Studio della convergenza nell'analisi di immagini digitali

Sintesi della tesi di laurea di

Claudia Mizzon

Relatore interno:Lucia Caramellino

Relatore esterno:Giovanni Sebastiani

Lo studio svolto in questa tesi é motivato da uno specifico problema applicativo nel campo dell'analisi di immagini biomediche. Il problema medico coinvolto é quello della caratterizzazione di alcune rilevanti patologie mammarie. I dati a disposizione consistono in sequenze temporali di immagini tomografiche di Risonanza Magnetica Nucleare (RMN)[2]. Le immagini relative ad una fissata sezione bidimensionale delle mammelle, vengono acquisite l'una dopo l'altra in modo consecutivo dopo un tempo opportuno dall'iniezione in vena del paziente di un mezzo di contrasto. L'intensitá dell'immagine di RMN dipende dalla concentrazione locale del mezzo di contrasto. Le sequenze temporali di RMN con mezzo di contrasto, forniscono cosí un modo efficace e non invasivo per lo studio della perfusione sanguigna nei vari organi [3]. L'uso di questa tecnica per lo studio di patologie mammarie, é molto promettente in relazione ai cambiamenti di vascolarizzazione dei tessuti mammari in presenza di tumori benigni o maligni [3].

Per ottenere una immagine RMN non si utilizzano radiazioni ionizzanti come nella tomografia a raggi X, ma si sfrutta l'interazione tra fotoni con una energia caratteristica delle radio-onde e i nuclei atomici. L'energia della radiazione utilizzata, é circa 9 ordini di grandezza piú piccola di quella dei raggi X. Le immagini RMN, sono costituite da dati digitali. L'energia radiante emessa del campione, viene captata da appositi sensori e, tramite un convertitore elettronico, trasformata in forma digitale. Un' immagine digitale consiste in un insieme finito di numeri reali o complessi, organizzato in forma di matrice. Questo é estremamente vantaggioso dal momento che si possono immagazzinare una quantitá enorme di dati in uno spazio molto piccolo e poiché tali immagini, dopo l'acquisizione, possono essere elaborate tramite calcolatori.

Nel processo di acquisizione peró, a causa delle imperfezioni dell'esperimento, del rumore casuale e della distorsione geometrica o non-linearitá degli strumenti tramite i quali avviene l'acquisizione, l'immagine misurata é affetta da distorsioni. Spesso i radiologi, o piú in generale gli specialisti che utilizzano immagini, tendono a voler eliminare tali distorsioni tramite la loro esperienza ad osservare ed interpretare le immagini. A volte peró, tali distorsioni mascherano rilevanti informazioni. In questi casi poco o nulla puó fare lo specialista da solo. Questo é sempre piú attuale a causa della grande quantitá di immagini disponibile per un fissato paziente (TAC ed RMN di sezioni multiple, ecc.). É quindi importante trovare metodi capaci di risolvere il problema dell' " image restoration", cioé rimuovere o minimizzare le distorsioni presenti nell'immagine.

Tra i vari metodi per risolvere il problema del restauro di immagini, [7] troviamo quelli basati sull'inferenza bayesiana [3] i quali hanno mostrato di dare stime dell'immagine migliori rispetto ad altre metodologie che sfruttano solo i dati sperimentali, dal momento che i primi utilizzano anche la conoscenza a priori che si ha dell'immagine prima di acquisire tali dati. I metodi bayesiani per stimare i parametri, utilizzano comunemente algoritmi Monte Carlo dinamici (MCMC) che hanno un costo computazionale molto elevato e quindi un lungo tempo di esecuzione al calcolatore. Lo scopo di questo lavoro é proprio quello di ridurre il costo computazionale e quindi il tempo di esecuzione di questi algoritmi. Ridurre il tempo di esecuzione in campo medico é importante in special modo quando é necessario il tempo reale. In questo caso, la riduzione del tempo di esecuzione, permette di migliorare la capacitá diagnostica delle immagini e anche di aumentare il numero di pazienti esaminati per unitá di tempo.

