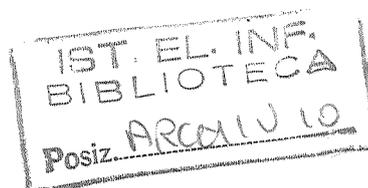


Consiglio Nazionale delle Ricerche

**ISTITUTO DI ELABORAZIONE
DELLA INFORMAZIONE**

PISA



L'uso di reti neurali per la
ricostruzione di immagini in
presenza di discontinuità

L. Bedini, A. Ribolini, E.
Salerno, A. Tonazzini

Nota Interna B4-53

dicembre 1994

L'USO DI RETI NEURALI PER LA RICOSTRUZIONE DI IMMAGINI IN PRESENZA DI DISCONTINUITA'

L. Bedini, A. Ribolini, E. Salerno e A. Tonazzini

Istituto di Elaborazione della Informazione

Consiglio Nazionale delle Ricerche

Via S. Maria, 46

56126 PISA

Abstract

In this paper, the edge-preserving reconstruction of piecewise smooth images is formulated within a probabilistic approach, where the image is modelled as two coupled Markov Random Fields, one representing the intensity and the other the discontinuities. In this way, prior knowledge about global and local features of the image can be introduced in the form of a Gibbs prior and a Maximum A Posteriori estimate can be assumed as the solution of the original problem. Since the posterior probability often presents local maxima, very expensive stochastic relaxation algorithms should be used to obtain the global maximizer. In order to reduce execution times, we propose a mixed-annealing algorithm, which obtains the solution iteratively by means of a sequence in which deterministic steps alternate with probabilistic ones. In particular, each deterministic step is designed to minimize a quadratic function of the intensity only, while each probabilistic step updates the discontinuities, using the Gibbs sampler. This algorithm can

exploit the computational power of neural networks in solving optimization problems; indeed, it can be implemented on a hybrid architecture made up of a grid of digital processors interacting with a linear neural network which supports most of the computational costs. Such a computational scheme can be adopted both for image restoration and for image reconstruction from projections, where the large size of the neighbourhoods would prevent the use of the Gibbs sampler.

1. Introduzione

La ricostruzione digitale di immagini, a partire da un insieme di dati osservati, è un problema inverso mal-posto, in quanto l'esistenza, l'unicità e la stabilità della soluzione non possono essere garantite. Soluzioni uniche e robuste al rumore si possono ottenere imponendo vincoli addizionali che sfruttino informazioni a priori sulle caratteristiche note dell'immagine e sulle proprietà statistiche del rumore presente nei dati. La teoria della regolarizzazione [Tik 77] consente di riformulare il problema della ricostruzione in termini di principi variazionali che incorporano vincoli derivati dalle conoscenze a priori. Svitati problemi dell'elaborazione di immagini possono essere riformulati in questo modo [Ber 88], [Pog 85]; il vincolo imposto è usualmente quello di "smoothness", cioè di lenta variabilità spaziale dell'intensità dell'immagine. Tale vincolo, tuttavia, non risulta adeguato in presenza di discontinuità legate a bruschi cambiamenti della geometria delle superfici o a variazioni repentine di intensità luminosa. Poiché le discontinuità hanno un ruolo importante nell'analisi e nel riconoscimento di immagini di scene reali, risulta fondamentale poterle ricostruire correttamente. A questo scopo, la teoria della regolarizzazione standard di Tikhonov presenta seri limiti poiché impone vincoli di "smoothness" anche dove non devono essere soddisfatti, e cioè in corrispondenza delle discontinuità [Bed 92a].

Recentemente, alcuni autori [Gem 84], [Bla 87a], [Bla 87b] hanno proposto approcci che consentono di spezzare i vincoli di "smoothness" in corrispondenza delle discontinuità, considerate anch'esse incognite (elementi di linea) del problema. In particolare, D. Geman e S. Geman sostengono il vantaggio di usare un approccio Bayesiano in cui il

comportamento locale dell'immagine è modellato per mezzo di Campi di Markov (MRF), in quanto particolarmente flessibile per l'introduzione di informazioni a priori sulle configurazioni degli elementi di linea. La soluzione viene ricondotta alla minimizzazione di un funzionale costo in cui compaiono esplicitamente sia le intensità che le discontinuità.

Nel presente lavoro, l'approccio Bayesiano sopra descritto viene proposto per la formulazione unificata dei problemi del restauro di immagini distorte e rumorose e della ricostruzione di immagini da un numero limitato di proiezioni. Con riferimento alla classe di immagini localmente "smooth", l'immagine è modellata come una coppia di MRF, uno per il processo di intensità e l'altro per il processo di linea. In tal modo le informazioni a priori sulla geometria delle discontinuità e sulle interdipendenze locali fra i valori di intensità possono essere descritte sotto forma di distribuzioni di Gibbs e la soluzione può essere definita come il punto di minimo globale di una funzione energia associata alla distribuzione di probabilità a posteriori.

