicità delle risonanze dinamiche al cuore rende la trasformata di Fourier di tali immagini temporalmente sparsa.

In conclusione, la sparsità a seguito di una trasformazione delle immagini MR e la natura codificata delle acquisizioni MR rendono la risonanza magnetica un sistema perfetto a cui applicare la teoria del CS.

La NMR rappresenta un caso speciale di applicazione del CS in cui le combinazioni lineari campionate sono semplicemente coefficienti di Fourier. Quindi, in tali condizioni, il CS è in grado di effettuare ricostruzioni accurate partendo da un sottoinsieme di misure del k -space. La risonanza magnetica, inoltre, soddisfa le proprietà chiave elencate nel paragrafo 2.2 dedicato al CS. La condizione di sparsità è soddisfatta, come indicato all'inizio del paragrafo. L'incoerenza degli artefatti da sottocampionamento è anch'essa soddisfatta: tali artefatti saranno infatti simili a rumore (quindi incoerenti) nel dominio della trasformata sparsificante. Per finire, gli algoritmi di ricostruzione sono non lineari e promuovono sia la sparsità della rappresentazione dell'immagine che la consistenza della ricostruzione coi campioni acquisiti.

La condizione di campionamento incoerente può essere vista in questo caso come una scelta adeguata di un sottoinsieme del dominio delle frequenze che possa essere campionato con successo e che sia incoerente rispetto alla trasformata di sparsificazione. I risultati migliori nel CS in generale sono stati ottenuti campionando in maniera completamente random, così da avere campioni con la minima coerenza possibile, in accordo con l'equazione 2.12 e con la definizione 1 [7]. Ma tale metodo di campionamento, nonostante sia teoricamente ottimo, non è attuabile in pratica. Campionare infatti una porzione completamente casuale del k -space è infatti impossibile, in quanto le traiettorie di campionamento devono soddisfare vincoli fisiologici e meccanici. Di conseguenza, le traiettorie di campionamento che dovranno essere seguite sono rappresentate da linee smooth e da curve. Inoltre, gli schemi di campionamento devono essere robusti per quanto riguarda situazioni non-ideali, molto frequenti in pratica. Metodi di campionamento non cartesiani possono essere molto sensibili a imperfezioni di sistema. Inoltre, distribuzioni uniformi e

Teoria

casuali dei nostri campioni nella frequenza spaziale non prendono in considerazione la distribuzione di energia delle immagini MR nel k -space, che è tutt'altro che uniforme. Essa infatti è concentrata al centro del k -space e decresce velocemente allontanandosi da esso. Quindi, modelli realistici per l'applicazione del CS nel campo delle MRI dovrebbero avere un campionamento a densità variabile (con densità maggiore al centro del k -space), traiettorie irregolari che seguano almeno in parte le proprietà di incoerenza del campionamento casuale, permettendo allo stesso tempo una raccolta veloce dei dati.

Passiamo ora alla valutazione dell'incoerenza: consideriamo inizialmente di essere in casi in cui l'immagine è già sparsa nel dominio spaziale. Supponiamo di campionare un sottoinsieme S del k -space. Con \mathcal{F}_S indichiamo la trasformata di Fourier valutata alle frequenze del sottoinsieme S . Con \mathcal{F}_S^* indichiamo l'aggiunta, che può essere ottenuta mediante zero-filling seguito da antitrasformata di Fourier. Definiamo la *point spread function* (PSF) semplicemente come $\text{PSF}(i, j) = (\mathcal{F}_S^* \mathcal{F}_S)(i, j)$. Se non vi fosse sottocampionamento e si campionasse con metodo Cartesiano, la PSF sarebbe l'identità (e quindi non esisterebbero termini fuori dalla diagonale). Sottocampionare il k -space invece introduce termini non nulli fuori dalla diagonale. Un termine non nullo in (i, j) significa che la ricostruzione lineare del pixel i soffre un'interferenza dal pixel $j \neq i$. Quindi, la PSF misura la tendenza di ricostruzioni lineari zero-filled di perdere energia dal pixel originale a pixel differenti. Questa energia si mostra nell'immagine come zone sfocate o come artefatti di aliasing. Lo scopo del campionamento irregolare è di rendere tale perdita di energia quasi uniforme nell'immagine, così da rendere trascurabile la perdita massima. La coerenza, in questa situazione, viene definita come il valore maggiore di un pixel non-diagonale in una PSF normalizzata.

Nel caso di immagini sparse in domini diversi da quello temporale si introduce il concetto di TPSF (Transform PSF) [27], definita come $\text{TPSF}(i, j) = (\Psi^* \mathcal{F}_S^* \mathcal{F}_S \Psi)(i, j)$. Con questa notazione, la coerenza è formalmente misurata come $\max_{i \neq j} |\text{TPSF}(i, j)|$, cioè nuovamente la dimensione maggiore di qualsiasi elemento non-diagonale nella TPSF.

Schema di acquisizione

Nella MRI Cartesiana 2-D, il campionamento completo viene implementato come una serie di n_{pe} acquisizioni (chiamate *phase encodes*, codifiche di fase) lungo traiettorie molto semplici: linee parallele equidistanti fra loro. Questo modello comporta n_{fe} campioni del *k-space* per traiettoria (*frequency encodes*, codifica di frequenza), il che produce una matrice Cartesiana di $n_{pe} \times n_{fe}$ campioni totali. Per le ricostruzioni tridimensionali si ha una terza dimensione, n_{se} (*slice encodes*), che porta il numero di acquisizioni lineari a $n_{pe} \times n_{se}$, risultando in una griglia $n_{pe} \times n_{se} \times n_{fe}$. È importante notare come le codifiche di frequenza non siano un fattore limitante in termini di sottocampionamento e di tempo di scansione, mentre il numero di linee di acquisizione (codifiche di fase e slice) è a tutti gli effetti limitante. Questo fatto suggerisce di utilizzare un minor numero di linee per velocizzare la scansione. Tale ragionamento ha senso: si ha infatti completa libertà nella scelta delle linee da acquisire ed è il numero di esse ciò che determina il tempo di scansione finale. L'implementazione di tale modello richiederebbe solo minime modifiche alle sequenze di impulso: basterebbe infatti semplicemente saltare alcune acquisizioni. Dato che la maggior parte delle sequenze in uso in ambito clinico sono cartesiane, è molto conveniente implementare un'acquisizione CS in questa maniera. Scegliere di saltare alcune linee però comporta una coerenza maggiore rispetto a un sottocampionamento puramente casuale del *k-space*. Inoltre, come si può vedere dalla Fig. 2.8, nel campo dell'imaging bidimensionale si sfrutta solamente il sottocampionamento in una dimensione.

Allontanandosi dallo schema cartesiano si acquisisce una maggior flessibilità nello sviluppare traiettorie di campionamento con bassa coerenza. Gli schemi non cartesiani più popolari sono campionamento radiali o spirali, che mostrano PSF incoerenti.

Le immagini dinamiche, invece, si possono vedere come segnali bidimensionali con il tempo come dimensione aggiuntiva. Le MRI dinamiche vengono acquisite nel dominio della frequenza spaziale sul tempo. Invece di campionare il dominio *k-t* su linee ordinate e congruenti, basta ordinare tali linee casualmente e il *k-t* spazio verrà campionato casualmente [28]. Il risultato di questa acquisizione a linee ordinate casualmente porta un'incoerenza col

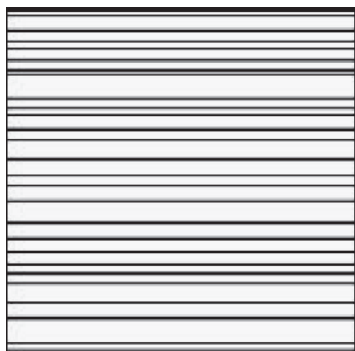


Figura 2.8: Come è evidente, utilizzando la traiettoria cartesiana 2D, si sfrutta solamente la sparsità monodimensionale, dato che solo una delle due dimensioni viene effettivamente sottocampionata.

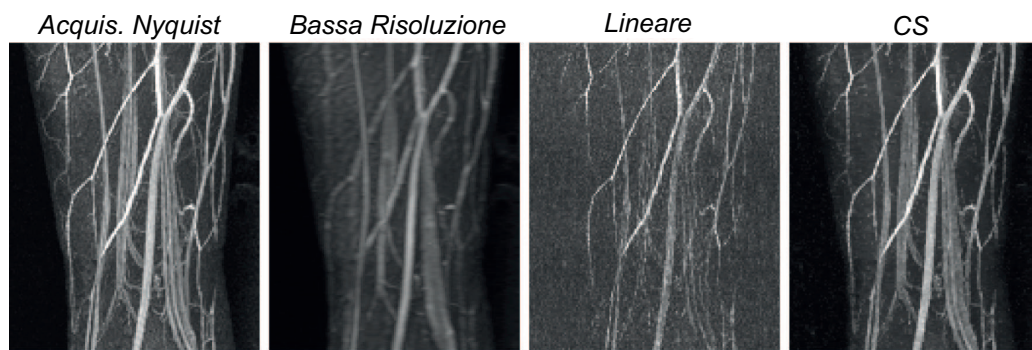
dominio spazio-frequenze temporali. Quindi, un ordinamento casuale delle linee è effettivamente un modo non dispendioso ed efficace di campionare incoerentemente dati dinamici.

2.3.3 Applicazioni del CS in MRI

Essendo il CS una teoria piuttosto generale, la sua applicazione a differenti ambiti porta a vincoli diversi dipendenti appunto dall'applicazione d'interesse, a seconda dell'hardware e delle considerazioni da fare sul paziente.

Una prima applicazione riguarda l'angiografia tridimensionale veloce. Come detto in precedenza, immagini provenienti da angiografie sono già piuttosto sparse nella rappresentazione spaziale [29] [30], e ancor di più alle differenze finite. In Fig. 2.9, vediamo l'applicazione del CS a un'angiografia tridimensionale Cartesiana con mezzo di contrasto. L'applicazione del CS a tale esame viene fatta scegliendo un sottoinsieme pseudocasuale del 10% delle traiettorie di acquisizione cartesiane originariamente utilizzate, con densità del k -space variabile. Tale scelta, come è evidente, porta a un sottocampionamento a bassa coerenza. Il CS è in grado di accelerare l'esame sensibilmente, dando la possibilità di avere una miglior risoluzione temporale o alternativamente una miglior risoluzione spaziale senza intaccare il tempo di scansione. Nel campo dell'angiografia delle coronarie, i raggi X rappresentano il gold standard per valutare malattie delle coronarie. Tale metodo è, però, invasivo. Un'alternativa

2.3 L'utilizzo del CS in Risonanza Magnetica



Configurazione di campionamento Cartesiana 3D

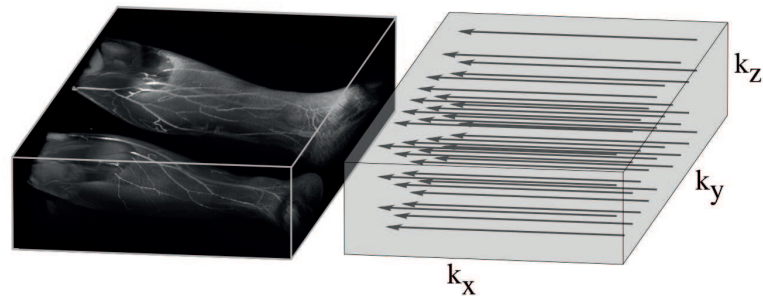


Figura 2.9: Angiografia 3D con mezzo di contrasto. In alto: Anche con un sottocampionamento pari a 10 volte l'originale, il CS può ricostruire la maggior parte dell'informazione dei vasi sanguigni rivelata dall'acquisizione alla frequenza di Nyquist. Il CS presenta molti meno artefatti rispetto alla ricostruzione lineare e una risoluzione molto migliore rispetto a un'acquisizione passa-basso del k -space. In basso: viene evidenziata la configurazione di campionamento casuale cartesiana tridimensionale.

Teoria

è rappresentata dalla tomografia computerizzata a raggi X multislice, ma tale tecnica pur non essendo invasiva richiede alte dosi di radiazioni ionizzanti. La MRI sta emergendo come alternativa non invasiva e non ionizzante. Tale esame è ricco di complicazioni: le arterie infatti sono costantemente in movimento, e il movimento del cuore combinato al respiro rendono l'acquisizione di immagini ad alta definizione una sfida molto complessa. Vincolando la scansione ai momenti in cui viene trattenuto il respiro e sincronizzandoci col battito cardiaco, però, si ottengono vincoli di tempo troppo penalizzanti per ottenere immagini a risoluzione accettabile. L'utilizzo del CS può permettere l'acquisizione di slice che coprano l'intera dimensione del cuore nell'unico periodo in cui al paziente viene fatto trattenere il respiro [31]: in tale campo viene utilizzata una traiettoria a spirale. Grazie all'uso del CS, inoltre, vengono soppressi i tipici artefatti che compaiono nelle ricostruzioni lineari a partire da misure sottocampionate [12].

L'applicazione clinica più frequente della MRI riguarda gli esami cerebrali. La maggior parte di essi viene fatta utilizzando acquisizioni multislice bidimensionali cartesiane. Tali immagini, come accennato precedentemente in questo paragrafo, sono sparse nel dominio wavelet. Il CS potrebbe permettere in linea teorica di ridurre sensibilmente il tempo di acquisizione e allo stesso tempo di ottenere una ricostruzione più accurata. Tale applicazione è stata oggetto di questa tesi e nei prossimi capitoli verranno presentati materiali e metodi utilizzati per poter applicare con successo la teoria del CS all'acquisizione di immagini DSC-MRI.

Capitolo 3

Materiali e metodi

3.1 Algoritmo veloce per la minimizzazione TV-L1-L2

In questo capitolo verrà presentato un algoritmo per la ricostruzione dei segnali partendo da dati parziali nel dominio della frequenza [1]. L'algoritmo minimizza la somma di tre termini corrispondenti a:

1. Total Variation (TV);
2. Minimizzazione della norma ℓ_1 ;
3. Aderenza ai dati con l'ausilio del metodo dei minimi quadrati.

Viene utilizzato uno schema di minimizzazione ad alternanza a doppio ciclo, nel quale il ciclo principale effettua le FFT (o, in caso i dati siano nel dominio DCT, la trasformata discreta del coseno). Verranno presentate anche le analisi delle condizioni di ottimalità di tale metodo.

3.1.1 Introduzione del problema e trattazione matematica

Definiamo il nostro segnale sconosciuto di interesse come $\bar{u} \in \mathbb{R}^N$. Nel caso in cui l'oggetto sotto analisi non sia un segnale ma altresì un'immagine (bidimensionale) o una matrice a tre o più dimensioni, utilizzeremo come \bar{u} la sua versione vettorizzata, quindi monodimensionale. Nella maggior parte dei casi, il numero di caratteristiche importanti in un segnale è molto minore

Algoritmo veloce per la minimizzazione TV-L1-L2

rispetto alla sua risoluzione, e questo significa che \bar{u} è solitamente sparso o comprimibile in una base appropriata. Definiamo $\Psi = [\psi_1, \psi_2, \dots, \psi_N] \in \mathbb{C}^{N \times N}$ come una base ortogonale a \mathbb{C}^N . Sappiamo quindi che esiste un unico $\bar{x} \in \mathbb{C}^N$ tale che

$$\bar{u} = \sum_{i=1}^N \psi_i \bar{x}_i = \Psi \bar{x}. \quad (3.1)$$

Il vettore \bar{u} è K -sparso in Ψ se $\|\bar{x}\|_0$, cioè il numero di elementi diversi da zero in \bar{x} , è K . Inoltre \bar{u} è comprimibile se \bar{x} ha solo qualche elemento con un'ampiezza relativamente "grande": il caso su cui ci si focalizzerà è quello in cui è verificata $K \ll N$ o \bar{u} è altamente comprimibile. Come abbiamo già spiegato nel capitolo 2 relativo alla Teoria, lo scopo del CS è ricostruire un segnale sparso da un piccolo numero di proiezioni di esso in un dominio definito. Definiamo M come un numero intero che soddisfi $K < M \ll N$. Ricordando che $\Phi \in \mathbb{C}^{M \times N}$ è una matrice generale di acquisizione non-adattativa (cioè indipendente da \bar{u}), lo scopo del CS è quello di acquisire

$$b = \Phi \bar{u} = A \bar{x} \in \mathbb{C}^M, \quad A = \Phi \Psi \quad (3.2)$$

e poi ricostruire \bar{x} (e quindi \bar{u} secondo l'equazione 3.1) dal vettore delle proiezioni b . Perché il CS sia pratico, si deve avere una matrice A che assicuri che b contenga una quantità sufficiente di informazione rispetto a quella contenuta in \bar{x} , e inoltre deve essere stabilito un algoritmo di decodifica (ricostruzione) efficiente, che sia in grado di ricostruire \bar{x} da b . A tal proposito, facendo riferimento alla RIP (introdotta nel paragrafo 2.2.3), abbiamo osservato che se Ψ è ortogonale, la matrice $A = \Phi \Psi$ soddisfa la RIP se Φ è una matrice *iid* gaussiana o una matrice parziale di Fourier. Dato che l'equazione 3.2 è indeterminata, ha infinite soluzioni. Sapendo a priori che b è acquisito da un segnale altamente sparso, un approccio ragionevole per l'identificazione della soluzione migliore è quello di cercare la soluzione più sparsa, cioè

$$\min_x \|x\|_0 : Ax = b. \quad (3.3)$$

In questa maniera il processo di decodifica è in grado di ricostruire esattamente un segnale K -sparso utilizzando solo $K+1$ misure *iid* gaussiane. Tale approccio, però, è NP-difficile e non è pratico computazionalmente per tutti i problemi

reali. Un sostituto valido alla 3.3 è

$$\min_x \|x\|_1 : Ax = b. \quad (3.4)$$

È dimostrato che sotto certe condizioni iniziali (vedi [32]), 3.3 e 3.4 hanno soluzioni comuni. Per il decoder ℓ_1 in 3.4, il numero di misure sufficienti per un'esatta ricostruzione di un segnale K -sparso è $O(K \log(N/K))$ quando A è ancora una matrice gaussiana a valori *iid* (da [8]) e $O(K \log N)$ [7] nel caso in cui A sia una matrice di Fourier parziale, come nel caso della MRI. Tale approccio, inoltre, è facilmente risolvibile in programmazione lineare, e rappresenta quindi un'ottima soluzione sia perchè promuove la sparsità sia per il fatto di essere computazionalmente trattabile. Se \bar{x} fosse comprimibile ma non sparso, o nel caso in cui le misure siano contaminate da rumore, è necessario rilassare la minimizzazione. Un'adeguata rappresentazione nel caso di rumore gaussiano è

$$\min_x \|x\|_1 : \|Ax - b\|_2 \leq \sigma, \quad (3.5)$$

dove $\sigma > 0$ dipende dal livello di rumore. È dimostrato che la soluzione della 3.5 è stabile e statisticamente può essere interpretata come \bar{x} come in equazione 3.1 [17].