Se si indica con $\mathbf{x}$ l'immagine "vera", ossia l'idealizzazione senza distorsioni dell'immagine acquisita $\mathbf{y}$, il processo di acquisizione fa si che l'immagine misurata sia affetta sia da errore sistematico che da errore di tipo casuale. In molti casi é possibile ricavare, dalla fisica dell'esperimento, un modello lineare $\mathbf{y}=\mathbf{Ax}+\mathbf{e}$ dove $\mathbf{A}$ é una trasformazione lineare che modellizza l'errore sistematico, mentre il termine additivo $\mathbf{e}$ rappresenta l'errore casuale. In questo lavoro viene considerato il caso $\mathbf{A}=\mathbf{I}$ dove $\mathbf{I}$ é la matrice identitá, poiché le immagini di RMN sono affette quasi esclusivamente da errore casuale additivo [1].

Uno dei problemi nell'analisi dell'immagine é appunto quello di stimare $\mathbf{x}$ a partire da $\mathbf{y}$. Per affrontare questo problema, un approccio molto valido é quello Bayesiano [5]. Seguendo tale approccio, si modellizza $\mathbf{x}$ tramite campi di Markov e si usano i metodi dell'inferenza bayesiana per stimare $\mathbf{x}$ a partire da $\mathbf{y}$. Si utilizzano le informazioni a priori note sulle caratteristiche dell'immagine vera $\mathbf{x}$ per costruire una distribuzione di probabilitá $p1(\mathbf{x})$ per $\mathbf{x}$. Tale distribuzione, unita alla distribuzione dei dati $p2(\mathbf{y} \mid \mathbf{x})$, determina, tramite il teorema di Bayes, la distribuzione a posteriori $p3(\mathbf{x} \mid \mathbf{y})$.

In questo lavoro come distribuzione a priori stata scelta la distribuzione di Geman e McClure:

$$p1(\mathbf{x} \mid \beta, \lambda) \propto \exp[\beta \sum_{i \sim j} \frac{1}{1+\lambda(x_i - x_j)^2}]$$

dove $i \sim j$ indica che x_i e x_j sono una coppia di pixels vicini; come distribuzione dei dati é stata scelta:

$$p2(\mathbf{y} \mid \mathbf{x}) = \frac{1}{(2\pi\sigma_e)^{\frac{m}{2}}} \prod_{i=1}^{m} \exp(-\frac{(y_i - x_i)^2}{2\sigma_e^2})$$

dove m é il numero di pixels che costituiscono l'immagine.

Tramite il teorema di Bayes, si ottiene la distribuzione a posteriori:

$$p3(\mathbf{x} \mid \mathbf{y}) \propto \exp[\beta \sum_{i \sim j} \frac{1}{1+\lambda(x_i - x_j)^2}] \frac{1}{(2\pi\sigma_e)^{\frac{m}{2}}} \prod_{i=1}^{m} \exp(-\frac{(y_i - x_i)^2}{2\sigma_e^2})$$

Uno stimatore dell'immagine $\mathbf{x}$ comunemente adottato, é la media a posteriori (MMSE). Purtroppo tale stimatore non é disponibile in forma chiusa. Questo é dovuto sia all'elevata dimensione dello spazio dove é definita la distribuzione a posteriori, che alla dipendenza tra le componenti di $\mathbf{x}$. Si ricorre quindi ad una approssimazione dello stimatore tramite la media campionaria su un campione dalla distribuzione a posteriori. Per calcolare il campione, vengono comunemente utilizzati algoritmi Monte Carlo basati sulle catene di Markov (MCMC). Il campione é costituito in pratica dagli ultimi valori di $\mathbf{x}$ lungo la catena. Sebbene tali valori non sono indipendenti e seguono solo approssimativamente la distribuzione a posteriori, la media campionaria converge alla media di $\mathbf{x}$ rispetto alla distribuzione a posteriori.

Gli algoritmi MCMC piú usati, sono il Gibbs sampler e il Metropolis [8]. Di solito ciascuna variabile x_i, viene aggiornata iterativamente in base a uno schema di visita casuale o deterministico. L'aggiornamento nel Gibbs sampler, avviene estraendo x_i secondo la distribuzione condizionata $p(x_i \mid \tilde{x_i})$, dove $\tilde{x_i}$ é l'insieme di tutte le variabili di x esclusa la x_i. Nell'algoritmo di Metropolis invece, viene proposto un nuovo valore $x_i^{'}$ estratto da una distribuzione $q(x_i \rightarrow x_i^{'})$ simmetrica rispetto allo scambio tra x_i e $x_i^{'}$ chiamata distribuzione di proposta. Il nuovo valore viene accettato con una probabilitá che dipende dal rapporto dei valori della distribuzione a posteriori corrispondenti a x_i e $x_i^{'}$. Se il nuovo valore non viene accettato, si mantiene il vecchio valore x_i.