Nel caso in cui siano utilizzate esplicitamente le discontinuità, la funzione energia a posteriori è in genere non convessa, e i classici algoritmi deterministici non assicurano il raggiungimento del minimo globale. In [Gem 84], [Bla 87a], e [Gei 91] sono stati proposti algoritmi sia stocastici che deterministici, in grado di sfuggire ai minimi locali, e quindi adatti a risolvere problemi di minimizzazione non convessa. La maggior parte degli algoritmi stocastici sono basati su tecniche di "annealing" simulato e generalmente presentano costi computazionali più elevati rispetto a quelli deterministici [Bla 89]. D'altro canto, la convergenza degli algoritmi stocastici è stata dimostrata in generale [Gem 84], [Aar 89], mentre la convergenza degli algoritmi deterministici è stata dimostrata solo per casi particolari [Bla 87a]. A causa della loro

intrinseca complessità e della grande dimensione dei tipici problemi di elaborazione di immagini, il costo computazionale di entrambi questi tipi di algoritmo tende ad essere molto alto. Per questa ragione sono stati fatti molti sforzi volti all'individuazione di algoritmi paralleli e nuove architetture parallele capaci di ridurre significativamente il tempo di esecuzione rispetto alle implementazioni seriali. In [Gem 84] è stata proposta l'applicazione di uno schema di "annealing" ad un algoritmo di rilassamento stocastico, il campionatore di Gibbs, che potrebbe essere implementato in parallelo su una griglia di processori interconnessi localmente. Tuttavia sono state considerate sperimentalmente solo immagini quantizzate su pochi livelli di grigio poichè all'aumentare dei livelli la complessità computazionale cresce notevolmente. Nel restauro di immagini, la possibilità di utilizzare il campionatore di Gibbs è legata principalmente al fatto che la distribuzione a posteriori è ancora Gibbsiana, con intorni relativamente piccoli. La sua applicazione al problema della ricostruzione di immagini da proiezioni non è invece possibile, poiché la probabilità a posteriori in questo caso non è più Gibbsiana. Infatti, a causa della relazione esistente fra l'immagine e i dati, il sistema di intorni associato alla distribuzione a posteriori è molto grande.

In [Bed 92b] è stato proposto un algoritmo che risolve il problema del restauro di immagini con discontinuità applicando, all'interno di uno schema iterativo di "annealing", un campionatore di Gibbs per l'aggiornamento del processo di linea e una minimizzazione deterministica per l'aggiornamento del processo di intensità. Lo schema globale dell'algoritmo, che porta alla minimizzazione della funzione energia a posteriori, può quindi essere considerato come risultante dalla cooperazione di due blocchi computazionali: il primo, fissata la

configurazione degli elementi di linea, trova il minimo rispetto al processo di intensità; il secondo aggiorna il processo di linea per mezzo di un campionatore di Gibbs, utilizzando la stima corrente del processo di intensità. Il primo blocco esegue una minimizzazione convessa e può essere efficientemente implementato per mezzo di una rete neurale lineare. Il campionatore di Gibbs può essere implementato su una griglia di processori che lavorano in parallelo. L'intrinseco parallelismo del campionatore di Gibbs, combinato con la potenza computazionale delle reti neurali nel risolvere problemi di ottimizzazione, consente quindi di ottenere un algoritmo molto efficiente.

Questo algoritmo di "annealing" misto viene qui applicato anche al caso della ricostruzione di immagini da proiezioni. L'aspetto interessante di questa applicazione risiede nel fatto che l'algoritmo consente di utilizzare il campionatore di Gibbs anche nel caso della ricostruzione da proiezioni, in quanto il sistema di interni relativo al processo di linea nella distribuzione a posteriori rimane lo stesso di quello adottato per la distribuzione a priori. Si vedrà che le differenze fra l'applicazione al restauro e quella alla ricostruzione consistono nel diverso grado di interconnessione della rete neurale che esegue la minimizzazione rispetto all'intensità.

2. Modelli a Campi di Markov nella Ricostruzione Bayesiana di Immagini

Per restauro digitale di immagini si intende il processo di stima dell'immagine ideale di un oggetto a partire da una sua versione degradata attraverso un operatore noto ed affetta da rumore usualmente additivo [And 77]. La ricostruzione di immagini da proiezioni (tomografia) è un procedimento mediante il quale si possono

ottenere mappe bidimensionali di una qualche proprietà fisica di un corpo, sfruttando la conoscenza dei suoi integrali di linea lungo un numero finito di percorsi [Her 80]. Il campo di applicazione forse più noto della tomografia è quello della radiologia diagnostica, in cui la proprietà fisica di interesse è la densità in una prefissata sezione del corpo in esame. In questo caso, misurando le perturbazioni indotte su alcuni tipi di radiazioni esploranti (tipicamente raggi X), si ottengono delle quantità legate agli integrali di linea della densità sulla sezione [Kak 79]. Mentre il problema della ricostruzione tomografica è stato risolto soddisfacentemente nel caso in cui si disponga di un numero elevato di proiezioni, esso rimane aperto nei casi in cui le proiezioni a disposizione siano limitate o mancanti su un certo intervallo angolare, o pesantemente affette da rumore [Cen 83]. Tali situazioni sono piuttosto frequenti in ambiente industriale e geologico, ma anche nel contesto medico, quando, ad esempio, si voglia diminuire la quantità di radiazioni nocive da somministrare al paziente.

Assumendo che il modello di formazione dei dati sia lineare, i problemi del restauro di immagini e della ricostruzione tomografica possono essere entrambi ricondotti alla risoluzione di sistemi di equazioni lineari. In applicazioni pratiche, tali sistemi risultano spesso indeterminati e sempre altamente mal-condizionati, per cui il calcolo di una soluzione soddisfacente richiede l'adozione di tecniche di regolarizzazione. Gli stabilizzatori principalmente usati per questo scopo, ad esempio le misure di entropia e di energia sulla soluzione o sulle sue derivate di vario ordine, producono soluzioni che risultano eccessivamente piatte. In tali situazioni, incorporare nel problema conoscenze riguardo le discontinuità può migliorare le ricostruzioni, consentendo la rottura del vincolo di "smoothness" nei punti in cui non

ha significato fisico. Tuttavia, la localizzazione preliminare delle discontinuità è alquanto difficoltosa se non addirittura impossibile, come quando i dati a disposizione sono valori di integrali di linea. Nella maggior parte dei casi è così preferibile trattare anche le discontinuità come incognite del problema.