Un problema collegato alla 3.5 è

$$\min_x \|x\|_1 + \lambda \|Ax - b\|_2^2, \quad (3.6)$$

con $\lambda > 0$. Dalla teoria dell'ottimizzazione, i problemi 3.5 e 3.6 sono equivalenti, nel senso che risolvere uno dei due determina il parametro nell'altro in maniera tale che entrambi i metodi forniscano la stessa soluzione.

Da questo punto in poi assumeremo che \bar{u} sia un'immagine bidimensionale in scala di grigi con N pixel. La sua corrispondente versione in frequenza parziale è data da

$$f_p = PT\bar{u} + \omega, \quad (3.7)$$

dove $\mathcal{T} \in \mathbb{C}^{N \times N}$ rappresenta una matrice di trasformazione specifica, $P \in \mathbb{R}^{p \times N}$ è una matrice di selezione contenente p righe della matrice identità di ordine N , e $\omega \in \mathbb{C}^p$ rappresenta rumore casuale. La matrice PT , nel CS, serve come matrice di acquisizione. Il modello 3.7 caratterizza la natura di molti

Algoritmo veloce per la minimizzazione TV-L1-L2

sistemi di acquisizione di dati, come nel caso della ricostruzione MRI. L'energia che viene registrata, proveniente dal sistema fisico, rappresenta infatti proprio la matrice $P\mathcal{T}\bar{u}$, o parte di essa. Nel caso della MRI, inoltre, \mathcal{T} non è altro che la trasformata di Fourier \mathcal{F} bidimensionale. L'algoritmo in esame prevede la ricostruzione di \bar{u} attraverso f_p utilizzando la teoria del CS.

Definiamo $\mathcal{F}_p = P\mathcal{F}$ e

$$\theta(u, f_p) = 1/2 \cdot \|\mathcal{F}_p u - f_p\|_2^2.$$

Nell'approccio utilizzato da Yang, \bar{u} viene ricostruito come la soluzione del seguente modello TVL1-L2:

$$\min_u \sum_i \|D_i u\|_2 + \tau \sum_i |\psi_i^\top u| + \lambda \theta(u, f_p), \quad (3.8)$$

dove \sum_i è presa su tutti i pixel, $\sum_i \|D_i u\|_2$ rappresenta una discretizzazione del TV di u , $\sum_i |\psi_i^\top u|$ è la norma ℓ_1 della rappresentazione di u attraverso Ψ e $\tau, \lambda > 0$ sono scalari che bilanciano regolarizzazione e fedeltà ai dati. Dato che le immagini di risonanza magnetica contengono generalmente al loro interno informazioni su diversi tessuti (a diversa intensità) confinanti fra loro, l'utilizzo del TV nella regolarizzazione sfrutta la sparsità dell'immagine e mantiene inalterate le transizioni (i bordi) fra più tessuti. È noto inoltre che le immagini MR hanno rappresentazione sparsa sotto certe basi wavelet [27]: proprio per questo motivo è stata scelta come Ψ la base wavelet. Il modello 3.8 è già stato analizzato da molti studi [33] e [27] ed è stato inoltre verificato che esso ricostruisce immagini MR di alta qualità partendo da un basso numero di coefficienti di Fourier [34].

3.1.2 Algoritmo risolutivo

Yang *et al.* propongono in [1] un algoritmo risolutivo per il modello 3.8. La difficoltà principale nella risoluzione di tale modello è causata dalla non differenziabilità dei suoi primi due termini. L'approccio messo in atto da Yang è quello di approssimare tali termini utilizzando il metodo della penalità quadratica, affinato grazie al lavoro di Courant [35]. Questo approccio può anche essere visto come l'applicazione della tecnica semi-quadratica [36] ad un'approssimazione appropriata del modello 3.8 [36] [37]. Introducendo come

Algoritmo veloce per la minimizzazione TV-L1-L2

variabili ausiliarie $\mathbf{w} = [\mathbf{w}_1; \dots; \mathbf{w}_N]$, dove ogni $\mathbf{w}_i \in \mathbb{R}^2$, e $z \in \mathbb{R}^N$ e definendo $\|\cdot\| = \|\cdot\|_2$, il modello 3.8 diventa

$$\begin{aligned} \min_{\mathbf{w}, z, u} \quad & \sum_i \|\mathbf{w}_i\| + \tau \sum_i |z_i| + \lambda \theta(u, f_p), \\ \text{s.t.} \quad & \mathbf{w}_i = D_i u, z_i = \psi_i^\top u, \quad \forall i. \end{aligned} \quad (3.9)$$

Per rilassare i vincoli di uguaglianza e penalizzare la loro violazione mediante funzioni quadratiche, vengono introdotte

$$\begin{aligned} \phi_1(s, t) &= |s| + (\beta/2) \cdot |s - t|^2, \quad s, t \in \mathbb{R} \\ \phi_2(\mathbf{s}, \mathbf{t}) &= \|\mathbf{s}\| + (\beta/2) \cdot \|\mathbf{s} - \mathbf{t}\|^2, \quad \mathbf{s}, \mathbf{t} \in \mathbb{R} \end{aligned}$$

per un dato $\beta > 0$. Il problema 3.9 può essere approssimato quindi da

$$\min_i \sum_i \phi_2(\mathbf{w}_i, D_i u) + \tau \sum_i \phi_1(z_i, \psi_i^\top u) + \lambda \theta(u, f_p). \quad (3.10)$$

Per semplicità, viene utilizzato lo stesso peso β per entrambi i termini $\|\mathbf{w}_i - D_i u\|^2$ e $|z_i - \psi_i^\top u|^2$. Per un u fissato, la minimizzazione rispettivamente a \mathbf{w} e z può essere effettuata in parallelo, dato che tutti i \mathbf{w}_i e z_i sono separati l'uno dall'altro nel modello 3.10. Quindi, l'algoritmo si sviluppa attraverso due passi. Nel primo, a u fissato, il minimizzatore z_i è ottenuto da un restringimento monodimensionale

$$z_i = s_1(\psi_i^\top u), \quad \forall i, \quad (3.11)$$

dove $s_1(t)$ minimizza $\phi_1(s, t)$ per un fissato t , dato da

$$s_1(t) \triangleq \max |t| - 1/\beta, 0 \cdot \text{sgn}(t), \quad t \in \mathbb{R}, \quad (3.12)$$

mentre il minimizzatore \mathbf{w}_i è dato dal restringimento bidimensionale

$$\mathbf{w}_i = s_2(D_i u), \quad \forall i, \quad (3.13)$$

dove $s_2(\mathbf{t})$ minimizza $\phi_2(\mathbf{s}, \mathbf{t})$ per un fissato \mathbf{t} , dato da

$$s_2(\mathbf{t}) \triangleq \max \|\mathbf{t}\| - 1/\beta, 0 \cdot \mathbf{t}/\|\mathbf{t}\|, \quad \mathbf{t} \in \mathbb{R}^2, \quad (3.14)$$

dove viene assunto che $0 \cdot (0/0) = 0$. Il costo computazionale per 3.11 e 3.13 è lineare in termini di N .

Algoritmo veloce per la minimizzazione TV-L1-L2

Il secondo passo dell'algoritmo consiste, per fissati \mathbf{w} e z , nella minimizzazione di 3.10 rispetto a u . Tale processo diventa un problema ai minimi quadrati

$$\min_u \sum_i (\|\mathbf{w}_i - D_i u\|^2 + \tau |z_i - \psi_i^\top u|^2) + 2\eta\theta(u, f_p), \quad (3.15)$$

dove $\eta = \lambda/\beta$.

Definiamo $w_j \triangleq (\mathbf{w}_1(j); \dots; \mathbf{w}_N(j)), j = 1, 2$. Dalla definizione 3 di $D^{(1)}$ e $D^{(2)}$ in appendice A segue che

$$\sum_i \|\mathbf{w}_i - D_i u\|^2 = \|w_1 - D^{(1)}u\|^2 + \|w_2 - D^{(2)}u\|^2.$$

A causa dell'ortogonalità di Ψ , le equazioni normali di 3.15 possono essere scritte come

$$Lu = r, \quad (3.16)$$

dove

$$\begin{aligned} L &= (D^{(1)})^\top D^{(1)} + (D^{(2)})^\top D^{(2)} + \tau I + \eta \mathcal{F}_p^\top \mathcal{F} \\ r &= (D^{(1)})^\top w_1 + (D^{(2)})^\top w_2 + \tau \Psi z + \eta \mathcal{F}_p^\top f_p. \end{aligned}$$

Dato che $D^{(1)}$ e $D^{(2)}$ sono operatori alle differenze finite, sotto le condizioni al contorno periodiche per u sono matrici circolanti e possono essere diagonalizzati dalla trasformata di Fourier \mathcal{F} . Viene definito quindi $\hat{D}^{(j)} = \mathcal{F}D^{(j)}\mathcal{F}^\top$, $j = 1, 2$, che è diagonale. Moltiplicando per \mathcal{F} entrambi i membri di 3.16 otteniamo

$$\hat{L}\mathcal{F}(u) = \hat{r}, \quad (3.17)$$

dove

$$\hat{L} = (\hat{D}^{(1)})^\top \hat{D}^{(1)} + (\hat{D}^{(2)})^\top \hat{D}^{(2)} + \tau I + \eta P^\top P$$

è diagonale dato che $P^\top P$ è diagonale, e

$$\hat{r} = (\hat{D}^{(1)})^\top \mathcal{F}(w_1) + (\hat{D}^{(2)})^\top \mathcal{F}(w_2) + \tau \mathcal{F}(\Psi z) + \eta P^\top f_p.$$

Quindi, risolvere 3.17 è possibile con un costo computazionale non elevato, il che significa che 3.16 può essere risolto semplicemente per w_1, w_2 e z fissati come segue:

1. Applicare la FFT a w_1, w_2 e z ;

Algoritmo veloce per la minimizzazione TV-L1-L2

2. Risolvere 3.17 per ottenere $\mathcal{F}(u)$, dove è assunto che le costanti $\hat{D}^{(1)}$ e $\hat{D}^{(2)}$ sono precalcolate;
3. Applicare la FFT inversa a $\mathcal{F}(u)$ per ottenere la soluzione u alla 3.16.

Dato che minimizzare la funzione obiettivo in 3.10 rispetto ad ogni variabile è computazionalmente economico, viene proposto il seguente algoritmo di minimizzazione alternata:

Algoritmo 1. *Inserisco $P, f_p; \tau, \lambda, \beta > 0$. Inizializzo $u = u^0$.*
while “non converge”, **do**
 1) *Dato u , calcola z e \mathbf{w} attraverso le 3.11 e 3.13.*
 2) *Dati z e \mathbf{w} , calcola u risolvendo la 3.16.*
end do

3.1.3 Criteri di stop e condizioni di ottimalità

I criteri di stop per l'algoritmo 1 vengono specificati derivando le condizioni di ottimalità da 3.8 e 3.10.

Nelle definizioni 4, 5 e 6 in Appendice A vengono riportati i significati di sottodifferenziale di $|t|, t \in \mathbb{R}$ e $\|v\|_1, v \in \mathbb{R}^N$ e due proposizioni che saranno necessarie per le dimostrazioni di questo paragrafo.

Dato che la funzione obiettivo è convessa, una tripletta (z, \mathbf{w}, u) fa parte delle soluzioni di 3.10 se e solo se il sottodifferenziale della funzione obiettivo in (z, \mathbf{w}, u) contiene l'origine. In riferimento alle proposizioni 5 e 6 in appendice A con A e B matrici di dimensioni appropriate, (z, \mathbf{w}, u) risolve 3.10 se e solo se

$$\begin{cases} \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|} + \beta(\mathbf{w}_i - D_i u) = 0, & i \in I_1 \triangleq \{i : \mathbf{w}_i \neq 0\}, \\ \|D_i u\| \leq 1/\beta, & i \in I_2 \triangleq \{i : \mathbf{w}_i = 0\} \\ \text{sgn}(z_i) + \beta(z_i - \psi_i^\top u) = 0, & i \in I_3 \triangleq \{i : z_i \neq 0\}, \\ |\psi_i^\top u| \leq 1/\beta, & i \in I_4 \triangleq \{i : z_i = 0\}, \end{cases}$$

$$D^\top(Du - w) + \tau(u - \Psi z) + \eta \nabla \theta(u, f_p) = 0,$$

Algoritmo veloce per la minimizzazione TV-L1-L2

dove $\nabla\theta(u, f_p) \triangleq \mathcal{F}_p^\top (\mathcal{F}_p u - f_p)$, $D = (D^{(1)}, D^{(2)})$ e $w = (w_1, w_2)$. Eliminando w e z dalle equazioni appena formulate, si ottiene

$$\sum_{i \in I_1} D_i^\top \frac{D_i u}{\|D_i u\|} + \sum_{i \in I_2} D_i^\top h_i + \tau \sum_{i \in I_3} \text{sgn}(\psi_i^\top u) \psi_i + \tau \sum_{i \in I_4} g_i \psi_i + \lambda \nabla\theta(u, f_p) = 0, \quad (3.18)$$

dove $h_i \triangleq \beta D_i u$ soddisfa $\|h_i\| \leq 1$ e $g_i \triangleq \beta \psi_i^\top u$ soddisfa $|g_i| \leq 1$.

In seguito si dimostra che la 3.18 è un'approssimazione dell'ottimalità di 3.8. Definiamo u^* come una qualsiasi soluzione di 3.8,

$$\begin{aligned} I_1^* &\triangleq \{i : D_i u^* \neq 0\}, & I_2^* &\triangleq \{i : D_i u^* = 0\}, \\ I_3^* &\triangleq \{i : \psi_i^\top u^* \neq 0\}, & I_4^* &\triangleq \{i : \psi_i^\top u^* = 0\}, \end{aligned}$$

Richiamando le definizioni 5 e 6 in appendice A, esistono $\{h_i^* \in \mathbb{R}^2 : \|h_i^*\| \leq 1, i \in I_2^*\}$ e $\{g_i^* \in \mathbb{R} : |g_i^*| \leq 1, i \in I_4^*\}$ tali che

$$\sum_{i \in I_1^*} D_i^\top \frac{D_i u^*}{\|D_i u^*\|} + \sum_{i \in I_2^*} D_i^\top h_i^* + \tau \sum_{i \in I_3^*} \text{sgn}(\psi_i^\top u^*) \psi_i + \tau \sum_{i \in I_4^*} g_i^* \psi_i + \lambda \nabla\theta(u^*, f_p) = 0. \quad (3.19)$$

L'equazione 3.18 si differenzia dalla 3.19 solo nei set di indici utilizzati. Al crescere di β , I_1 e I_3 convergeranno a I_1^* e I_3^* rispettivamente. Il criterio di stop dell'Algoritmo 1 è basato, come detto, sulle condizioni di ottimalità di 3.10. Definiamo

$$\begin{cases} r_1(i) = (\mathbf{w}_i / \|\mathbf{w}_i\|) / \beta + \mathbf{w}_i - D_i u & i \in I_1, \\ r_2(i) = \|D_i u\| - 1 / \beta & i \in I_2, \\ r_3(i) = \text{sgn}(z_i) / \beta + z_i - \psi_i^\top u & i \in I_3, \\ r_4(i) = |\psi_i^\top u| - 1 / \beta & i \in I_4, \end{cases}$$

$$r_5 = D^\top (Du - w) + \tau(u - \Psi z) + \eta \nabla\theta(u, f_p).$$

Definiamo poi

$$\begin{aligned} \text{res}_1 &\triangleq \max\{\|r_1(i)\|, i \in I_1; r_2(j), j \in I_2\}, \\ \text{res}_2 &\triangleq \max\{|r_3(i)|, i \in I_3; r_4(j), j \in I_4\}. \end{aligned}$$

L'Algoritmo 1 termina quando si verifica

$$\text{Res} \triangleq \max\{\text{res}_1, \text{res}_2, \|r_5\|_\infty\} \leq \epsilon, \quad (3.20)$$

3.2 Miglioramento del metodo

dove $\epsilon > 0$ è la tolleranza. In un'implementazione pratica dell'algoritmo 1, viene assegnato a β un valore inizialmente piccolo che viene poi gradualmente aumentato. Per un β fissato, viene applicato l'algoritmo 1 a 3.10 finchè la condizione 3.20 è soddisfatta. La soluzione ottenuta è quindi utilizzata per far ripartire l'algoritmo 1 corrispondente al valore di β successivo. L'algoritmo che è stato quindi utilizzato è il seguente:

Algoritmo 2. *Inserisco P, f_p ; Parametri del modello $\tau, \lambda > 0$ e tolleranza $\epsilon > 0$. Inizializzo $\beta = 2^5$ e $u = u^0$.*

while $\beta \leq 2^{10}$, **do**

1) *Partendo da u , risolvi 3.10 utilizzando l'algoritmo 1.*

2) *Aggiorna $u \leftarrow u_\beta, \beta \leftarrow 2\beta$.*

end do

Per ottenere performance migliori si sarebbe potuto certamente incrementare β e stoppare le iterazioni interne in una maniera maggiormente adattativa, ma Yang [1] ha deciso di mantenere tale algoritmo. Il valore finale di β pari a 2^{10} è stato scelto perchè, grazie a prove empiriche, è abbastanza grande da rendere 3.10 un'approssimazione ottima di 3.8, nel senso che un valore di β maggiore non avrebbe fornito una soluzione evidentemente più accurata.