Un estensione di tale algoritmo é quello di Metropolis -Hastings [5]. In questo caso la distribuzione di proposta non é simmetrica.

Purtroppo gli algoritmi MCMC convergono lentamente. Diversi approcci sono stati proposti per aumentare la velocitá di convergenza di questi algoritmi. Ad esempio, si possono cercare configurazioni iniziali di x ottimali o studiare la migliore strategia di visita [11] [12]. Un altro modo é quello di introdurre uno o piú parametri di rilassamento [13] [14] [15]. É inoltre possibile introdurre delle nuove variabili $\mathbf{u}$, dette ausiliarie e una nuova distribuzione a posteriori sullo spazio degli stati ampliato $U \times X$. La catena di Markov viene in questo caso definita sullo spazio esteso. La distribuzione sullo spazio esteso puó essere definita in diversi modi. Nel Simulated Tempering ad esempio, la distribuzione congiunta di $\mathbf{x}$ e $\mathbf{u}$ é tale che la condizionata di $\mathbf{x}$ rispetto a $\mathbf{u_0}$ é l'originaria distribuzione a posteriori per $\mathbf{x}$ [16]. Il valore $\mathbf{u_0}$ é un particolare valore di $\mathbf{u}$. Seguendo un altro approccio invece, la distribuzione congiunta é tale che la sua marginale rispetto a $\mathbf{x}$ é proprio la distribuzione a posteriori originaria per $\mathbf{x}$. Quest'ultimo approccio ha il vantaggio che tutti i campioni estratti secondo la distribuzione estesa tramite la catena di Markov, vengono utilizzati per calcolare la media campionaria. Nell'altro caso invece solo i campioni con $\mathbf{u}=\mathbf{u_0}$ vengono utilizzati. Un esempio molto valido del secondo approccio é costituito dall'algoritmo di Swendsen-Wang [9]. In questo caso le variabili ausiliarie sono connesse a dei "cluster" di pixel. L'algoritmo procede aggiornando alternativamente $\mathbf{x}$ e $\mathbf{u}$ secondo le rispettive distribuzioni condizionate rispetto a $\mathbf{u}$ e $\mathbf{x}$. Questo algoritmo é stato utilizzato con successo per esplorare distribuzioni di Gibbs corrispondenti ai modelli di Ising e di Potts. Recentemente Higdon ha proposto un algoritmo seguendo il secondo approccio di cui sopra per applicazioni a immagini in toni di grigio e quindi, al contrario dei modelli di Ising e di Potts, con un numero grande di livelli per x_i. Nell'algoritmo proposto peró, Higdon usa le variabili ausiliarie per immagini in toni di grigio solo per formare i clusters. Il successivo aggiornamento di $\mathbf{x}$ avviene invece secondo Metropolis-Hastings.

In questo lavoro sono stati studiati algoritmi MCMC che utilizzano variabili ausiliarie in modo effettivo seguendo il secondo approccio di cui sopra. In particolare é stato implementato l'algoritmo proposto da Higdon, in cui le variabili ausiliarie si utilizzano solo per formare i clusters, e sono stati proposti qui per la prima volta degli algoritmi per il metodo del disaccoppiamento parziale che utilizzano in modo effettivo le variabili ausiliarie. Inoltre é stata studiata l'influenza, sulla velocitá di convergenza dell'algoritmo di Metropolis-Hastings, della distribuzione iniziale e della distribuzione di proposta. Le prestazioni dei diversi algoritmi sono state quantificate sulla base di simulazioni al calcolatore per la riduzione del rumore di due tipi di immagini: l'immagine del cervello umano e l'immagine della simulazione del cervello del topo.