Nell'approccio Bayesiano qui adottato l'immagine è considerata come una coppia (F,L) di Campi di Markov, dove F (processo di intensità) è la matrice $N \times N$ di variabili casuali a valori continui $f_{i,j}$, associate ai valori dei pixel, e L (processo di linea) è una matrice di variabili casuali binarie associate alle discontinuità. Tipicamente, gli elementi di linea sono considerati localizzati sulla griglia dei siti interpixel, e sono distinti in orizzontali e verticali, come mostrato in Fig.1. Un elemento di linea assume valore 1 per indicare la presenza di una discontinuità, 0 altrimenti. Nel seguito ogni particolare realizzazione del campo L sarà rappresentata dalla coppia (h,v), dove h e v sono le matrici $(N-1) \times N$ e $N \times (N-1)$ associate agli elementi di linea orizzontali e verticali, rispettivamente.

Indicando con f la notazione lessicografica per la matrice f , realizzazione del processo F, il modello di formazione dei dati è [And 77]:

$$\mathbf{g} = \mathbf{H} \mathbf{f} + \mathbf{n} \quad (2.1)$$

dove \mathbf{g} è il vettore dei dati, \mathbf{n} è una realizzazione del processo di rumore, indipendente da f ed l , e \mathbf{H} rappresenta la matrice di degradazione. Nel caso del restauro, \mathbf{H} rappresenta un operatore di convoluzione, mentre, nel caso della ricostruzione, \mathbf{H} è la *matrice di Radon*, il cui elemento (i,j) rappresenta la lunghezza dell'intersezione del percorso i -esimo con il pixel j -esimo [Her 80].

Il problema della ricostruzione è ricondotto alla determinazione di una realizzazione (f,l) che sia ottima in qualche senso, avendo a disposizione una realizzazione dei dati. La soluzione si può ottenere esprimendo le informazioni a priori sul problema in termini di opportune distribuzioni di probabilità e scegliendo come stima ottima quella che massimizza la probabilità a posteriori per (f,l) dato g (stima MAP). Per il teorema di Bayes, tale probabilità a posteriori è data da:

$$P(f,l | g) = \frac{P(g | f,l) P(f,l)}{P(g)} \quad (2.2)$$

Nell'usuale assunzione di un processo di rumore le cui componenti sono indipendenti, bianche e gaussiane, con varianza σ^2 nota e media nulla si ha:

$$P(g | f,l) = (2\pi \sigma^2)^{-M/2} \exp \left[-\frac{1}{2\sigma^2} \|g - Hf\|^2 \right] \quad (2.3)$$

dove M è la dimensione del vettore g . Ignorando il termine costante $P(g)$ e considerando la (2.3), la (2.2) diventa:

$$P(f,l | g) = (2\pi \sigma^2)^{-M/2} \exp \left[-\frac{1}{2\sigma^2} \|g - Hf\|^2 \right] P(f,l) \quad (2.4)$$

Il ruolo chiave è così giocato dall'espressione scelta per la distribuzione a priori $P(f,l)$ e cioè, nel nostro caso, dal modello MRF adottato.

Poiché la probabilità condizionale di un elemento di un MRF rispetto a tutti gli altri dipende solo dalla configurazione degli elementi associati a siti in un intorno prestabilito, il campo è completamente determinato specificando tutte le probabilità condizionali. Questo dà una estrema flessibilità nel modellare le informazioni a priori circa il comportamento locale del campo stesso. Inoltre, il teorema di Clifford-

Hammersley [Gem 84] stabilisce che la probabilità congiunta di un MRF può sempre essere espressa nella forma di una distribuzione di Gibbs:

$$P(f,l) = \frac{1}{Z} \exp \left(\frac{-U(f,l)}{\beta} \right)$$

$$U(f,l) = \sum_c V_c(f,l) \quad (2.5)$$

dove Z è la costante di normalizzazione, β è una costante positiva, $U(f,l)$ è la funzione energia e i potenziali $V_c(f,l)$ sono funzioni con supporto sulle "clique" associate al sistema di intorni. Un sottoinsieme c di siti è chiamato "clique" se due siti distinti in c stanno uno nell'intorno dell'altro. L'unione di tutte le "clique" che contengono uno stesso sito è l'intorno di tale sito. Dall'eq. (2.5) è così immediato verificare che la probabilità condizionale di un elemento del campo in un sito s dipende solo dagli elementi associati ai siti dell'intorno di s . Questo importante risultato evidenzia che il comportamento locale del campo può essere direttamente modellato attraverso la probabilità congiunta, specificando i potenziali piuttosto che le singole probabilità condizionali. Un compito essenziale a questo scopo è la definizione di un appropriato sistema di intorni. In questo lavoro, per modellare la classe di immagini localmente "smooth" con discontinuità connesse, si è utilizzato il sistema di intorni mostrato in Fig. 2. Con riferimento a questo sistema, la dipendenza fra intensità dei pixel e configurazioni degli elementi di linea, e la mutua dipendenza fra elementi di linea, sono state espresse quantitativamente per mezzo dei potenziali $V_c(f,l)$ associati al sistema di "clique". In particolare, sono stati associati potenziali diversi da zero solo a "clique" costituite da due e quattro elementi di linea (Figg. 3a, 3b e 3c) e a "clique" miste costituite da tre elementi (Figg. 3d e 3e). Le possibili

configurazioni, a meno di rotazioni, per le "clique" di linea sono mostrate in Fig.4.