3.2 Miglioramento del metodo

L'idea che sta alla base del nostro metodo nasce dalla natura della tecnica in questione: la DSC-MRI infatti, essendo composta da una sequenza di immagini dinamiche acquisite per tempi successivi molto ravvicinati, contiene in tale rappresentazione più informazione rispetto ad un insieme di immagini MRI tradizionali statiche senza correlazione temporale. Ciò che interessa non è l'informazione in sè, ma la ridondanza che deriva dal fatto di avere immagini molto simili fra loro. Il presupposto ideale è quello di considerare quasi costanti nel tempo le intensità dei voxel dove il contrasto ha un apporto minore, mentre si considerano tempo-varianti quelle dei voxel in cui l'agente paramagnetico ha effetti maggiori: in particolare, si ha un abbattimento del segnale in tali voxel. Secondo questo assunto, che si può considerare con buona approssimazione

Algoritmo veloce per la minimizzazione TV-L1-L2

vero, si ha a disposizione una conoscenza a priori (*prior*) sui valori del segnale nelle regioni “statiche”.

L'obiettivo del nostro algoritmo è quello di sfruttare questa ridondanza di informazione per aumentare il fattore di sottocampionamento senza inficiare la qualità di ricostruzione rispetto ad un sottocampionamento con CS tradizionale. Questo consentirebbe nell'ambito delle sequenze dinamiche di ridurre sensibilmente i tempi di acquisizione per ogni singola fetta, e di ottenere quindi un numero maggiore di scansioni nell'intervallo di tempo limitato che caratterizza il passaggio del mezzo di contrasto nei tessuti. Parimenti si può sfruttare il vantaggio in termini qualitativi acquisendo un *k-space* più esteso, che significherebbe ricostruire immagini con risoluzione maggiore. Inoltre questo lavoro ha come ulteriore scopo quello di verificare se l'aumento della risoluzione temporale, a scapito anche della fedeltà nella ricostruzione, possa migliorare la stima dei parametri emodinamici. Di seguito verrà illustrato nel dettaglio il modello matematico sviluppato e la sua introduzione nell'algoritmo 2 già illustrato in precedenza.

3.2.1 Aggiunta dell'informazione spaziale a priori

L'idea di utilizzare un prior nasce dalla necessità di integrare un'informazione di per sé carente, dovuta al campionamento molto rado del *k-space*. Se con i normali algoritmi usati nel CS si possono avere buone ricostruzioni con 1/2 o 1/4 dei campioni, con questo metodo si ha la possibilità di scendere a valori significativamente inferiori. L'acquisizione in risonanza magnetica corrisponde ad una griglia di coefficienti di Fourier (*k-space*), pertanto la ricostruzione di un'immagine avviene tramite trasformata di Fourier inversa dei campioni acquisiti. Aumentando di molto il fattore di sottocampionamento si ottiene un *k-space* troppo sparso per poter ricostruire correttamente l'immagine: l'aggiunta del prior mira proprio a colmare questo deficit di informazione. In base a quanto detto nel paragrafo precedente, il nostro prior risulta essere nel dominio spaziale e come tale deve essere utilizzato. La trasformata di Fourier è un operatore globale i cui singoli coefficienti non conservano informazioni sulla localizzazione spaziale dei pixel: ciò implica che effettuando la trasformata di Fourier di una parte dei pixel dell'immagine i coefficienti risultanti non

possono essere collegati direttamente in alcun modo ai singoli pixel considerati. Questo è fondamentale poiché la nostra informazione a priori è incentrata su una parte dei pixel dell'immagine di riferimento (corrispondenti ai voxel con intensità costante dove non arriva il mezzo di contrasto) e non sull'immagine globale. Si è quindi pensato di aggiungere l'informazione spaziale come un ulteriore vincolo di *data fitting* nella funzione obiettivo da minimizzare: ciò costringe l'ottimizzatore ad avvicinarsi di più ai dati che abbiamo a disposizione come prior, sebbene la scarsa numerosità del *k-space* tenda ad introdurre un errore consistente.

Maschera per il prior spaziale

Di essenziale importanza per lo sviluppo del metodo è stata la definizione corretta di una maschera per poter utilizzare correttamente il prior nel minimizzatore. Essendo questi test basati esclusivamente su simulazioni di immagini, la maschera è stata presa come nota a priori e fissata. Come già discusso in precedenza, esiste corrispondenza biunivoca fra i punti appartenenti alla maschera e i punti in cui vi è passaggio diretto di mezzo di contrasto. Essendo in condizioni altamente ideali, vi è assenza di fenomeni di leakage o di drift di luminosità. Tale scelta è stata effettuata tenendo a mente lo scopo ultimo di queste simulazioni: valutare il comportamento dell'algoritmo nella ricostruzione delle due zone distinte: quelle appartenenti al prior e quelle non appartenenti ad esso. Tale argomento verrà approfondito in sede di discussione, nel capitolo 5.

3.2.2 Modello matematico modificato

Approccio basato sulla Baseline

Per integrare l'informazione a priori nel modello già visto in 3.10, bisogna effettuare delle opportune modifiche alla funzione che esso minimizza. Tale integrazione è stata effettuata integrando nel modello l'informazione a priori rappresentata dalla baseline, cioè un'immagine di riferimento in condizione di stato stazionario, calcolata come la media delle prime 8 acquisizioni (come spiegato in 2.1.4). Data l'incompatibilità fra prior spaziale e gli altri termini

Algoritmo veloce per la minimizzazione TV-L1-L2

della minimizzazione (che si trovano nel dominio delle frequenze e in quello wavelet), si è deciso di utilizzare il dominio wavelet per rappresentare il prior spaziale. Per sfruttare la ridondanza nei dati spaziali, si introduce un nuovo termine da minimizzare in tale dominio: rappresentando con U l'immagine al passo corrente e con V la baseline (rappresentante l'informazione a priori, come spiegato nel paragrafo 2.1.4), si vuole che la differenza fra esse sia minima nei punti in cui non vi è passaggio rilevante di mezzo di contrasto. Matematicamente, si deve minimizzare

$$|\Psi^\top V_i - \Psi^\top U|_{\ell_1, i} \quad i \notin C,$$

dove C rappresenta l'insieme dei punti dove arriva il mezzo di contrasto (cioè quei punti supposti in stato non stazionario dove l'informazione a priori non deve essere utilizzata). Per evitare una discontinuità nei dati è stato necessario mediare nei punti interni all'insieme C l'informazione presa dalle ricostruzioni con le informazioni provenienti dalla baseline sui punti appartenenti all'insieme stesso. Come si può intuire, minimizzare il termine sopra introdotto corrisponde ad utilizzare le informazioni estratte dalla baseline in tutti i punti dell'insieme complementare a C . Il termine da noi sviluppato deve essere introdotto nelle equazioni 3.8 e 3.10: è necessario quindi trattarlo allo stesso modo del termine ℓ_1 considerato in 3.11. Definiamo quindi una nuova variabile ausiliaria $\xi \in \mathbb{R}^N$ che assuma questi valori:

$$\xi_i^b = \begin{cases} \Psi^\top V_i & i \notin C \\ \Psi^\top (\alpha U_i + (1 - \alpha) V_i) & i \in C \end{cases} \quad (3.21)$$

dove $0 < \alpha < 1$. Dovendo minimizzare il termine $\|\xi_i^b\|_1$, espandendolo ai minimi quadrati come in 3.15 si ottiene

$$\|\xi_i^b - \Psi^\top U\|^2. \quad (3.22)$$

La nostra funzione di aggiornamento di ξ_i^b è analoga alla 3.12.

Nello sviluppo 3.16, quindi, compaiono dei termini in più dovuti alla nostra aggiunta di informazione a priori. Tali equazioni si modificano quindi in

$$\begin{aligned} L &= (D^{(1)})^\top D^{(1)} + (D^{(2)})^\top D^{(2)} + \tau I + \eta \mathcal{F}_p^\top \mathcal{F} + \chi I, \\ r &= (D^{(1)})^\top w_1 + (D^{(2)})^\top w_2 + \tau \Psi z + \eta \mathcal{F}_p^\top f_p + \chi \Psi \xi, \end{aligned}$$

avendo definito χ il coefficiente che pesa il nostro termine rispetto alla minimizzazione di Yang.

Tali termini, secondo l'equazione 3.17, diventano quindi

$$\begin{aligned}\hat{L} &= (\hat{D}^{(1)})^\top \hat{D}^{(1)} + (\hat{D}^{(2)})^\top \hat{D}^{(2)} + \tau I + \eta P^\top P + \chi I, \\ \hat{r} &= (\hat{D}^{(1)})^\top \mathcal{F}(w_1) + (\hat{D}^{(2)})^\top \mathcal{F}(w_2) + \tau \mathcal{F}(\Psi z) + \eta P^\top f_p + \chi \mathcal{F}(\Psi \xi).\end{aligned}$$

Si può notare come \hat{L} mantenga la proprietà di diagonalità, in quanto il nostro termine aggiuntivo non è altro che un'identità pesata da scalare, quindi diagonale. Tale approccio verrà identificato in seguito con l'acronimo *RecPFmod*.

Approccio alternativo con pesatura aggiuntiva

Un approccio alternativo (che denomineremo *RecPFalt*) consiste nel considerare non solo la baseline (V) nella pesatura al momento della costruzione di ξ , ma anche l'informazione ottenuta dalla ricostruzione al passo precedente, definita come U_{-1} . Tale metodo deriva dall'idea teorica che, con i dati reali, sarebbe necessario aggiornare la maschera durante le acquisizioni sottocampionate, per ottimizzare l'informazione spaziale a priori e identificarne eventuali cambiamenti temporali. Per ovviare a questo problema, abbiamo introdotto la pesatura sulla ricostruzione al passo precedente, in modo tale da avere la possibilità di aggiornare la nostra ricostruzione in caso di eventi non previsti in fase di definizione della maschera. Tale approccio, comunque, sarà più significativo nell'utilizzo con i dati reali [38], visto il modo in cui è stata definita la maschera in questa sede. La nuova definizione di ξ diventa quindi la seguente:

$$\xi_i^a = \begin{cases} \Psi^\top [\gamma U_{-1,i} + (1 - \gamma)V_i] = \Psi^\top \Upsilon & i \notin C \\ \Psi^\top [\beta U_i + (1 - \beta)\Upsilon] & i \in C \end{cases} \quad (3.23)$$

dove $0 < \beta < 1$ e $0 < \gamma < 1$. I passaggi matematici presentati nel paragrafo precedente rimangono validi anche per questo approccio.

3.2.3 Aggiornamento dei criteri di stop

I nuovi criteri di stop da inserire a causa dell'aggiunta del nostro termine sono analoghi ed omogenei a quelli già visti in 3.1.3 riguardanti il termine

Algoritmo veloce per la minimizzazione TV-L1-L2

di norma ℓ_1 , cioè il termine in ϕ_1 in 3.10 e 3.12. Per prima cosa, si può affermare che anche i nuovi criteri di stop che andremo a definire soddisfano le condizioni di ottimalità. Questo si può affermare a priori per quanto affermato in apertura di paragrafo: essendo i nuovi criteri di stop omogenei ed analoghi a quelli del termine di norma ℓ_1 , essi soddisferanno per forza le condizioni già trattate nella sezione 3.1.3. Una volta chiarito questo aspetto, si può proseguire all'implementazione dei due nuovi criteri di stop e alla modifica del residuo globale che nel modello 3.1.3 è denominato r_5 ma che nella nostra trattazione diventerà, dopo le opportune modifiche, r_7 . Definiamo quindi in analogia con l'equazione 3.11

$$y_i = s_3(\psi_i^\top V), \quad \forall i, \quad (3.24)$$

dove $s_3(t)$ minimizza $\phi_1(s, t)$ per un fissato t (qui rappresentato da $\psi^\top V$ che altro non è il prior rappresentato dalla baseline nel dominio wavelet), dato dall'equazione 3.12. Quindi, si ha che

$$s_3(\psi_i^\top V) \triangleq \max\{|\psi_i^\top V| - 1/\beta, 0\} \cdot \text{sgn}(\psi_i^\top V), \quad \psi_i^\top V \in \mathbb{R}. \quad (3.25)$$

Definendo poi in analogia con I_3 e I_4

$$I_5 \triangleq \{i : \psi_i^\top V \neq 0\}, \quad I_6 \triangleq \{i : \psi_i^\top V = 0\},$$

ottengo i nuovi residui, espressi come

$$\begin{cases} r_1(i) = (\mathbf{w}_i / \|\mathbf{w}_i\|) / \beta + \mathbf{w}_i - D_i u & i \in I_1, \\ r_2(i) = \|D_i u\| - 1/\beta & i \in I_2, \\ r_3(i) = \text{sgn}(z_i) / \beta + z_i - \psi_i^\top u & i \in I_3, \\ r_4(i) = |\psi_i^\top u| - 1/\beta & i \in I_4, \\ r_5(i) = \text{sgn}(y_i) / \beta + y_i - \psi_i^\top V & i \in I_5, \\ r_6(i) = |\psi_i^\top V| - 1/\beta & i \in I_6, \end{cases}$$

$$r_7 = D^\top (Du - w) + \tau(u - \Psi z) + \iota(V - \Psi y) + \eta \nabla \theta(u, f_p).$$

dove con ι è stata indicata la pesatura del nostro termine aggiuntivo. Definiamo poi il nuovo

$$\text{res}_3 \triangleq \max\{|r_5(i)|, i \in I_5; r_6(j), j \in I_6\}.$$

L'Algoritmo 1 ora termina quando

$$\text{Res} \triangleq \max\{\text{res}_1, \text{res}_2, \text{res}_3, \|r_7\|_\infty\} \leq \epsilon \quad (3.26)$$

è verificato, dove $\epsilon > 0$ è la tolleranza. Rimane valido, in definitiva, l'algoritmo 2, con i nuovi criteri di stop per il ciclo interno (algoritmo 1) appena definiti in 3.26.

3.3 Dati a disposizione

Per poter studiare il comportamento dell'algoritmo su sequenze dinamiche, si è fatto ricorso ad una simulazione MATLAB col noto "Shepp-Logan phantom" [39], comunemente utilizzato nelle simulazioni in ambito MRI, al quale sono stati sovrapposti del rumore bianco e alcune aree di intensità variabile definite arbitrariamente. Tali aree imitano il cambiamento di intensità prodotto dal passaggio del tracciante in determinate zone dell'encefalo. La varianza del rumore è stata scelta in modo da ottenere un livello di SNR (rapporto segnale-rumore) di 15 dB. La simulazione del passaggio del mezzo di contrasto è stata effettuata sovrapponendo alle aree di interesse una funzione gamma mono-variata con valore massimo fissato: tale funzione ricalca l'andamento del segnale al momento del passaggio del tracciante. Le dimensioni scelte per le immagini simulate sono di 256x256 pixel. Vengono considerati 51 tempi virtuali, e i primi 8 di essi vengono utilizzati per creare l'informazione spaziale nota a priori (baseline, come citato nel paragrafo 2.1.4) su cui si basa il metodo. L'esecuzione di un esame DSC-MRI infatti richiede che l'acquisizione inizi ben prima che venga iniettato l'agente di contrasto. Questo ci permette di avere delle acquisizioni non sottocampionate che possiamo sfruttare dopo averle mediate fra di loro per minimizzare il rumore.

Per la verifica del nostro metodo abbiamo considerato tre valori di accelerazione: uno basso (4), uno medio/alto (8) ed uno molto alto (16). L'accelerazione corrisponde al sottocampionamento del *k-space*: un'accelerazione pari a 4 corrisponde infatti ad un'acquisizione di un quarto dei campioni del *k-space*. Naturalmente, a valori alti di accelerazione corrisponde un maggior rumore di ricostruzione che porta, inevitabilmente, ad un maggiore errore. Per quanto riguarda la trattazione dei dati reali, si rimanda a [38].



(a) $t=1$

(b) $t=19$

Figura 3.1: Phantom: stato stazionario (a) e durante il picco di tracciante (b)

3.4 Metodo di campionamento del k -space

Nel Cap.2 sono presentate le traiettorie di acquisizione del k -space più comuni. Essendo questo un tema di fondamentale importanza ed allo stesso tempo molto vasto da affrontare, non rientra negli scopi di questa tesi affrontare nel dettaglio tale problematica. Non verrà quindi presentata alcuna traiettoria specifica per l'acquisizione DSC-MRI. Infatti, oltre ai problemi di ottimizzazione della traiettorie per avere una ricostruzione ottimale, bisogna sempre confrontarsi con i problemi pratici dell'implementazione fisica sullo scanner MRI. Perciò d'ora in poi si farà riferimento ad un solo metodo di acquisizione del k -space che, come già dimostrato teoricamente nel Cap.2, risulta teoricamente ottimo per ricostruire dati parziali con un errore minimo.

Si è già evidenziata l'importanza di acquisire con un alto grado di incoerenza rispetto alla base di sparsificazione: un campionamento totalmente random permette di massimizzare l'incoerenza ed ottenere solamente artefatti di ricostruzione assimilabili a rumore bianco. Inoltre, una caratteristica del k -space è quella di avere la maggior parte dell'energia del segnale concentrata alle basse frequenze. Per sfruttare appieno questa proprietà e massimizzare l'efficacia del CS, è quindi opportuno adottare uno schema di campionamento random con densità variabile. In questo modo, concentrando il campionamento alle basse frequenze, a parità di campioni acquisiti si estrae più informazione rispetto ad un campionamento random con densità costante su tutto il k -space. Si suppone inoltre che questo *pattern* di acquisizione sia ottenuto con una traiettoria di tipo cartesiano, cioè con un "rastrellamento" ordinato per linee

del k -space. Si tratta tuttavia di una situazione ideale che non tiene conto di eventuali limiti fisici della macchina per la messa a punto della traiettoria di acquisizione.