Le conclusioni di questo studio sono le seguenti: una prima conclusione riguarda gli algoritmi MCMC con variabili ausiliarie secondo lo schema proposto da Higdon. Implementando gli algoritmi con un effettivo uso delle variabili ausiliarie, i vantaggi in termini di velocitá di convergenza sono stati modesti. Questo é vero per tutti e due i tipi di immagini.

Il fatto che un vantaggio apprezzabile avvenga nel caso dell'immagine simulata del cervello del topo, é legato al maggior significato che hanno i clusters rispetto all'immagine del cervello umano. Anche delle nuove proposte per il metodo del disaccoppiamento parziale non hanno purtroppo prodotto vantaggi significativi.

Per quanto riguarda invece lo studio sull'algoritmo di Metropolis-Hastings, sono stati riscontrati vantaggi significativi in termini di velocitá di convergenza nel caso di distribuzione di proposta indipendente e non simmetrico uguale alla distribuzione del rumore. Questo é vero per tutti i tipi di immagini utilizzate, per tutte le distribuzioni iniziali scelte. Naturalmente queste considerazioni dipendono dal rapporto segnale-rumore delle immagini. Per valori tipici del rapporto segnale-rumore e del parametro di smoothing confrontabili con quelli utilizzati in questo lavoro, la funzione di proposta secondo il modello del rumore Gaussiano é piú efficente ad esempio del random-walk Gaussiano ottimale. Infatti questa scelta ha prodotto un aumento della probabilitá di accettazione per Metropolis significativo.

Bibliografia

[1] G.Sebastiani ,1997, Mathematical and Statistical Methods for Medical Magnetic Resonance Imaging, *The Norvegian University of Science and Tecnology Trondheim,Norvay, Dr.Philos Thesis*.

[2] S.H.Heywang-Kobrunner R.Beck ,1996, *Contrast-Enhancement MRI of the breast* Springer.

[3] R.Highnam, M.Brady ,1999, *Mammographic Image Analysis* Kluver Academic Publishers .

[4] P.Baldi ,1998, *Calcolo delle probabilità e statistica* McGraw-Hill Libri Italia srl.

[5] G.Winkler ,1995, *Image Analysis Random Fields and Dynamic Monte Carlo Methods* Springer .

[6] D. Geman, S. Geman ,1984, Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, *Trans. Patt. Anal.Machine Intell.*, **6**, 721-741.

[7] S. Geman, D. Mc Clure ,1987, Stastistical methods for tomographic image reconstructions *Int. Stat. Inst. Bullettin* **52** 4-20.

[8] N. Metropolis, A.W.Rosenbluth, M.N.Rosenbluth, A.H.Teller, E.Teller ,1953, Equations of state calcolution by fast computing machines *J. Chem. Phys.*, **21** 1087-1091.

[9] D. Higdon, 1998, Auxiliary Variable Methods for Markov Chain Monte Carlo whit Application, *American Statistical Association*, **93**, No 442, 585-595 Theory and Methods.

[10] R.M. Neal, 2000, Slice Sampling, *Department of Statistics, University of Toronto* No. 2005.

[11] W.R. Gilks, G.O. Roberts, 1996, Strategies for improving MCMC, in: W.R. Gilks, S. Richardson and D.J. Spiegelhalter, eds. em Markov Chain Monte Carlo in Practice, (Chapman and Hall: London), ch. 6,pp 89-114.

[12] G.O. Roberts, S.K. Sahu,1997, Updating schemes, correlation structure, blocking and parametrization for the Gibbs sampler,*Journal of the Royal Statistical Society*, Series B **59**, 291-317.

[13] P. Barone, A. Frigessi,1990, Improving stochastic relaxation for Gaussian random fields, *Probability in the Engineering and Informational Sciences* , **4**, 369-389.

[14] P. Barone, G. Sebastiani, J. Stander, Over-relaxation methods a Markov chains for Monte Carlo simulation, *Accepted for publication Statistics and Computing.*

[15] P. Barone, G. Sebastiani, J. Stander,2001, General over-relaxation methods Monte Carlo algorithms for Gaussian densities, simulation, *Statistics and Probability Letters* **52**, 115-124.

[16] E. Marinari, G. Parisi,1992, Simulated tempering: a new Monte Carlo scheme, *Europhysics Letters*, **19** ,451-458.