Una tipica scelta per $U(f,l)$ è la seguente [Mar 87]:

$$U(f,l) = \lambda \sum_{i=1}^N \sum_{j=1}^{N-1} [(f_{i,j} - f_{i,j+1})^2 - \theta^2] (1 - v_{i,j}) + \quad (2.6 a)$$

$$+ \lambda \sum_{i=1}^{N-1} \sum_{j=1}^N [(f_{i,j} - f_{i+1,j})^2 - \theta^2] (1 - h_{i,j}) + \quad (2.6 b)$$

$$+ \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} V_a(h_{i,j}, v_{i,j}, h_{i,j+1}, v_{i+1,j}) + V_b(v_{i,j}, v_{i,j+1}) + V_b(h_{i,j}, h_{i+1,j}) \quad (2.6 c)$$

in cui le (2.6 a) e (2.6 b) rappresentano le somme dei potenziali relativi alle "clique" miste di Figg. 3e e 3d, e la (2.6 c) rappresenta le somme dei potenziali sulle "clique" di linea di Figg. 3a, 3b e 3c. La costante θ rappresenta una soglia sui gradienti orizzontali e verticali, e λ è un parametro di regolarizzazione positivo. La presenza dei termini (2.6 a) e (2.6 b) nella probabilità a priori dell'immagine, in accordo alla (2.5), rende più probabili soluzioni in cui gli elementi di linea hanno valore 1 dove il gradiente è alto. Nella (2.6 c), i potenziali V_a e V_b sono funzioni di tutte le possibili configurazioni delle "clique" di linea a quattro e a due elementi rispettivamente; tali funzioni sono specificabili per mezzo di tabelle. Un basso valore del potenziale rende la corrispondente configurazione più probabile e viceversa. In immagini reali, le discontinuità sono usualmente connesse e sottili. Questo vincolo può essere imposto associando un potenziale alto alle configurazioni in cui un elemento di linea si presenta isolato, e alle configurazioni in cui due elementi di linea adiacenti e paralleli sono contemporaneamente attivi.

La soluzione MAP al problema della ricostruzione di immagini si ottiene massimizzando la probabilità a posteriori di eq. (2.4) che, tenendo conto di eq. (2.5) diventa:

$$P(f,l | g) = \frac{1}{Z} (2\pi\sigma^2)^{-M/2} \exp \left(- \frac{\|g - Hf\|^2}{2\sigma^2} - U(f,l) \right) \quad (2.7)$$

dove $U(f,l)$ è data dalla (2.6) e β è stato posto uguale a 1. Lo stesso scopo può essere equivalentemente raggiunto minimizzando la funzione energia a posteriori definita come:

$$E_g(f,l) = \frac{\|g - Hf\|^2}{2\sigma^2} + U(f,l) \quad (2.8)$$

E' da notarsi che, a causa della sua dipendenza dal processo di linea, la funzione $E_g(f,l)$ risulta non convessa.

3. L'algoritmo di "annealing" misto

Per la soluzione generale di problemi di minimizzazione non convessa, come quello qui trattato, sono stati proposti algoritmi iterativi, basati sul rilassamento stocastico e l'"annealing" simulato. In questi algoritmi sono consentiti cambi casuali nello stato delle variabili, in accordo ad una legge di probabilità che dipende da un parametro di controllo τ detto temperatura [Aar 83]. Nel caso presente, la legge di probabilità si ottiene introducendo il parametro τ in (2.7) nel seguente modo:

$$P(f,l|g) = \frac{1}{Z} \exp \left(\frac{-E_g(f,l)}{\tau} \right) \quad (3.1)$$

La temperatura viene fatta decrescere gradualmente in accordo ad un opportuno schema di "annealing" e, per ogni suo valore, si calcola uno stato delle variabili che rappresenta un campione tratto dalla

probabilità a posteriori di eq. (3.1). Nell'algoritmo di Metropolis [Met 53], questo campione è ottenuto generando un gran numero di transizioni di stato casuali. E' noto che, al tendere a zero della temperatura, questa procedura converge allo stato più probabile in accordo alla (2.7), cioè al minimo globale dell'energia (2.8). In [Kir 83] viene usato l'algoritmo di Metropolis per raggiungere uno stato stabile a ogni temperatura di uno schema di "annealing" derivato ad hoc per analogia con l'"annealing" chimico. In [Gem 84] viene proposto un campionatore di Gibbs per raggiungere stati stabili all'interno di uno schema in cui la temperatura è abbassata secondo una legge logaritmica.

Questi metodi hanno generalmente un costo computazionale molto elevato. La ragione principale è la lenta convergenza, che richiede spesso parecchie centinaia di iterazioni. Una seconda ragione è la dimensione molto grande del problema trattato; nel nostro caso, per esempio, il numero di variabili è $N^2+2N(N-1)$, dove un tipico valore per N è 512. In [Gem 84] è evidenziato che il campionatore di Gibbs potrebbe essere efficientemente implementato su sistemi di calcolo parallelo. Tuttavia questo richiederebbe un elevato numero di processori, uno per ogni variabile del problema. Inoltre deve essere notato che, per la minimizzazione dell'energia a posteriori che risulta nella ricostruzione da proiezioni, la dimensione molto elevata del sistema di intorni rende impraticabile l'uso del campionatore di Gibbs. Infatti, per la presenza della matrice di Radon come operatore di distorsione, l'intorno di ogni pixel diventa virtualmente l'intera immagine.