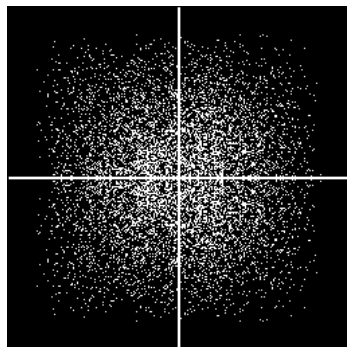


Figura 3.2: Campionamento random con densità variabile del k -space

Come si vede in Fig. 3.2 i campioni acquisiti (indicati in bianco) sono maggiormente concentrati intorno al centro, che corrisponde alle basse frequenze del k -space (rappresentazione traslata dello spettro di un segnale 2D). È da notare che la metà superiore originariamente non è stata campionata, in quanto si è sfruttata la proprietà di simmetria hermitiana dello spettro ($a_{ij} = a_{ji}^*$, dove con a^* si indica il complesso coniugato dell'elemento a) per riprodurre i campioni di quella metà dai campioni effettivamente acquisiti.

3.5 Ambiente software

L'intera opera di implementazione dei differenti approcci CS di questa tesi è stata svolta utilizzando la suite MATLAB 2011b nella sua versione per i processori a 64-bit. Le macchine utilizzate sono dotate rispettivamente di CPU quad-core AMD FX 3.6 GHz, 8 GB di RAM, Windows 7 Professional 64-bit e di CPU quad-core Intel Q9550 2.83 GHz, 4 GB di RAM e Windows 7 Professional 64-bit. Diversi toolbox sono stati utilizzati per compiti specifici, dalla lettura delle immagini alla soluzione del problema di ottimizzazione: in tutti i casi il toolbox preso in considerazione viene distribuito gratuitamente dagli autori attraverso il proprio sito internet o repository esterne. Di seguito viene presentata una lista dettagliata dei software:

Algoritmo veloce per la minimizzazione TV-L1-L2

- *Nifti* toolbox, utilizzato per leggere le immagini in formato DICOM del volume di acquisizione, è disponibile all'indirizzo:

<http://www.rotman-baycrest.on.ca/~jimmy/NIFTI/>

- *cvx*, toolbox per la risoluzione di problemi di ottimizzazione convessa. Permette l'ottimizzazione delle norme ℓ_1 o ℓ_2 di un segnale soggetto a vincoli di consistenza. La sua semplicità tuttavia lo rende adatto all'applicazione in problemi basilari di CS, ma la non ottimizzazione dell'algoritmo e la bassa efficienza nell'utilizzo della memoria ne limitano l'uso a segnali monodimensionali o al più ad immagini di dimensioni molto ridotte (64x64). È disponibile all'indirizzo:

<http://cvxr.com/cvx/>

- *l1magic*, come il precedente risolve il problema di ottimizzazione convessa, ma in questo caso è rivolto specificatamente all'applicazione del CS. Una funzione di particolare interesse presente all'interno del pacchetto permette di ricostruire i dati tramite regolarizzazione del Total Variation (TV). È reperibile all'indirizzo:

<http://users.ece.gatech.edu/~justin/l1magic/>

- *l1-ls*, risolve problemi di minimizzazione della norma ℓ_1 regolarizzata con termine ai minimi quadrati. Si trova all'indirizzo:

<http://www.stanford.edu/~boyd/l1-ls/>

- *SparseMRI-v0.2* una suite di funzioni matlab dedicate alla ricostruzione di immagini di risonanza magnetica utilizzando vari metodi di ottimizzazione, fra cui anche la sparsificazione tramite trasformata Wavelet. In quest'ultimo caso è necessario utilizzare un toolbox aggiuntivo (WaveLab850). I due software si trovano rispettivamente agli indirizzi:

<http://www.stanford.edu/~mlustig/SparseMRI.html>

<http://www-stat.stanford.edu/~wavelab/>

- *RecPF-v1.1*, un toolbox per la ricostruzione di immagini da acquisizioni parziali nel dominio di Fourier. Oltre ad essere pensato specificatamente per la ricostruzione di immagini MRI, la sua formulazione ottimizzata

3.5 Ambiente software

(soluzione lineare del problema) lo rende estremamente veloce rispetto a tutti gli altri toolbox, i quali attuano una ricerca non-lineare del minimo. Dato il suo ruolo centrale in questa tesi, questo metodo è stato esaminato approfonditamente in questo capitolo. È disponibile all'indirizzo:

<http://www.caam.rice.edu/~optimization/L1/RecPF/>

Capitolo 4

Risultati

4.1 Analisi statistica

Prima di presentare i risultati ottenuti in questo lavoro è bene introdurre le metodologie di valutazione degli errori e di confronto con i dati veri. In questa tesi si valuterà solamente l'efficacia della ricostruzione delle immagini di simulazione di una risonanza magnetica, lasciando la trattazione degli errori nei dati reali e nella stima dei parametri emodinamici a [38]. Il lavoro svolto in questa tesi è quindi incentrato sulla ricostruzione, cercando di migliorare il risultato in termini assoluti rispetto al metodo già esistente.

4.1.1 Definizione degli errori

Per avere un riscontro diretto sull'efficacia del metodo e poter confrontare i risultati in modo semplice, è stato scelto di utilizzare un solo tipo di errore. Analizzando una singola immagine si definisce la radice dello scarto quadratico medio (Root Mean Square Error, RMSE) come:

$$RMSE = \sqrt{\frac{1}{N} \sum_i (U_i - I_i)^2}, \quad (4.1)$$

dove U è l'immagine ricostruita, I è l'immagine originale, i sono i campioni appartenenti all'encefalo ed N il loro numero.

La scelta di considerare solo parte dei pixel dell'immagine risulta fondamentale per il confronto fra i metodi. Infatti i pixel dello sfondo, con valore di intensità pari a zero, vengono ricostruiti con un certo grado di incertezza, che

Risultati

dipende dal metodo utilizzato e dal livello di sottocampionamento. Includere anche tali campioni nel calcolo dell'errore potrebbe falsare la stima, qualora ci fosse una buona ricostruzione all'interno dell'encefalo ma una elevato rumore nello sfondo. Per il calcolo dell'errore è necessario quindi definire preventivamente una ROI (Region Of Interest) che comprenda solo i campioni d'interesse.

Se si considera ora tutta la sequenza dinamica e non più la singola immagine, si ottiene un vettore di RMSE di lunghezza pari al numero di immagini nella sequenza. Per continuità con le scelte effettuate in [38], vista la presenza di più data-set da confrontare è conveniente considerare la media del RMSE su tutti i tempi di acquisizione. D'ora in poi sarà questo l'errore di riferimento presente nelle tabelle di confronto, il quale viene definito come:

$$meanRMSE = \frac{1}{T} \sum_{j=1}^T RMSE_j, \quad (4.2)$$

dove RMSE è calcolato con la formula 4.1 e T è il numero di acquisizioni temporali. Per comodità in seguito si farà riferimento, con abuso di notazione, all'errore appena definito con l'acronimo RMSE.

4.1.2 Analisi temporale su ROI

Uno strumento maggiormente qualitativo per confrontare i risultati, che vada oltre alla semplice visualizzazione delle immagini ricostruite, è il grafico dell'andamento temporale di una ROI (nel nostro caso di dimensioni 3×3) o di un singolo pixel. Tale grafico si ottiene ponendo in ascissa i tempi di ricostruzione delle immagini e in ordinata il livello di grigio del pixel considerato (corrispondente all'attenuazione del segnale NMR nel voxel associato). Giustapponendo l'andamento dei dati veri con quello delle ricostruzioni si capisce intuitivamente quanto, come e dove le ricostruzioni compiono errori.

4.1.3 Scelta dei parametri

L'algoritmo di ricostruzione delle immagini necessita di un opportuno setup dei parametri per funzionare al meglio. Prima di procedere alla ricostruzione delle immagini, è bene quindi effettuare un *benchmark* per individuare i

4.2 Risultati delle simulazioni

parametri ottimali. Questi corrispondono, come spiegato nel paragrafo 3.2, al peso dell'informazione a priori nella funzione di minimizzazione, al peso della baseline nei punti in cui non vi è prior e al peso della ricostruzione precedente (solo nel caso *RecPFalt*).

È stato creato un iteratore per le immagini della simulazione a cui vengono passate sotto forma di matrice tutte le combinazioni possibili dei valori scelti per i parametri (20 combinazioni per il modello semplice e 80 per quello con pesatura aggiuntiva). I valori possibili sono i seguenti:

- $\alpha = 10^{-7}, 2.5 \cdot 10^{-7}, 5 \cdot 10^{-7}, 7.5 \cdot 10^{-7}, 10^{-6}$;
- $peso_0 = 0.2, 0.4, 0.6, 0.8$;
- $peso_1 = 0.2, 0.4, 0.6, 0.8$.

Con $peso_0$ ci si riferisce al valore di α nell'equazione 3.21 ed a quello di β nell'equazione 3.23, mentre con $peso_1$ si indica il valore di γ nell'equazione 3.23.

All'iteratore è stato quindi passato per la ricostruzione il data-set del phantom ottenuto tramite accelerazione a 8. Una volta completata la ricostruzione della sequenza per ogni set di parametri si è calcolato l'errore medio secondo la formula 4.2: ordinando in maniera crescente gli errori ottenuti si è scelto semplicemente il set di parametri che ha generato l'errore di ricostruzione più basso. I risultati sono riassunti in tabella 4.1.

PHANTOM				
	α	$peso_0$	$peso_1$	$RMSE$
<i>RecPFmod</i>	$7.5 \cdot 10^{-7}$	0.8	/	6.7430
<i>RecPFalt</i>	$5 \cdot 10^{-7}$	0.8	0.8	10.1030

Tabella 4.1: Parametri di ricostruzione scelti per le simulazioni

4.2 Risultati delle simulazioni

Come già spiegato nel paragrafo 3.3, per la valutazione preliminare dei diversi metodi sono state generate delle simulazioni con il metodo del *phantom* di

Risultati

Shepp-Logan [39]. La nona acquisizione (chiamata acquisizione al tempo T_1 o al nono campione) è la prima ad essere effettivamente sottocampionata. Nei grafici 4.1 e 4.2 vengono riportate le immagini originali e le immagini ricostruite dai vari metodi alle diverse accelerazioni. La tabella 4.2 completa inoltre le informazioni sugli errori di ricostruzione.

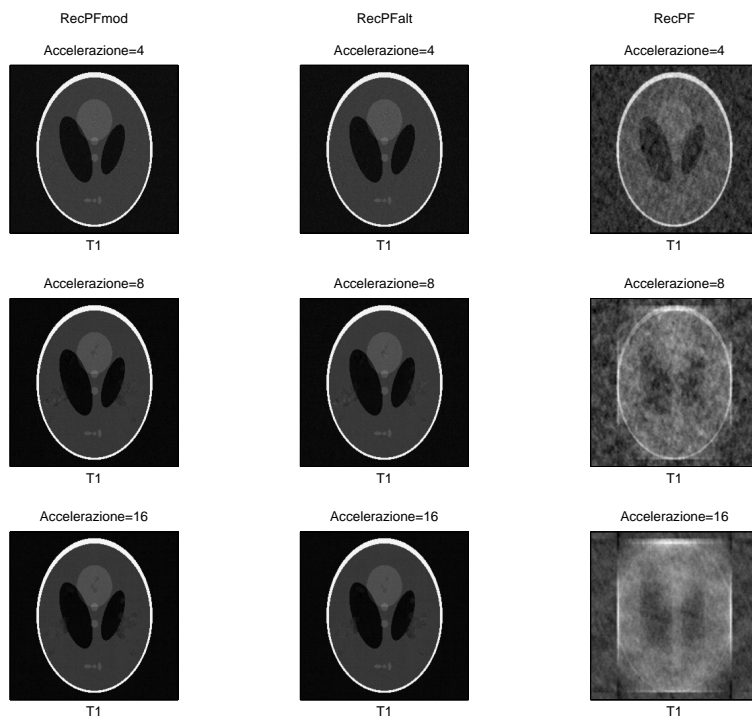


Figura 4.1: Immagini ottenute dal Phantom originale in Fig. 3.1 al tempo T_1 (prima acquisizione sottocampionata) con i vari metodi e le varie accelerazioni.

Verifichiamo ora il comportamento dei tre metodi a confronto andando ad analizzare prima l'andamento temporale delle intensità dei pixel appartenenti a due ROI 3×3 non appartenenti alla maschera indicate con N1 e N2 e poi quelle di altre due ROI analoghe chiamate C1 e C2 in cui ci sia passaggio simulato di Gadolinio, come evidenziato in Fig. 4.3. Le ROI C1 ed N1 appartengono, come si può notare, alla stessa struttura, così come le ROI C2 ed N2 sono state definite in una regione omogenea dell'encefalo, per valutare la robustezza dei tre algoritmi in maniera adeguata. I risultati sono stati raggruppati per ROI e per accelerazione, per sottolineare le differenze nella ricostruzione della stessa

4.2 Risultati delle simulazioni

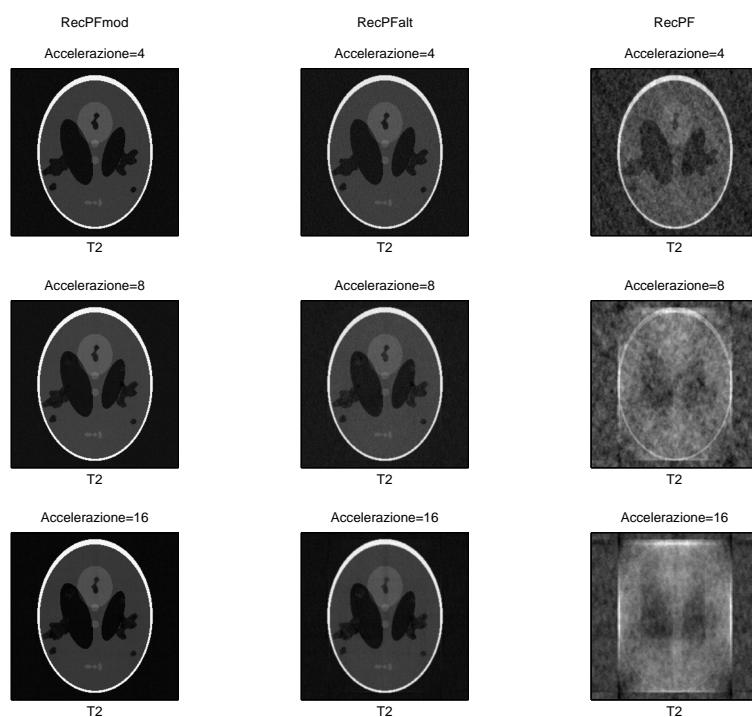


Figura 4.2: Immagini ottenute dal Phantom originale in Fig. 3.1 al tempo T_2 (momento di picco del mezzo di contrasto) con i vari metodi e le varie accelerazioni.

zona fra i vari metodi. Non sono state prese in considerazione analisi su pixel singoli a causa dell'elevata quantità di rumore presente in risultati di questo tipo: analizzando ROI invece che pixel singoli permette di ridurre di molto il rumore, grazie all'operazione di media che viene effettuata. La scelta delle ROI è stata fatta in modo tale da avere la possibilità di confrontare tessuti omogenei in cui vi sia la contemporanea presenza di zone comprese nella maschera e di zone da essa escluse. Per chiarezza di esposizione, i primi 8 campioni di ogni ricostruzione (gli 8 campioni utilizzati per ottenere la media delle baseline, il nostro prior spaziale) sono stati uguagliati ai dati originali.

Risultati

RMSE su PHANTOM			
Metodi	<i>RecPFmod</i>	<i>RecPFalt</i>	<i>RecPF</i>
Accelerazione 4	2.5515	4.4193	9.1435
Accelerazione 8	2.7123	5.4178	18.2003
Accelerazione 16	2.7908	5.9397	20.6354

Tabella 4.2: Errori di ricostruzione con simulazioni Shepp-Logan phantom

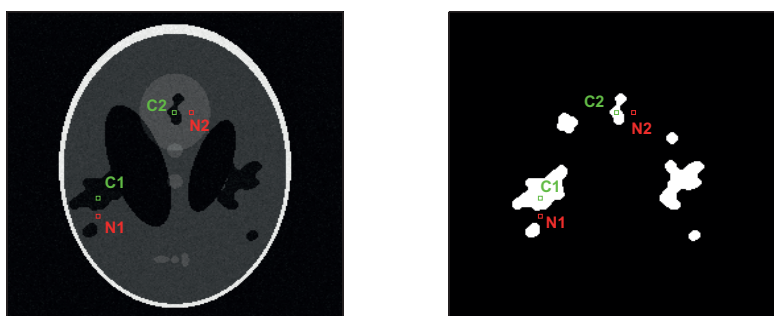


Figura 4.3: Le quattro ROI 3×3 prese in considerazione per la discussione dei risultati: le due ROI rappresentative delle zone interessate dal passaggio diretto del mezzo di contrasto, sono centrate in $(70, 155)$ e in $(128, 84)$. Le due ROI evidenziate in rosso invece saranno quelle rappresentative del comportamento dei tre metodi in punti non direttamente interessati dal passaggio del mezzo di contrasto, e sono centrate in $(70, 170)$ e in $(141, 84)$.