In questo lavoro è suggerito uno schema di "annealing" misto in cui le variabili continue sono aggiornate in modo deterministico, mentre quelle binarie sono aggiornate in modo stocastico. Il procedimento è

basato sull'osservazione che l'energia a posteriori di eq. (2.8) è una funzione convessa della sola variabile f e così il suo minimo $f^*(l)$ può essere calcolato in modo deterministico. In accordo a questa considerazione, il problema di minimizzare $E_g(f,l)$ rispetto sia ad f che ad l può essere ricondotto al problema della minimizzazione di $E_g(f^*(l),l)$ rispetto alla sola l . Poiché la $E_g(f^*(l),l)$ è ancora Gibbsiana con lo stesso sistema di intorno della distribuzione a priori, la sua minimizzazione può essere eseguita mediante un campionatore di Gibbs accoppiato con uno schema di "annealing", la cui efficienza è ora molto migliorata a causa dell'enorme riduzione del numero di possibili configurazioni. Si noti infatti che le variabili interessate dalla minimizzazione sono solamente le $2N(N-1)$ componenti binarie di l . Il Campionatore di Gibbs può essere usato per trarre campioni da una data distribuzione di Gibbs, in modo parallelo, aggiornando ripetutamente tutti gli elementi di linea simultaneamente ed in modo asincrono. Ogni elemento di linea l_s associato al sito s è aggiornato in accordo alla probabilità condizionale locale, dati i valori osservati degli elementi di linea nei siti r , $r \neq s$, $r \in A_s$, dove A_s è l'intorno di s . I passi dell'algoritmo di "annealing" misto sono i seguenti:

1. Si scelgono un valore iniziale τ_0 per la temperatura τ e una legge di decremento; si fissano il numero L_0 di transizioni da eseguire alla temperatura τ_0 e una configurazione iniziale l_0 per gli elementi di linea.
2. Dati τ_k e L_k , iniziando dalla configurazione precedente l_k , il campionatore di Gibbs trae un campione l_{k+1} del processo di linea dalla seguente distribuzione di Gibbs:

$$P(l_{k+1}|g) = \frac{1}{Z} \exp \left(\frac{-E_g(f^*(l_{k+1}), l_{k+1})}{\tau_k} \right)$$

3. Si selezionano nuovi valori di τ_{k+1} e L_{k+1} e si ripete l'algoritmo dal passo 2, finché non sia soddisfatto un opportuno criterio di terminazione.

E' stato provato [Aar 89] che, se le sequenze $\{\tau_k\}$ e $\{L_k\}$ sono scelte convenientemente, algoritmi come quello descritto sopra sono in grado di sfuggire a minimi locali e quindi di convergere al minimo globale.

Poiché la dimensione di A_s è generalmente modesta e la variabile l_s è binaria, il costo computazionale per generare un campione del processo di linea in accordo alla probabilità condizionale locale è normalmente basso. Tuttavia, per ogni possibile configurazione del processo di linea, la probabilità condizionale locale dipende dal valore di $f^*(l)$, risultante dalla minimizzazione dell'energia a posteriori rispetto a f . Anche se la dimensione di questo problema di minimo è ridotta da $N^2+2N(N-1)$ a N^2 , il suo costo computazionale rimane ancora molto elevato, e tale da impedire l'implementazione dell'algoritmo sopra descritto su calcolatore seriale. Nel paragrafo seguente si mostrerà che è possibile progettare una rete neurale che prende come ingresso la configurazione corrente degli elementi di linea l e dà come uscita il valore $f^*(l)$.

Di seguito forniamo una versione approssimata dell'algoritmo di "annealing" misto, che, oltre a rimanere adatta ad una implementazione facente uso di reti neurali, presenta una complessità ridotta compatibile con una implementazione su calcolatore seriale. Detto l'_k l'aggiornamento corrente generato dal campionatore di Gibbs nel passo 2 partendo dalla configurazione iniziale l_k , si possono ragionevolmente adottare le seguenti approssimazioni:

- a. l'_k differisce da l_k solo per pochi elementi;

b. per ogni l'_k vale $E_g(f^*(l'_k), l'_k) \approx E_g(f^*(l_k), l'_k)$.

Il passo 2 dell'algoritmo può così essere diviso nel seguente modo:

2a. dato l_k , si calcola $f^*(l_k)$;

2b. dati τ_k e L_k , iniziando dalla configurazione precedente l_k , il campionatore di Gibbs trae un campione l_{k+1} del processo di linea dalla seguente distribuzione di Gibbs:

$$P(l_{k+1}|g) = \frac{1}{Z} \exp\left(\frac{-E_g(f^*(l_k), l_{k+1})}{\tau_k}\right)$$

L'approssimazione descritta sopra aumenta l'efficienza dell'algoritmo in quanto il calcolo di $f^*(l)$ viene eseguito non per ogni aggiornamento del singolo elemento di linea, ma solo dopo aver raggiunto uno stato stazionario per l'intero processo di linea. L'algoritmo esegue in successione una minimizzazione di E_g rispetto ad f , fissato l_k , quindi, mantenendo costante f , aggiorna l_k generandone un campione in accordo alla sua probabilità condizionale. Per valori bassi della temperatura l'aggiornamento di l_k viene praticamente eseguito deterministicamente in modo da abbassare il valore dell'energia, quindi si avrà almeno la convergenza ad un minimo locale. D'altra parte, specialmente ad alte temperature, la possibilità di un aumento del valore dell'energia può consentire di scavalcare i minimi locali; non è dunque escluso che l'algoritmo possa convergere anche al minimo globale. I risultati di simulazione, riportati parzialmente nel paragrafo 5, hanno evidenziato che ricostruzioni di buona qualità possono essere ottenute con un numero di iterazioni dei passi 2a e 2b dell'ordine della decina, e con un numero di aggiornamenti completi degli elementi di linea dell'ordine delle centinaia, per ogni esecuzione del passo 2b.

4. Schema computazionale basato su una rete neurale

Ogni singola iterazione dell'algoritmo di minimizzazione proposto nel paragrafo precedente prevede due passi principali: l'aggiornamento degli elementi di linea e la minimizzazione dell'energia a posteriori rispetto a f .

Gli elementi di linea sono aggiornati usando il campionatore di Gibbs: tutti i siti di linea sono visitati ripetutamente e i loro valori sono ottenuti traendo campioni dalla distribuzione condizionale locale. Un tale schema può essere implementato nella pratica usando un'architettura parallela composta da una griglia di processori, ognuno dei quali è relativo ad un sito di linea ed è connesso solo ai processori relativi ai siti del suo intorno [Gem 84]. Mentre il numero di siti e di processori distinti può essere molto grande, il numero di connessioni a un dato processore è usualmente basso, in dipendenza dall'ordine del sistema di intorni scelto.

La minimizzazione dell'energia a posteriori rispetto agli elementi di intensità trarrebbe anch'essa vantaggio dall'uso di architetture parallele. Recentemente, è stato evidenziato che alcune reti neurali artificiali, per la cui pratica implementazione sono stati proposti modelli elettrici analogici, rappresentano uno strumento potente per risolvere problemi di ottimizzazione. Poggio e altri [Pog 85] hanno sviluppato un modello di rete analogica resistiva per la minimizzazione delle funzioni energia quadratiche che derivano dalla regolarizzazione standard di tipici problemi della Computer Vision. In [Bed 92a], il modello di rete neurale non lineare e continuo proposto da Hopfield [Hop 82], [Hop 85], [Hop 86] è applicato al restauro di immagini con il metodo della massima entropia. Le reti di Hopfield possono essere adattate anche per risolvere problemi di minimizzazione quadratica,

assumendo che tutti gli elementi abbiano caratteristica lineare. Nel seguito, per la minimizzazione dell'energia a posteriori rispetto all'intensità, faremo riferimento alla rete di Hopfield già illustrata in [Bed 92a]. Tale rete può essere realizzata mediante un circuito elettrico composto di amplificatori O_k , con resistenze e capacità di ingresso R_k e C_k , rispettivamente, interconnessi tra loro mediante una rete caratterizzata da una matrice di transconduttanze T . La dinamica di tale circuito elettrico è descritta dal seguente sistema di equazioni differenziali:

$$C_k \frac{du_k}{dt} = \sum_{i=1}^N T_{ki} v_i - \frac{u_k}{R_k} + I_k$$

$$v_k = \varphi_k(u_k) \quad k=1,2,\dots,N \quad (4.1)$$

dove φ_k esprime la caratteristica ingresso-uscita dell'amplificatore O_k , e la transconduttanza T_{ki} interconnette l'uscita dell'amplificatore O_k e l'ingresso dell'amplificatore O_i . I_k è la corrente iniettata dall'esterno sull'ingresso di O_k . E' possibile mostrare che, sotto opportune condizioni [Hop 84], la funzione di Liapunov associata al circuito elettrico descritto dalle (4.1) è la seguente:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} v_i v_j + \sum_{i=1}^N \frac{1}{R_i} \int_0^{v_i} \varphi_i^{-1}(v) dv - \sum_{i=1}^N I_i v_i \quad (4.2)$$

Nel seguito sarà mostrato che i parametri di questa rete possono essere scelti in modo che la funzione (4.2) prenda la forma della funzione energia a posteriori del problema qui trattato. Assumendo per tutti gli amplificatori un guadagno unitario, l'eq. (4.2), in forma matriciale, diventa:

Il grado di interconnessione della rete neurale dipende dalla forma assunta dalla matrice T . In particolare, come già detto in precedenza, nel caso del restauro si hanno degli intorno di dimensione ridotta e pertanto la rete risulterà caratterizzata da connessioni solo locali; nel caso della ricostruzione da proiezioni, la struttura della matrice H è tale che l'intorno di ogni pixel nell'energia a posteriori diventa praticamente l'intera immagine e quindi la rete neurale risulterà totalmente interconnessa.

5. Risultati sperimentali

Le prestazioni dell'algoritmo descritto nei paragrafi precedenti sono state valutate su un computer seriale, utilizzando immagini sia sintetiche che reali. Dato il costo computazionale dell'algoritmo, ne è stata implementata solo la versione approssimata fornita nel paragrafo 4. Inoltre, come ulteriori semplificazioni, sono state trattate immagini di dimensione massima 64×64 , e, nello schema di "annealing", sono state eseguite solo poche iterazioni, mantenendo la temperatura ad un valore fissato sufficientemente basso.

In tutte le prove effettuate, il valore del parametro di regolarizzazione λ è stato scelto empiricamente in modo che le immagini restaurate presentassero un buon compromesso fra la proprietà di "smoothness" e la consistenza con i dati. Il valore di soglia θ dovrebbe essere scelto sulla base dell'effettivo valore minimo del gradiente, corrispondente alle discontinuità tipiche della classe di immagini trattate. In generale, un tale valore non è disponibile per immagini reali; una stima approssimata può tuttavia essere ottenuta sulla base di informazioni a priori o, nel caso del restauro, di una preliminare analisi delle immagini degradate. Particolare attenzione è stata dedicata alla

scelta dei potenziali di linea V_a e V_b di eq. (2.6 c). Questi potenziali determinano la probabilità condizionale di un elemento di linea data la configurazione degli elementi nel suo intorno. Per le immagini considerate, la scelta di V_a e V_b è stata fatta in modo tale da scoraggiare la formazione di configurazioni poco probabili come incroci di linee, terminazioni di linee e linee doppie e, viceversa, incoraggiare le continuazioni di linee. I valori qui adottati sono riportati nella Tabella I.

Nel seguito sono riportati alcuni dei risultati ottenuti applicando l'algoritmo al restauro di immagini distorte e affette da rumore e alla ricostruzione di immagini da un numero limitato di proiezioni.