4.2 Risultati delle simulazioni

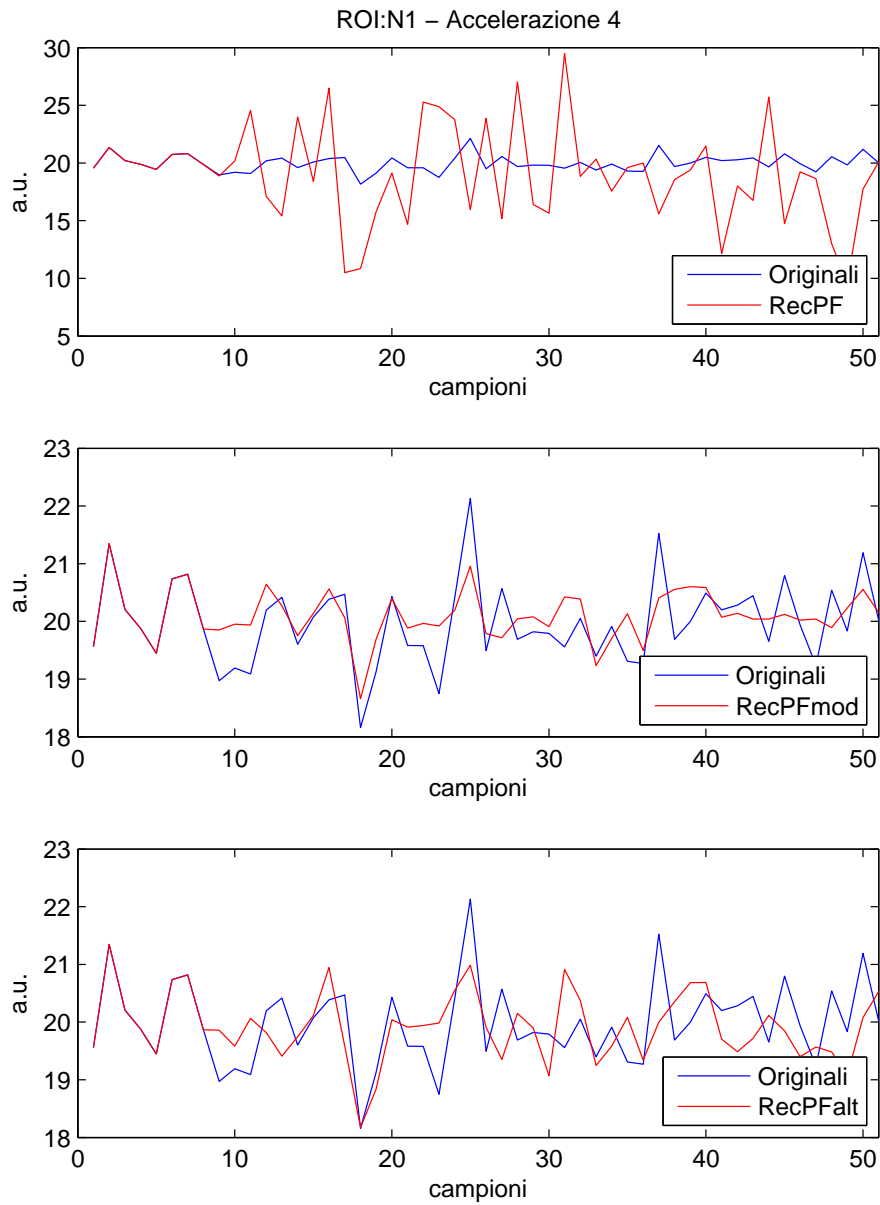


Figura 4.4: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 4 nella ROI N1 3×3 .

Risultati

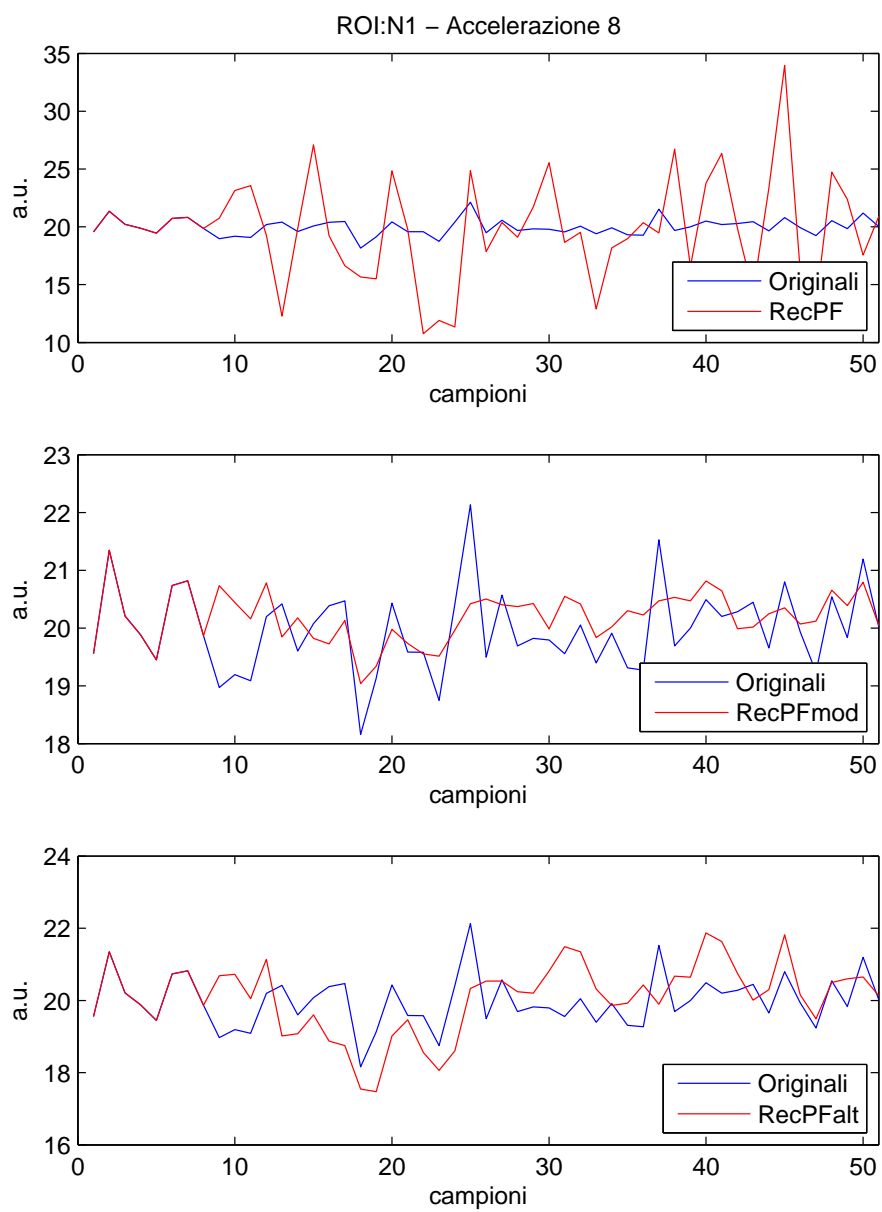


Figura 4.5: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 8 nella ROI N1 3×3 .

4.2 Risultati delle simulazioni

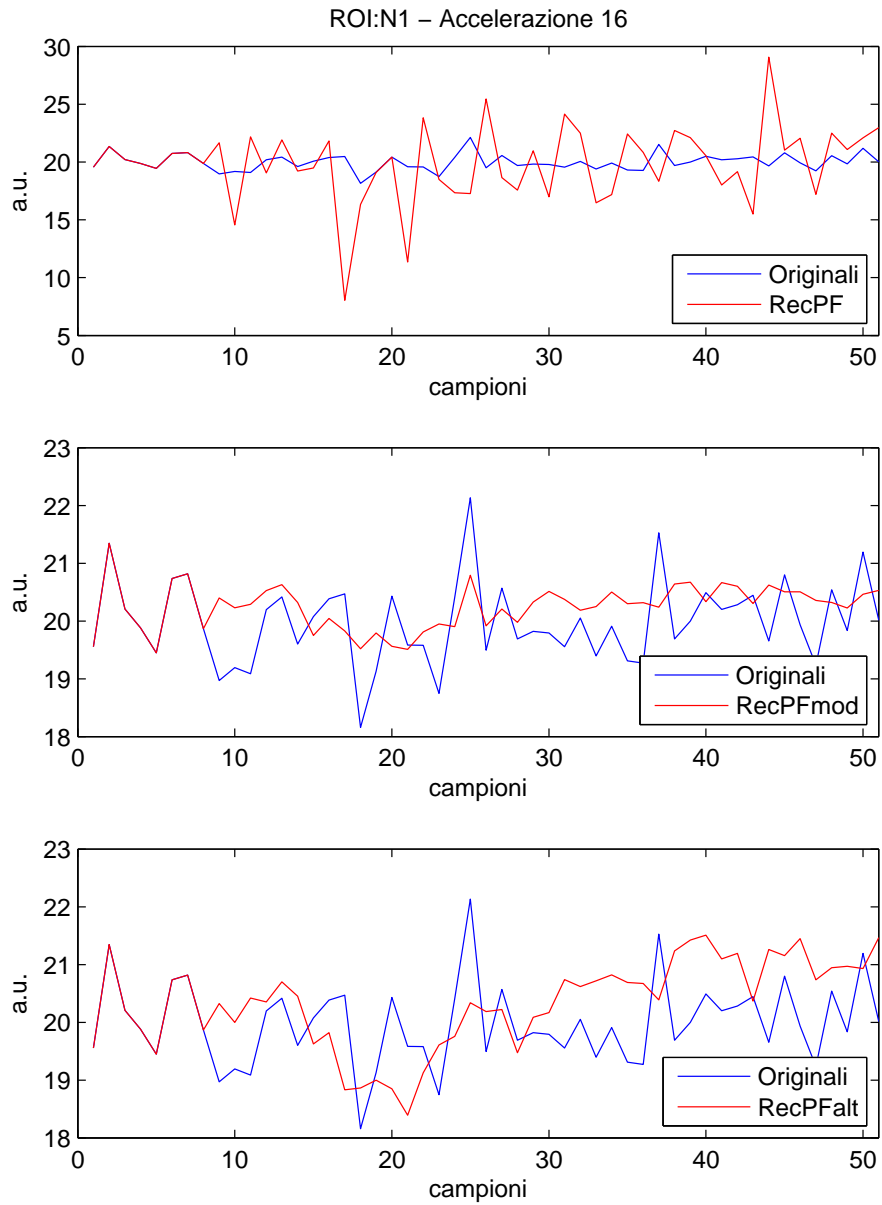


Figura 4.6: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 16 nella ROI N1 3×3 .

Risultati

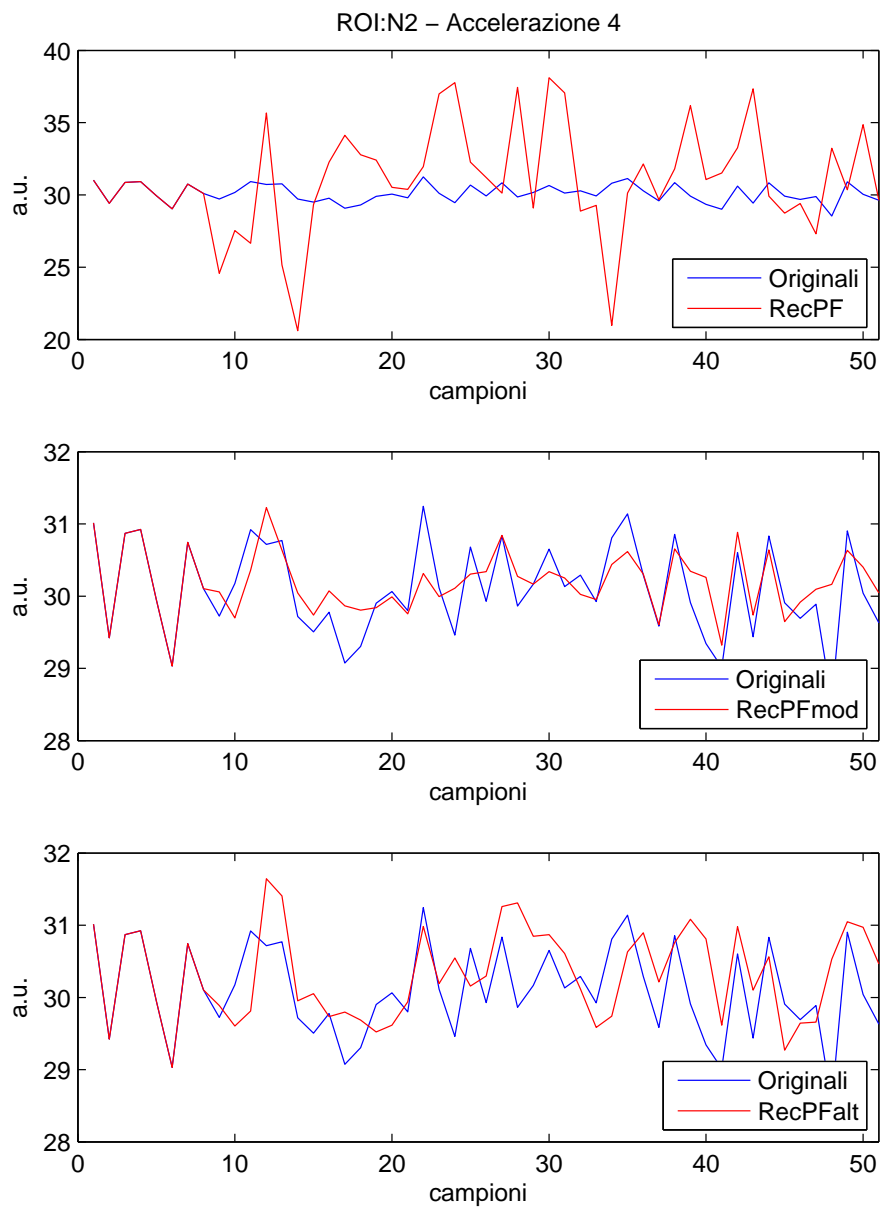


Figura 4.7: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 4 nella ROI N2 3×3 .

4.2 Risultati delle simulazioni

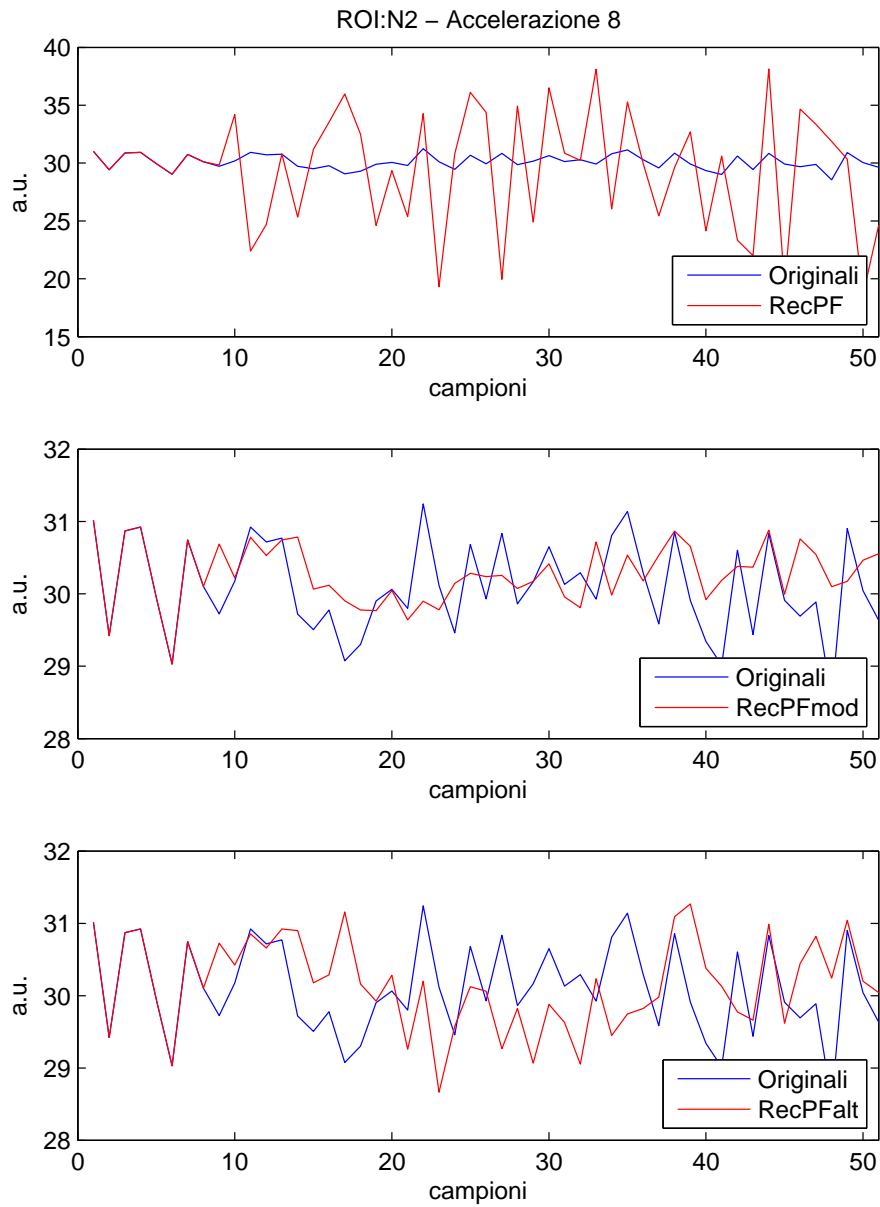


Figura 4.8: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 8 nella ROI N2 3×3 .