5.1 Restauro di Immagini

In questo caso, i dati sono costituiti da immagini degradate, ottenute mediante convoluzione delle immagini originali con un filtro passa-basso, definito da una maschera uniforme, e contaminate con rumore gaussiano additivo di media zero e varianza σ^2 . L'immagine sintetica considerata è mostrata in Fig. 6a, mentre la corrispondente immagine degradata con una maschera uniforme 3×3 e rumore additivo con SNR pari a 20dB è mostrata in Fig. 6b. L'immagine restaurata (Fig. 6c) è stata ottenuta dopo 7 iterazioni usando come parametri $\lambda=0.003$, $\theta=20$ e $\tau=0.5$; per ogni iterazione è stata eseguita una minimizzazione rispetto a f , ed è stato tratto un campione del processo di linea visitando consecutivamente tutti gli elementi di linea 400 volte. In Fig. 6d è riportata la mappa finale degli elementi di linea.

Come secondo esempio riportiamo un'immagine reale (Fig. 7a), degradata artificialmente con un filtro uniforme di dimensione 5×5 e aggiungendo rumore con SNR pari a 20dB. L'immagine degradata è mostrata in Fig. 7b. Per gli stessi valori dei parametri usati nel caso

sintetico, e dopo 8 iterazioni, sono stati ottenuti i risultati mostrati in Fig. 7c; per ogni iterazione sono stati effettuati 400 aggiornamenti completi degli elementi di linea. In Fig. 7d è riportata la mappa della configurazione finale degli elementi di linea.

5.2 Ricostruzione di immagini da proiezioni

Per questo caso sono riportati i risultati relativi alla ricostruzione di un'immagine sintetica ispirata ad una versione circolare del fantoccio di testa umana proposto da Shepp e Logan [Kak 83] (Fig. 8a). I valori del processo di intensità sono stati scelti nell'intervallo $[0, 0.4]$ e corrispondono ai valori di attenuazione incontrati nel cranio da un fascio di raggi X di 60 keV [Her 80]. I dati a disposizione sono stati calcolati come valori di proiezioni uniformemente distribuite in $[0, \pi]$, ognuna relativa a 150 percorsi equispaziati. Considerando che nei casi pratici si riescono ad ottenere dati con un buon SNR, il rumore sui dati di proiezione è stato fissato a 40dB. In Fig. 8b è mostrata la ricostruzione ottenuta con 8 iterazioni complete dell'algoritmo; i valori dei parametri utilizzati sono stati $2\lambda\sigma^2=3$, $\theta=0.048$, $\tau=0.7$. Nelle Figg. 8c e 8d sono riportati, per confronto, i risultati ottenuti con i metodi di ricostruzione ART (Algebraic Reconstruction Technique) [Cen 83] e CBP (Convolution Backprojection) [Lew 83], utilizzando lo stesso numero di dati. Il numero delle proiezioni è stato poi portato a 150, e sono state eseguite le ricostruzioni ART e CBP (Figg. 8e e 8f). Si può osservare che la qualità di queste ultime ricostruzioni è confrontabile con quella ottenuta, mediante il metodo proposto, utilizzando un numero di proiezioni ridotto di un fattore 5.

6. Conclusioni

E' stato proposto un algoritmo di "annealing" misto per la ricostruzione MAP di immagini localmente "smooth". L'immagine è stata modellata come una coppia di MRF, uno per il processo di intensità e l'altro per il processo di linea. La stima MAP è stata ottenuta come il minimo globale di una funzione energia a posteriori $E_g(f,l)$. Poichè l'energia è una funzione convessa di f per ogni configurazione fissata di elementi di linea, essa presenta un unico minimo $f^*(l)$. Il minimo globale appartiene così all'insieme di coppie (f,l) tale che $f=f^*(l)$ e può essere raggiunto minimizzando la funzione $E_g(f^*(l),l)$ rispetto a l . Questa minimizzazione è eseguita attraverso un algoritmo di "annealing" simulato che aggiorna solo il processo di linea, mentre il processo di intensità è vincolato ad assumere il valore $f^*(l)$. Il metodo proposto può essere utilizzato per risolvere sia il problema del restauro che quello della ricostruzione da proiezioni, in cui la dimensione degli intorni rende impraticabile l'uso diretto del campionatore di Gibbs. E' stata anche proposta un'implementazione dell'algoritmo su una architettura parallela ibrida, costituita da una griglia di processori digitali interagenti con una rete neurale analogica. La griglia di processori digitali aggiorna il processo di linea usando un campionatore di Gibbs. La rete neurale consiste di un circuito analogico lineare che, per ogni fissata configurazione l degli elementi di linea, raggiunge il valore $f^*(l)$ come suo stato di equilibrio. E' stata proposta inoltre una versione approssimata dell'algoritmo, basata sull'assunzione che l'energia a posteriori vari poco per pochi aggiornamenti degli elementi di linea.

Le prestazioni di questo algoritmo approssimato sono state analizzate per mezzo di un computer seriale su immagini sia sintetiche che reali, con applicazione al restauro di immagini e alla ricostruzione

tomografica da proiezioni. I risultati ottenuti si sono rivelati soddisfacenti anche con un numero di iterazioni inferiore a 10.