Risultati

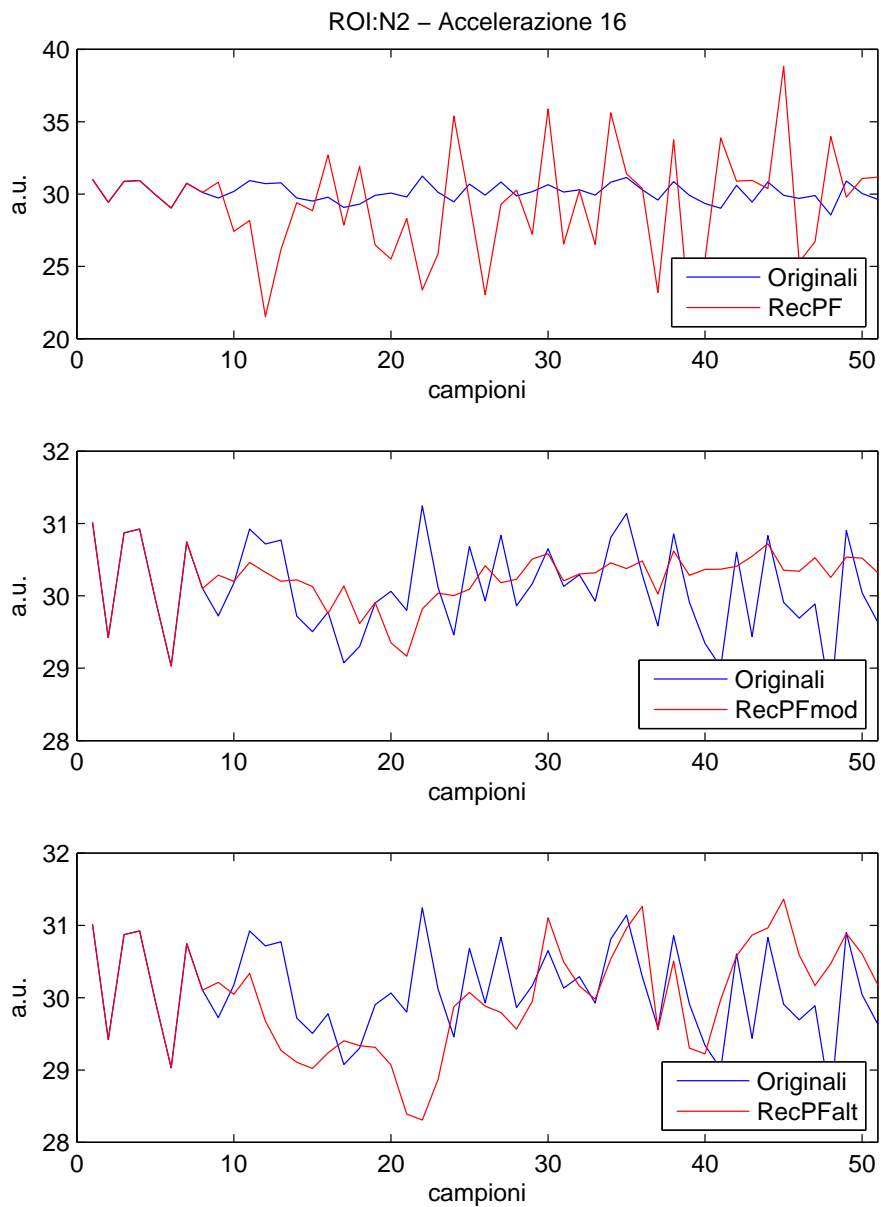


Figura 4.9: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 16 nella ROI N2 3×3 .

4.2 Risultati delle simulazioni

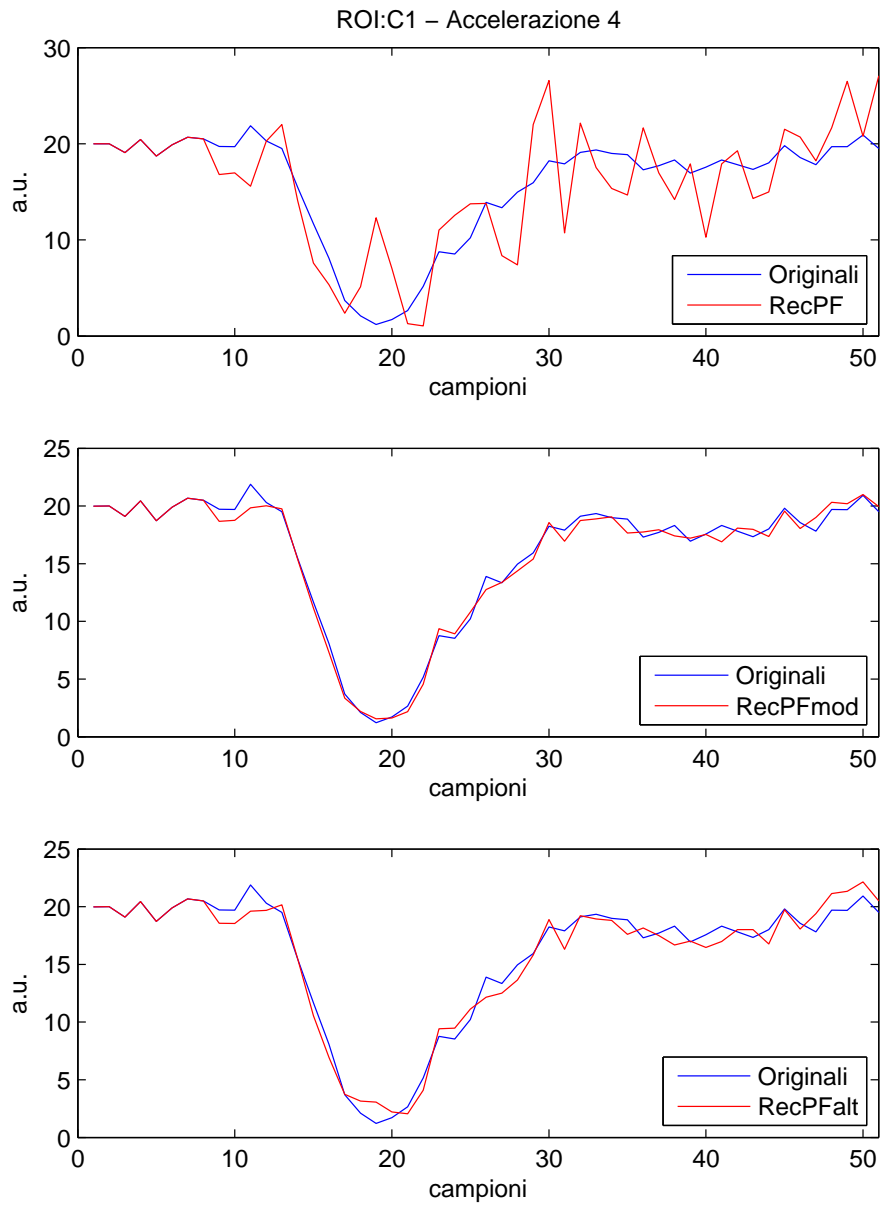


Figura 4.10: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 4 nella ROI C1 3×3 .

Risultati

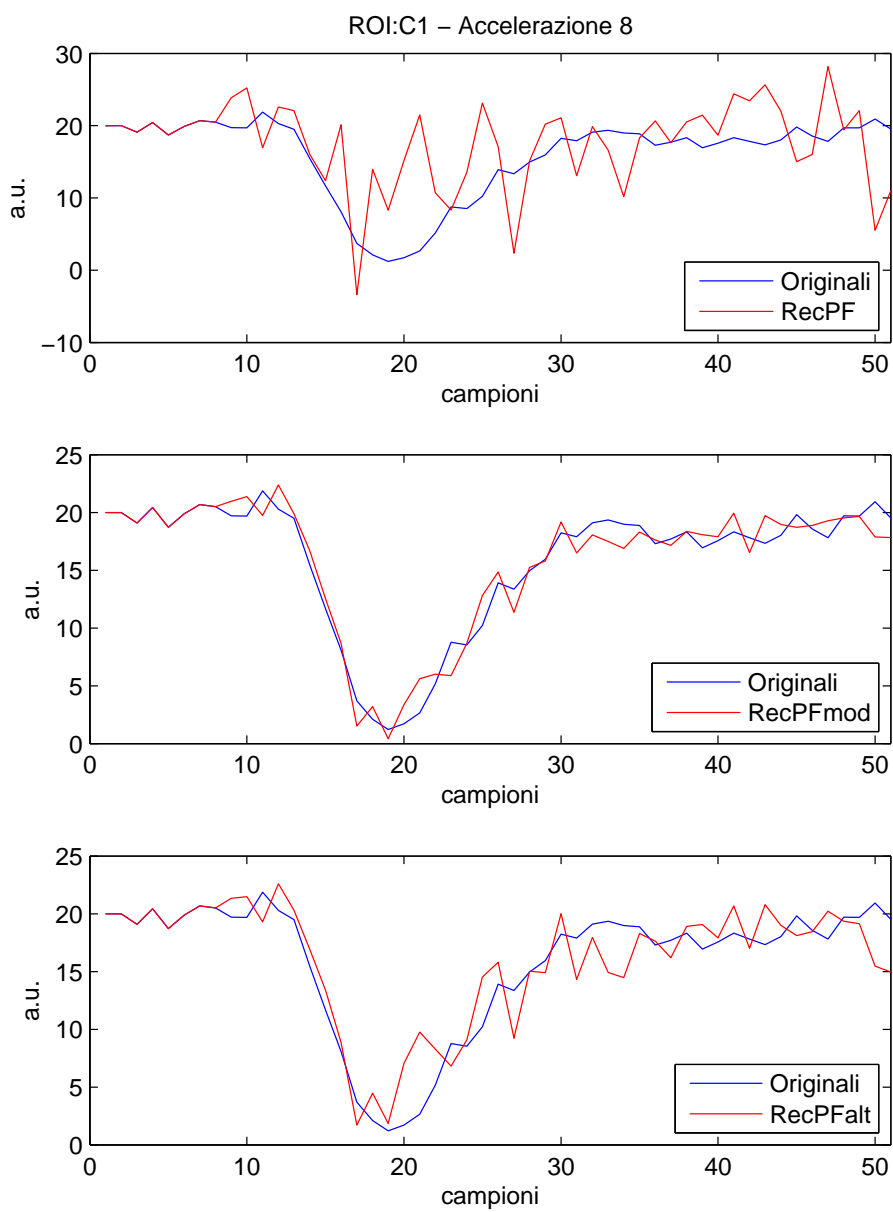


Figura 4.11: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 8 nella ROI C1 3×3 .

4.2 Risultati delle simulazioni

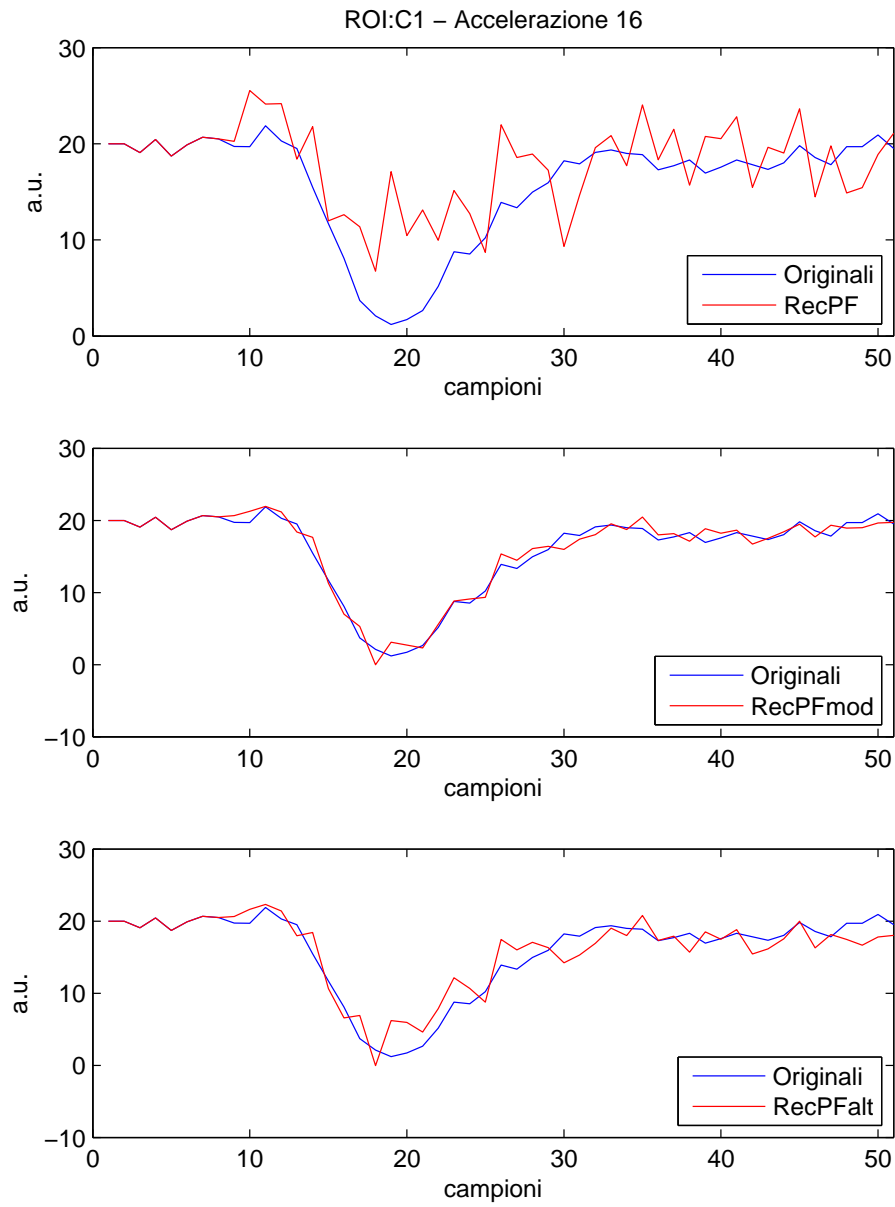


Figura 4.12: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 16 nella ROI C1 3×3 .

Risultati

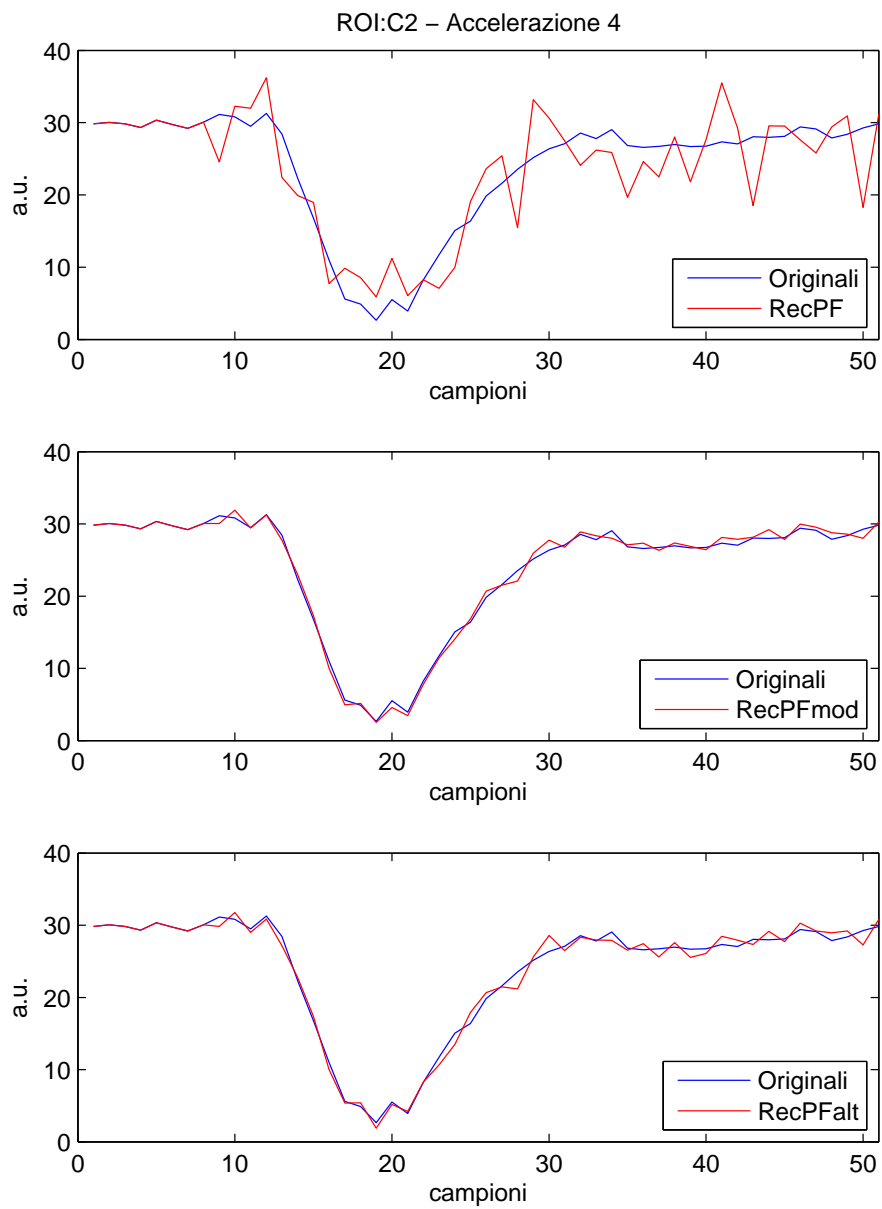


Figura 4.13: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 4 nella ROI C2 3×3 .

4.2 Risultati delle simulazioni

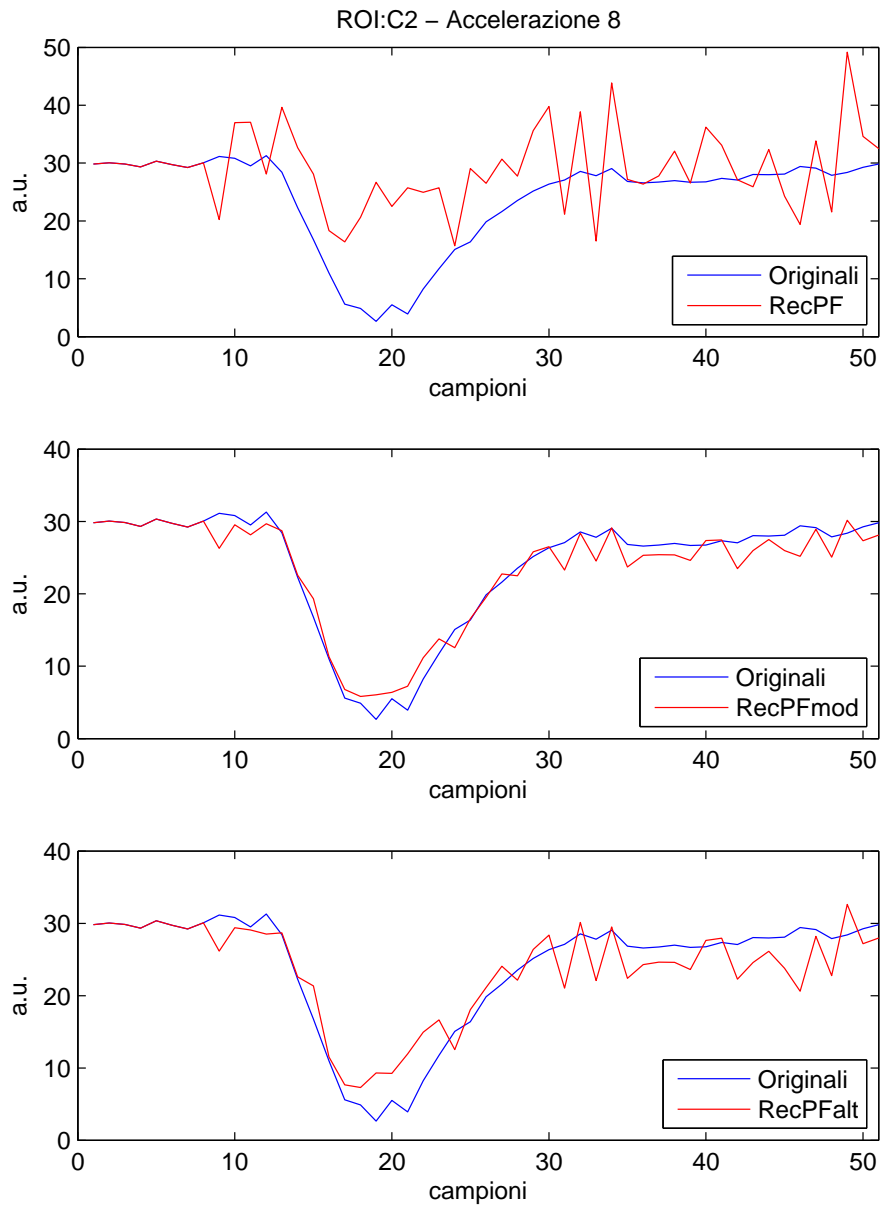


Figura 4.14: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 8 nella ROI C2 3×3 .

Risultati

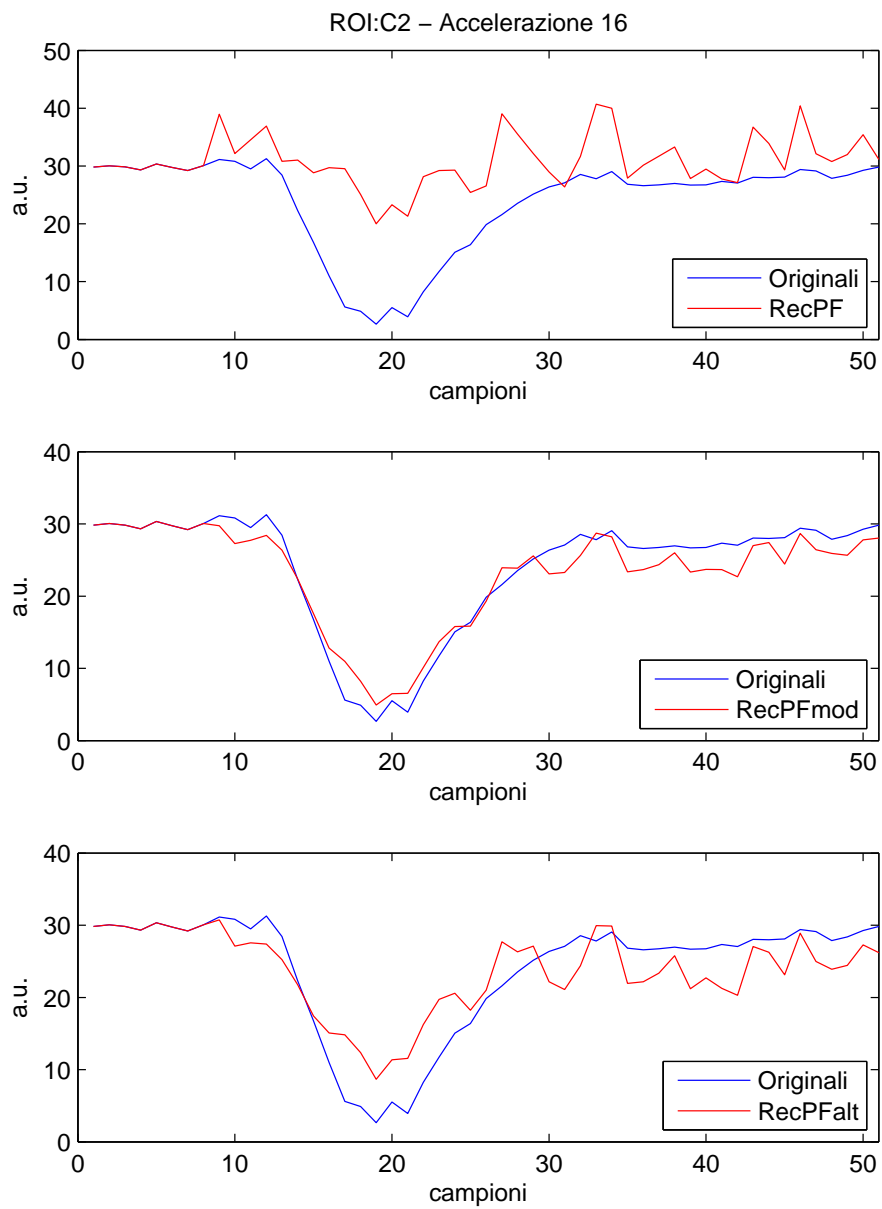


Figura 4.15: Confronto fra i dati originali e quelli ricostruiti dai tre diversi metodi ad accelerazione 16 nella ROI C2 3×3 .

Capitolo 5

Discussione

5.1 Analisi visiva spaziale dei risultati

Confrontando i dati originali mostrati in figura 3.1 con le ricostruzioni ottenute utilizzando i tre metodi, rappresentate nelle Figg. 4.1 e 4.2, è possibile analizzare visivamente la fedeltà delle ricostruzioni. In questo capitolo definiremo con T_1 l'istante della prima acquisizione sottocampionata e con T_2 l'istante del picco del mezzo di contrasto.

5.1.1 Ricostruzione a T_1

In riferimento alla Fig. 4.1 si può discutere della qualità delle ricostruzioni spaziali al nono campione (tempo T_1), cioè prima dell'arrivo del mezzo di contrasto. La differenza fra la ricostruzione con il metodo originale (*RecPF*, nella colonna 3) e i due metodi da noi modificati (*RecPFmod* e *RecPFalt*, rispettivamente nelle colonne 1 e 2) è oggettivamente evidente e salta immediatamente all'occhio. Già ad accelerazione 4, *RecPF* ha un comportamento insufficiente, presentando evidenti artefatti tipici da sottocampionamento e perdendo buona parte della sua informazione. La zona del phantom in basso, che presenta tre strutture ellissoidali affiancate, non è visibile già al livello di sottocampionamento meno intenso. Ad accelerazione 8 si distinguono a malapena solamente le due strutture scure (che simulano la presenza dei ventricoli), mentre tutte le altre strutture più chiare non risultano visibili. Gli artefatti sono ovviamente più evidenti, e anche il bordo del phantom inizia a perdere definizione. Ad

Discussione

accelerazione 16 non si riconosce più la forma dell'encefalo, che acquisisce forma rettangolare a causa dell'elevato fattore di sottocampionamento. Le uniche strutture vagamente riconoscibili sono i due ventricoli scuri al centro. In nessuna delle tre immagini accelerate è possibile ricavare eventuale informazione presente nei dati originali su parametri emodinamici o strutture patologicamente danneggiate.

Il comportamento di *RecPFmod* e *RecPFalt* è, invece, soddisfacente. Tutte le strutture rimangono visibili persino ad accelerazione 16, anche se si presentano artefatti all'interno delle zone appartenenti alla maschera (dove, cioè, verrà simulato il passaggio del mezzo di contrasto). Questi artefatti, probabilmente, sono causati dalla transizione netta fra interno ed esterno alla maschera e da una continuità non ideale dell'algoritmo fra le zone interessate dal passaggio del mezzo di contrasto e quelle da esso non coinvolte. Gli artefatti sono comparabili fra i due metodi, che si comportano in maniera molto simile, confermandosi temporalmente robusti.

5.1.2 Ricostruzione a T_2

In riferimento alla Fig. 4.2 si può discutere della qualità delle ricostruzioni spaziali al campione 18 (tempo T_2), cioè nell'istante del picco del mezzo di contrasto.

Ancora una volta *RecPF* si dimostra inadeguato: come già indicato nel paragrafo precedente, all'aumentare dell'accelerazione si ha una progressiva perdita di informazione nell'immagine e una scomparsa delle strutture dell'encefalo. Ad accelerazione 4 le zone in cui è presente il contrasto sono visibili, ma gli artefatti sono già uniformemente presenti ed impediscono qualsiasi valutazione clinicamente rilevante. Ad accelerazioni pari a 8 e 16 si ha, nuovamente, una scomparsa di tutte le strutture ad eccezione di quelle più scure (che, però, soffrono di un blurring eccessivo che fa scomparire la transizione netta sui loro bordi) ed una modifica della barriera emato-encefalica che, come già visto in precedenza, da essere ellissoidale diventa rettangolare.

RecPFmod e *RecPFalt* si comportano egregiamente: a tutte le accelerazioni viene mantenuta l'informazione sui bordi di ogni struttura, le zone di contrasto rimangono visibili ed evidenti e non vi è presenza evidente di alia-

5.2 Andamento temporale della ricostruzione

sing da sottocampionamento. Fra i due metodi quello migliore sembra essere *RecPFmod* che non presenta blurring nemmeno al livello di accelerazione massimo. *RecPFalt*, invece, pur mantenendo inalterata l'informazione sui bordi, presenta un leggero fenomeno di blurring nelle zone non appartenenti alla maschera e qualche artefatto da sottocampionamento nelle zone dove il picco del contrasto è evidente. Eventuali strutture danneggiate o patologie risulterebbero evidenti utilizzando *RecPFmod* e *RecPFalt*, al contrario di quanto accadrebbe con l'uso di *RecPF*.

5.2 Andamento temporale della ricostruzione

Passiamo ora all'analisi dell'intensità dei pixel, andando ad osservare il suo andamento temporale nelle due coppie di ROI di dimensione 3×3 pixel definite in precedenza ed evidenziate in Fig. 4.3.

5.2.1 ROI non appartenenti alla maschera

L'analisi dei risultati nelle ROI non appartenenti alla maschera ha lo scopo di verificare che gli algoritmi non stravolgano la natura fisiologica del segnale nelle zone in cui, teoricamente, non vi sia passaggio di mezzo di contrasto.

Il comportamento dei tre metodi analizzato nella ROI N1 (Fig. 4.4, 4.5 e 4.6) è sostanzialmente simile a quello visibile nella ROI N2 (Fig. 4.7, 4.8 e 4.9).

Accelerazione minore

Ad accelerazione 4 (Fig. 4.4 e 4.7 nelle ROI N1 ed N2 rispettivamente), risulta evidente la differenza fra il metodo originale e i due metodi modificati in questa tesi. Già a prima vista, infatti, è possibile notare che la scala in ordinata dei valori dell'intensità dei pixel sia nel caso di *RecPF* non comparabile a quella dei valori originali, mentre nel caso dei due metodi modificati essa si mantenga omogenea a quella presentata dai dati non sottocampionati.

Analizzando più nel dettaglio il comportamento dei metodi, si nota come *RecPF* mantenga un andamento molto rumoroso rispetto ai dati originali. Questo comportamento sarebbe comprensibile se l'analisi fosse stata effettuata su

Discussione

un singolo pixel (a causa del rumore uniforme), ma essendo l'andamento valutato mediando il segnale su una ROI di 9 pixel, si può definire insufficiente la regolarità di *RecPF*. Tali picchi, simili a rumore uniforme, sono probabilmente dovuti agli artefatti da sottocampionamento visibili in Fig. 4.1 e 4.2 e discussi nel paragrafo 5.1.

Il comportamento di *RecPFmod* e *RecPFalt*, invece, è coerente ai dati originali. Vi è, temporalmente, quasi un'irregolarità minore (e quindi uno smoothing maggiore) rispetto ai dati originali. Questo risultato è prevedibile: il contributo della baseline del metodo *RecPFmod* (discusso nel paragrafo 3.2.2), infatti, tende a ricostruire i dati fuori dalla regione d'interesse in maniera più simile possibile a quelli del prior (la media del livello basale), attenuando così il rumore e rendendo di conseguenza il segnale molto più smooth rispetto agli altri due metodi.

Accelerazione media

Anche a questa accelerazione (Fig. 4.5 e 4.8) i risultati sono simili a quelli discussi nel paragrafo precedente: *RecPF* mantiene andamento molto irregolare mentre i due metodi modificati sono molto più smooth e rimangono nel range di valori di intensità fisiologico.

Accelerazione massima

Ad accelerazione massima, corrispondente a un sottocampionamento del *k-space* di un fattore sedici, i tre metodi si comportano come osservato in precedenza. *RecPF*, a causa degli artefatti da sottocampionamento sempre più elevati, esibisce un andamento temporale altamente irregolare. *RecPFmod* e *RecPFalt* continuano a ricostruire molto bene, anche se a questa accelerazione nella ROI N1 (Fig. 4.6) sembra esserci un bias fra la media dei valori originali e quella dei dati ricostruiti nella seconda metà dei campioni, anche se assolutamente non significativa. Nella ROI N2 di Fig. 4.9, invece, si può notare come *RecPFalt* abbia uno strano comportamento subito dopo il ventesimo campione, probabilmente dovuto ad un artefatto di ricostruzione che, comunque, non ha comportato errori evidenti nelle ricostruzioni successive nonostante

5.2 Andamento temporale della ricostruzione

la presenza nel modello 3.23 del termine che pesa la ricostruzione al passo precedente.

5.2.2 ROI appartenenti alla maschera

Passiamo ora all'analisi dei risultati nelle ROI C1 e C2 appartenenti alla maschera, descritti nelle Figg. 4.10 - 4.15.

Accelerazione minore

Come è evidente dalle Figg. 4.10 e 4.13, mentre i due metodi modificati si comportano in maniera praticamente perfetta rispetto ai dati originali, vi è una differenza nel comportamento del *RecPF* nelle due ROI. Solo in C2, infatti, è possibile distinguere in maniera chiara il picco negativo dovuto al passaggio del Gadolinio. In C1, invece, già ad accelerazione 4 il picco è meno visibile a causa di un'oscillazione del valore del segnale nel campione 19 rispetto ai valori nei campioni adiacenti.

Accelerazione media

Ad accelerazione media il comportamento dei tre metodi è molto simile nelle due ROI (Fig. 4.11 e 4.14). Il metodo originale, a causa degli elevati artefatti di sottocampionamento, perde qualsiasi informazione relativa alla localizzazione del picco del mezzo di contrasto. I due metodi modificati, invece, continuano a fittare bene i dati originali, nonostante un livello di rumore leggermente più elevato. Fra i due, *RecPFalt* è quello che sembra soffrire maggiormente di irregolarità, in accordo alle analisi effettuate nel paragrafo 5.1 relativo all'analisi visiva delle ricostruzioni.

Accelerazione massima

Ad accelerazione 16, mentre le ricostruzioni ottenute attraverso *RecPF* si confermano completamente irregolari e prive di informazioni utili, si accentua la differenza fra i due metodi modificati. *RecPFmod*, infatti, continua a fittare in maniera molto soddisfacente i dati in maniera piuttosto regolare, nonostante un leggero bias nella seconda parte dei campioni temporali nella ROI C2

Discussione

in Fig. 4.15. *RecPFalt*, invece, sembra soffrire di un'irregolarità maggiore. Mentre nella ROI C1 (Fig. 4.12) il picco negativo del contrasto rimane evidente nonostante l'irregolarità, questa gioca un ruolo più importante in C2, in quanto il picco viene leggermente mascherato (rimanendo comunque distinguibile) dal rumore presente nei campioni ad esso successivi. Questo fatto concorda con l'analisi effettuata nel paragrafo 5.1, quando si è affermato che nelle immagini ricostruite ad accelerazione 16 il metodo *RecPFalt* presentava artefatti da sottocampionamento nelle aree interne alla maschera (cioè dove sono state scelte C1 e C2).

5.3 Riflessioni generali

Va sottolineato nuovamente che i risultati ottenuti sono da prendere come un'indicazione di massima del comportamento dei tre metodi in un ambito simulato ed altamente ideale. Infatti, uno dei punti più critici degli algoritmi modificati è rappresentato dalla definizione e costruzione della maschera sull'immagine in analisi. Come già definito nella sezione 3.2.1, infatti, la maschera contenente le aree d'interesse su cui andrà direttamente ad agire il mezzo di contrasto è predeterminata ed esattamente nota: tale assunzione verrà a mancare in presenza di dati reali, comportando un probabile aumento sia dell'errore di ricostruzione sia degli artefatti da sottocampionamento. In condizioni reali infatti non si conosce a priori né in che zone il contrasto si manifesterà né in che modo il suo passaggio possa causare un drift di luminosità media, un ritardo temporale fra i due emisferi o, ancora, fenomeni di leakage. In questi dati simulati la maschera è calcolata invece a priori in maniera esatta: solo in quella zona si avrà passaggio di mezzo di contrasto, senza alcun fenomeno di leakage o drift di intensità. Inoltre, nella simulazione si hanno transizioni nette fra "tessuti" differenti, condizione che nei test su dati reali ovviamente non si verifica. I vasi sanguigni interessati dal passaggio del mezzo di contrasto nei dati reali, infatti, non presentano bordi netti come invece accade nelle strutture del phantom di Shepp-Logan. Nel caso di applicazione a dati reali, quindi, non si potrebbe utilizzare una maschera fissa e predeterminata come nel caso di questa simulazione, ma si dovrebbe eventualmente utilizzare una maschera variabile, adattativa, calcolata ba-

sandosi su acquisizione intermedie non sottocampionate a tempi noti scelti opportunamente attraverso studi specifici prima dell'esame diagnostico.

Per quanto riguarda l'analisi specifica degli algoritmi, si può spendere una parola riguardo la differenza del comportamento fra i due metodi modificati. *RecPFalt*, infatti, sembra fittare i dati in maniera leggermente peggiore rispetto a *RecPFmod*. Questa differenza è causata probabilmente dal modo in cui è stato definito *RecPFalt* utilizzando l'equazione 3.23. La presenza del termine relativo alla ricostruzione al passo precedente, infatti, comporta una propagazione degli errori di ricostruzione che aumenta con l'aumentare dell'accelerazione o con l'avanzamento temporale. Come si può notare dalle Figg. 4.1 e 4.2 infatti vi è un peggioramento in *RecPFalt* rispetto a *RecPFmod* sia all'aumentare dell'accelerazione allo stesso istante temporale sia nel confronto a tempi diversi (T_1 e T_2 in questo caso) mantenendo costante l'accelerazione.

Capitolo 6

Conclusioni e Sviluppi Futuri

Come già evidenziato nel capitolo precedente, i risultati di questa analisi sono da considerarsi indicativi, provenienti da simulazioni effettuate in condizioni altamente ideali rispetto alla situazione di analisi normale con data-set reali. Gli algoritmi finora analizzati saranno testati su dati reali provenienti da DSC-MRI in [38].

I risultati di questa tesi confermano che *RecPFmod* rappresenta un'ottima soluzione per l'applicazione del CS alla DSC-MRI. Attraverso il suo utilizzo vengono sensibilmente migliorati i risultati ottenuti con un algoritmo tradizionale come *RecPF*. Sottocampionamenti ridotti di 16 volte rispetto all'originale con *RecPFmod* hanno errori di ricostruzione temporale del segnale su ROI 3×3 minori rispetto a sottocampionamenti di 4 volte del *k-space* ricostruiti con *RecPF*. Anche l'approccio *RecPFalt* ricostruisce ottimamente: il suo utilizzo con dati reali, probabilmente, potrà dare risultati ancora migliori di quanto visto nelle simulazioni presentate in questa tesi, a causa del setting diverso della maschera e delle transizioni fra strutture differenti più smooth e meno nette. Con tale metodo, grazie alla pesatura aggiuntiva, si potranno ottenere ricostruzioni dove si noteranno in maniera minore le transizioni nette ai bordi della maschera utilizzata in questi dati simulati. La semplice analisi visiva delle ricostruzioni alle diverse accelerazioni conferma infatti la bontà dei due approcci da noi ideati: al contrario di *RecPF* (in cui ad accelerazioni medio-alte ogni informazione utile viene perduta) i nostri due algoritmi ricostruiscono in maniera soddisfacente le immagini da set anche altamente sottocampionati,

Conclusioni e Sviluppi Futuri

non presentando artefatti da sottocampionamento, blurring o aliasing evidente ad accelerazioni medio-basse. Solo con l'utilizzo di *RecPFalt*, ad accelerazioni alte, si nota un accenno di aliasing nelle zone interessate dal contrasto.

Una volta dimostrato che i metodi modificati hanno il potenziale per fornire moltissimi vantaggi al paziente e possono portare a un miglioramento sensibile negli esami di DSC-MRI, lo scopo su cui si dovranno intensificare le ricerche sarà quello della costruzione di una maschera (fissa, multipla o adattativa) in grado di identificare al meglio le zone coinvolte nel passaggio del mezzo di contrasto. Tale miglioramento è necessario, in quanto si è osservato in questa tesi che la presenza di una maschera esatta comporta ovviamente un miglioramento netto della ricostruzione attraverso *RecPFmod*. In generale, comunque, gli algoritmi di ricostruzione *mod* e *alt* rimangono validi per qualunque esame di MRI dinamico: basterà implementare una maschera diversa (ed eventualmente delle pesature diverse) a seconda delle zone interessate dal mezzo di contrasto per ottenere risultati notevolmente migliori in tutti gli ambiti d'interesse rispetto alle modalità in cui questi esami vengono effettuati al giorno d'oggi. Nel futuro, con grande probabilità, gli schemi di acquisizione EPI o cartesiani verranno soppiantati dalla produzione in massa di macchinari in grado di effettuare scansioni radiali o a spirale. Quando questo upgrade verrà effettuato, si potranno utilizzare schemi innovativi come k-t SENSE [40], k-t GRAPPA [41] e k-t FOCUSS [42], che utilizzano scansioni radiali nel *k-space* e ridondanza informativa temporale [43]. Visto che tale upgrade di tecnologia non avverrà in tempi brevi, considerando il momento in cui versa l'economia mondiale e in particolare quella italiana, il nostro approccio potrà essere considerato attuale ed interessante per molti anni, dato che non richiede modifiche hardware ai macchinari che vengono tutt'oggi utilizzati negli ospedali mondiali.

Appendice A

Definizioni

Definizione 3. Per qualsiasi i , D_i in 3.8 è una matrice $2 \times N$ tale che i due valori in $D_i u$ rappresentano le differenze finite locali orizzontali e verticali di u al pixel i , dove D_i è definito essere compatibile con la trasformata \mathcal{T} in uso. Gli operatori per le differenze finite globali orizzontali e verticali sono chiamati $D^{(1)}, D^{(2)} \in \mathbb{R}^{N \times N}$ in maniera tale che l' i -esima riga di $D^{(j)}$ sia identica alla j -esima riga di $D_i, j = 1, 2$.

Definizione 4. Indicando con $\text{sgn}(t)$ la funzione segno di $t \in \mathbb{R}$, viene definita la funzione

$$\text{SGN}(t) = \begin{cases} \{\text{sgn}(t)\} & t \neq 0, \\ [-1, 1] & t = 0, \end{cases}$$

che è il sottodifferenziale di $|t|$.

Per un $v \in \mathbb{R}^N$ viene definita invece la funzione

$$\text{SGN}(v) = \{\zeta \in \mathbb{R}^N : \zeta_i \in \text{SGN}(v_i), \forall i\},$$

che è il sottodifferenziale di $\|v\|_1$.

Definizione 5. Per ogni $A \in \mathbb{R}^{p \times n}$ si può affermare che

$$\partial_x \|Ax\| = \begin{cases} A^\top Ax / \|Ax\| & \text{se } Ax \neq 0; \\ A^\top h : \|h\| \leq 1, h \in \mathbb{R}^p & \text{altrimenti.} \end{cases}$$

Definizione 6. Per ogni $B \in \mathbb{R}^{m \times n}$ si può affermare che

Definizioni

$$\partial_x \|Bx\|_1 = \{B^\top g : g \in \text{SGN}(Bx)\}.$$

Definizione 7. Per definire il TV discreto, viene introdotto un operatore lineare del gradiente discreto. Se $u \in X$, il gradiente ∇u è un vettore in $Y = X \times X$ dato da

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2)$$

indicando con

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{se } i < N, \\ 0 & \text{se } i = N, \end{cases}$$

$$(\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{se } j < N, \\ 0 & \text{se } j = N, \end{cases}$$

per $i, j = 1, \dots, N$. Quindi, la Total Variation di u è definita, come in [44], da

$$TV(u) = \sum_{1 \leq i, j \leq N} |(\nabla u)_{i,j}|,$$

con $|y| = \sqrt{y_1^2 + y_2^2}$ per ogni $y = (y_1, y_2) \in \mathbb{R}^2$.

Definizione 8. Viene definita norma ℓ_1 di un vettore $x = \{x_1, \dots, x_n\}$ lo scalare

$$|x|_1 = \sum_{r=1}^n |x_r|.$$

Definizione 9. Viene definita norma ℓ_2 di un vettore $x = \{x_1, \dots, x_n\}$ lo scalare

$$\|x\|_2 = \sqrt{\sum_{k=1}^n |x_k|^2},$$

dove con $|x_k|$ si indica il modulo complesso.

Definizione 10. Viene definita matrice ortogonale una matrice quadrata M_1 che sia invertibile e la cui trasposta coincida con la sua inversa, cioè

$$M_1 \cdot M_1^\top = M_1^\top \cdot M_1 = I_n$$

dove con I_n si indica l'identità di ordine n .

Definizione 11. Viene definita matrice sparsa una matrice M_2 i cui elementi siano in gran maggioranza nulli.

Definizione 12. Due matrici quadrate M_3 e M_4 vengono definite simili quando esiste una matrice invertibile A tale che

$$M_3 = A^{-1}M_4A.$$

Definizione 13. Viene definita matrice diagonalizzabile una matrice quadrata M_5 che sia simile ad una matrice diagonale.

Definizione 14. Una matrice di Toeplitz M_6 (detta anche matrice a diagonali costanti) è una matrice in cui ogni diagonale da sinistra a destra è formata da un unico elemento $a_{i,j}$ in essa ripetuto. Matematicamente, si ha che

$$a_{i,j} = a_{i-1,j-1}$$

dove i e j rappresentano rispettivamente gli indici delle righe e delle colonne.

Definizione 15. Una matrice circolante è un caso particolare di matrice di Toeplitz in cui ogni riga è shiftata di un elemento a destra rispetto alla riga precedente. Se viene applicata la DFT ad una matrice circolante, questa viene diagonalizzata.

Definizione 16. La legge di Faraday-Lenz afferma che la forza elettromotrice f.e.m. indotta in un circuito chiuso da un campo magnetico B è pari all'opposto della variazione del flusso magnetico Φ_B del campo attraverso l'area abbracciata dal circuito nell'unità di tempo:

$$f.e.m. = -\frac{d\Phi_B}{dt}.$$

Appendice B

Elenco degli acronimi

Acronimo	Significato
ADC	Analog to Digital Converter
AIC	Analog to Information Converter
CS	Compressed Sensing
DCT	Discrete Cosine Transform
DICOM	Digital Imaging and COmmunications in Medicine
DSC	Dynamic Susceptibility Contrast
DWT	Discrete Wavelet Transform
EPI	Echo-Planar Imaging
FFT	Fast Fourier Transform
FID	Free Induction Decay
IID	Independent and Identically Distributed
JPEG	Joint Photographic Experts Group
LASSO	Least Absolute Shrinkage and Selection Operator
MATLAB	MATrix LABoratory
MPEG	Moving Pictures Experts Group
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
PSF	Point Spread Function
RecPF	Reconstruction from Partial Fourier data
RecPFmod	RecPF modificato
RecPFalt	RecPF alternativo

Elenco degli acronimi

Acronimo	Significato
RF	Radio Frequency
RIP	Restricted Isometry Property
RMSE	Root Mean Square Error
ROI	Region Of Interest
SNR	Signal Noise Ratio
TE	Time to Echo
TPSF	Transform Point Spread Function
TR	Time to Repeat
TV	Total Variation
VOI	Voxel of Interest

Bibliografia

- [1] J. Yang Y., Zhang, and W. Yin. A fast TVL1-L2 minimization algorithm for signal reconstruction from partial Fourier data. April 2010.
- [2] F. Bistolfi. *Campi Magnetici in Medicina*. 1986.
- [3] E.M. Purcell, H.C. Torrey, and R.V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 1946.
- [4] F. Bloch and H. Staub. Fission spectrum. *U.S. Atomic Energy Commission, Technical Information Service*, 1951.
- [5] D.K. Sodickson and W.J. Manning. Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays. *Magn. Reson. Med.*, 1997.
- [6] K.P. Pruessmann, M. Weiger, M.B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magn. Reson. Med.*, 1999.
- [7] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. on Inform. Theory*, 2006.
- [8] E. Candès and T. Tao. Near optimal signal recovery from random projections: Universal encoding strategies. *IEEE Trans. on Inform. Theory*, 2006.
- [9] D.L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 2006.
- [10] D.L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inform. Theory*, 2001.

BIBLIOGRAFIA

- [11] J. Tropp and A.C. Gilbert. Signal recovery from partial information via orthogonal matching pursuit,. *IEEE Trans. Inform. Theory*, 2007.
- [12] M. Lustig, D.L. Donoho, J.M. Santos, and J.M. Pauly. Compressed sensing MRI. *IEEE Signal Processing Magazine*, March 2008.
- [13] E.J. Candès and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, March 2008.
- [14] E. Candès and J. Romberg. Sparsity and incoherence in compressive sampling,. *Inverse Probab.*, 2007.
- [15] D.L. Donoho and P.B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, 1989.
- [16] E. Candès and T. Tao. Decoding by linear programming,. *IEEE Trans. Inform. Theory*, 2005.
- [17] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate information. *Communications on Pure and Applied Mathematics*, 2005.
- [18] A. Cohen, W. Dahmen, and R. DeVore. Compressed sensing and best k-term approximation,. *J. Amer. Math. Soc.*, 2009.
- [19] R. Tibshirani. Regression shrinkage and selection via the LASSO. *J. of the Royal Statistical Society Ser. B*, 58(1):267–288, 1996.
- [20] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 2008.
- [21] M. Rudelson and R. Vershynin. On sparse reconstruction from fourier and gaussian measurements,. *Communications on Pure and Applied Mathematics*, 2008.
- [22] D. Takhar, V. Bansal, M. Wakin, M. Duarte, D. Baron, K.F. Kelly, and R.G. Baraniuk. A compressed sensing camera: New theory and an implementation using digital micromirrors. In *Proc. Comp. Imaging IV SPIE Electronic Imaging*, San Josè.

- [23] D.S. Taubman and M.W. Marcellin. *JPEG 2000: Image compression fundamentals, standards and practice*. Norwell, MA: Kluwer, 2001.
- [24] L. He and L. Carin. Exploiting structure in wavelet-based bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 2009.
- [25] S. Mallat. *A Wavelet Tour of Signal Processing, 2nd ed.* 1998.
- [26] G. Wright. Magnetic resonance imaging. *IEEE Signal Processing Mag.*, 1997.
- [27] M. Lustig, D.L. Donoho, and J.M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR Imaging. *Magnetic Resonance in Medicine*, 2007.
- [28] T. Parrish and X. Hu. Continuous update with random encoding (CURE): A new strategy for dynamic imaging,. *Magn. Reson. Med.*, 1995.
- [29] D.C. Peters, F.R. Korosec, T.M. Grist, W.F. Block, J.E. Holden, K.K. Vigen, and C.A. Mistretta. Undersampled projection reconstruction applied to MR angiography,. *Magn. Reson. Med.*, 2000.
- [30] F.R. Korosec, R. Frayne, T.M. Grist, and C.A. Mistretta. Time-resolved contrast-enhanced 3D MR angiography,. *Magn. Reson. Med.*, 1996.
- [31] J.M. Santos, C.H. Cunningham, M. Lustig, B.A. Hargreaves, B.S. Hu, D.G. Nishimura, and J.M. Pauly. Single breath-hold whole-heart MRA using variable-density spirals at 3T,. *Magn. Reson. Med.*, 2006.
- [32] D. Donoho. For most large underdetermined systems of linear equations, the minimal l_1 -norm solution is also the sparsest solution. *Communications on pure and applied mathematics*, 2006.
- [33] L. He, T.C. Chang, S. Osher, T. Fang, and P. Speier. MR image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods. *UCLA CAM Report 06-35*, 2006.
- [34] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty. An efficient algorithm for compressed MR imaging using total variation and wavelets. *IEEE Conference on CVPR*, 2008.

BIBLIOGRAFIA

- [35] R. Courant. Variational methods for the solution of problems with equilibrium and vibration. *Bull. Amer. Math. Soc.*, 1943.
- [36] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE Trans. Image Process.*, 1995.
- [37] Y. Yang, W. Yin, Y. Zhang, and Y. Wang. A fast algorithm for edge-preserving variational multichannel image restoration. *SIAM J. on Imaging Sciences*, 2009.
- [38] P. Di Prima. Compressed Sensing in DSC-MRI, Parte II: Risultati sperimentali. *Università degli studi di Padova*, 2012.
- [39] L.A. Shepp and B.F. Logan. The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science*, 21(3):21–43, June 1974.
- [40] J. Tsao, P. Boesiger, and K.P. Pruessmann. k-t BLAST and k-t SENSE: dynamic MRI with high frame rate exploiting spatiotemporal correlations. *Magn. Reson. Med.*, 2003.
- [41] F. Huang, J. Akao, S. Vijayakumar, G.R. Duensing, and M. Limkeman. k-t GRAPPA: A k-space implementation for dynamic MRI with high reduction factor. *Magn. Reson. Med.*, 2005.
- [42] H. Jung, K. Sung, K.S. Nayak, E.Y. Kim, and J.C. Ye. k-t FOCUSS: a general compressed sensing framework for high resolution dynamic MRI. *Magn. Reson. Med.*, 2009.
- [43] F. Waier. Non-cartesian MRI scan time reduction through sparse sampling. *Delft University of Technology, Dip. Fisica Applicata, Olanda*, 2001.
- [44] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 2004.

Elenco delle figure

2.1	Immagine MR e relativo k -space	11
2.2	Esempio di domini x - t e x - f	13
2.3	La serie di impulsi spin-echo.	13
2.4	La serie di impulsi gradient-echo.	14
2.5	La sequenza di acquisizione Echo Planar Imaging.	14
2.6	Sottocampionamento in coefficienti wavelet	30
2.7	Struttura ad albero dei coefficienti wavelet.	32
2.8	Traiettoria cartesiana 2D implica sparsità 1D	36
2.9	Confronto ricostruzioni in angiografia 3D.	37
3.1	Phantom a T1 e T2	54
3.2	Campionamento random del k -space	55
4.1	Phantom ricostruito a T1	62
4.2	Phantom ricostruito a T2	63
4.3	ROI prese in analisi.	64
4.4	ROI N1 - Accelerazione 4	65
4.5	ROI N1 - Accelerazione 8	66
4.6	ROI N1 - Accelerazione 16	67
4.7	ROI N2 - Accelerazione 4	68
4.8	ROI N2 - Accelerazione 8	69
4.9	ROI N2 - Accelerazione 16	70
4.10	ROI C1 - Accelerazione 4	71
4.11	ROI C1 - Accelerazione 8	72
4.12	ROI C1 - Accelerazione 16	73
4.13	ROI C2 - Accelerazione 4	74

ELENCO DELLE FIGURE

4.14 ROI C2 - Accelerazione 8	75
4.15 ROI C2 - Accelerazione 16	76

Elenco delle tabelle

4.1	Parametri di ricostruzione scelti per le simulazioni	61
4.2	Errori di ricostruzione	64