BIBLIOGRAFIA

- Aar 89** Aarts E., Korst J. *Simulated Annealing and Boltzmann Machines*, John Wiley & Sons, 1989, New York.
- And 77** Andrews H. C., Hunt B. R. *Digital Image Restoration*, Prentice-Hall, 1977, Englewood Cliffs.
- Bed 92a** Bedini L., Tonazzini A. "Reti di Hopfield per il restauro di immagini", *Rivista di Informatica*, Vol XXII, No 1, 1992, pp. 29-41.
- Bed 92b** Bedini L., Tonazzini A. "Image restoration preserving discontinuities: the Bayesian approach and neural networks", *Image and Vision Comput.*, Vol 10, No 2, 1992, pp. 108-118.
- Ber 88** Bertero M., Poggio T., Torre V. "Ill-posed problems in early vision", *Proc. IEEE*, Vol 76, No 8, 1988, pp. 869-889.
- Bla 87a** Blake A., Zisserman A. *Visual Reconstruction*, MIT Press, 1987, Cambridge, Massachusetts.
- Bla 87b** Blake A., Zisserman A. "Localising discontinuities using weak continuity constraints", *Pattern Recogn. Letters*, Vol 6, No 1, 1987, pp. 51-59.
- Bla 89** Blake A. "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 11, No 1, 1989, pp. 2-12.
- Cen 83** Censor Y. "Finite Series-expansion reconstruction methods", *Proc. IEEE*, Vol 71, No 3, 1983, pp. 409-419.

- Gei 91** Geiger D., Girosi F. "Parallel and deterministic algorithms for MRFs: surface reconstruction", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 13, No 5, 1991, pp. 401-412.
- Gem 84** Geman S., Geman D. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 6, No 6, 1984, pp. 721-741.
- Her 80** Herman G. T. *Image Reconstruction from Projections: the Fundamentals of Computerized Tomography*, Academic Press, 1980, London.
- Hop 82** Hopfield J. J. "Neural networks and physical systems with emergent collective computational abilities", *Proc. Natl. Acad. Sci. USA*, Vol 79, 1982, pp. 2554-2558.
- Hop 84** Hopfield J. J. "Neurons with graded response have collective computational properties like those of two-state neurons", *Proc. Natl. Acad. Sci. USA*, Vol 81, 1984, pp. 3088-3092.
- Hop 85** Hopfield J. J. "Neural computation of decisions in optimization problems", *Biol. Cybern.*, Vol 52, 1985, pp. 141-152.
- Hop 86** Hopfield J. J., Tank D. W. "Computing with neural circuits: a model", *Science*, Vol 233, 1986, pp. 625-633.
- Kak 79** Kak A. C. "Computerized tomography with X-ray, emission and ultrasound sources", *Proc. IEEE*, Vol 67, No 9, 1979, pp. 1245-1272.

- Kak 83** Kak A. C., Pan S. X. "A computational study of reconstruction algorithms for diffraction tomography: interpolation versus filtered backprojection", *IEEE Trans. Acoust. Speech Signal Processing*, Vol 31, No 5, 1983, pp. 1262-1275.
- Kir 83** Kirkpatrick S., Gellatt C. D., Vecchi M. P. "Optimisation by simulated annealing", *Science*, Vol 220, No 4598, 1983, pp. 671-680.
- LaS 61** La Salle J., Lefschetz S. *Stability by Liapunov's Direct Method with Applications*, Academic Press, 1961, London.
- Lew 83** Lewitt R. M. "Reconstruction algorithms: transform methods", *Proc. IEEE* , Vol 71, No 3, 1983, pp. 390-408.
- Mar 87** Marroquin J., Mitter S., Poggio T. "Probabilistic solution of ill-posed problems in computational vision", *J. Am. Stat. Ass.*, Vol 82, No 397, 1987, pp. 76-89.
- Met 53** Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller E. "Equations of state calculations by fast computing machines", *J. Chem. Phys.*, Vol 21, 1953, pp. 1087-1091.
- Pog 85** Poggio T., Torre V., Koch C. "Computational vision and regularization theory", *Nature* , Vol 317, 1985, pp. 314-319.
- Tik 77** Tikhonov A. N., Arsenin V. Y. *Solution of Ill-posed Problems*, Winston-Wiley, 1977, Washington.

DIDASCALIE DELLE FIGURE

Fig.1 Griglie dei pixel e degli elementi di linea.

Fig.2 Sistema di intorni:

- (a) intorno per un elemento di linea orizzontale;
- (b) intorno per un elemento di linea verticale;
- (c) intorno per un pixel.

Fig.3 "clique" con valori di potenziale non nulli:

- (a) "clique" di linea a quattro elementi;
- (b) "clique" di linea a due elementi (verticale);
- (c) "clique" di linea a due elementi (orizzontale);
- (d) "clique" mista orizzontale;
- (e) "clique" mista verticale.

Fig.4 Possibili configurazioni (a meno di rotazioni) per le "clique" di linea:

- (a) nessuna linea; (b) terminazione; (c) angolo;
- (d) continuazione; (e) configurazione a "T"; (f) croce;
- (g) nessuna linea; (h) linea semplice; (i) linea doppia.

Fig.5 Il sistema computazionale.

Fig.6 Restauro di un'immagine sintetica (64×64):

- (a) immagine originale; (b) immagine degradata (filtro uniforme 3×3, SNR=20dB); (c) immagine restaurata dopo 7 iterazioni; (d) mappa degli elementi di linea rilevati.

Fig.7 Restauro di un'immagine reale (64×64):

(a) immagine originale; (b) immagine degradata (filtro uniforme 5×5, SNR=20dB); (c) immagine restaurata dopo 8 iterazioni; (d) mappa degli elementi di linea rilevati.

Fig.8 Ricostruzione di un'immagine sintetica (64×64):

(a) immagine originale; (b) ricostruzione Bayesiana dopo 8 iterazioni (30 proiezioni, SNR=40dB); (c) ricostruzione ART dopo 10 iterazioni (30 proiezioni, SNR=40dB); (d) ricostruzione CBP (30 proiezioni, SNR=40dB); (e) Ricostruzione ART dopo 10 iterazioni (150 proiezioni, SNR=40dB); (f) ricostruzione CBP (150 proiezioni, SNR=40dB).

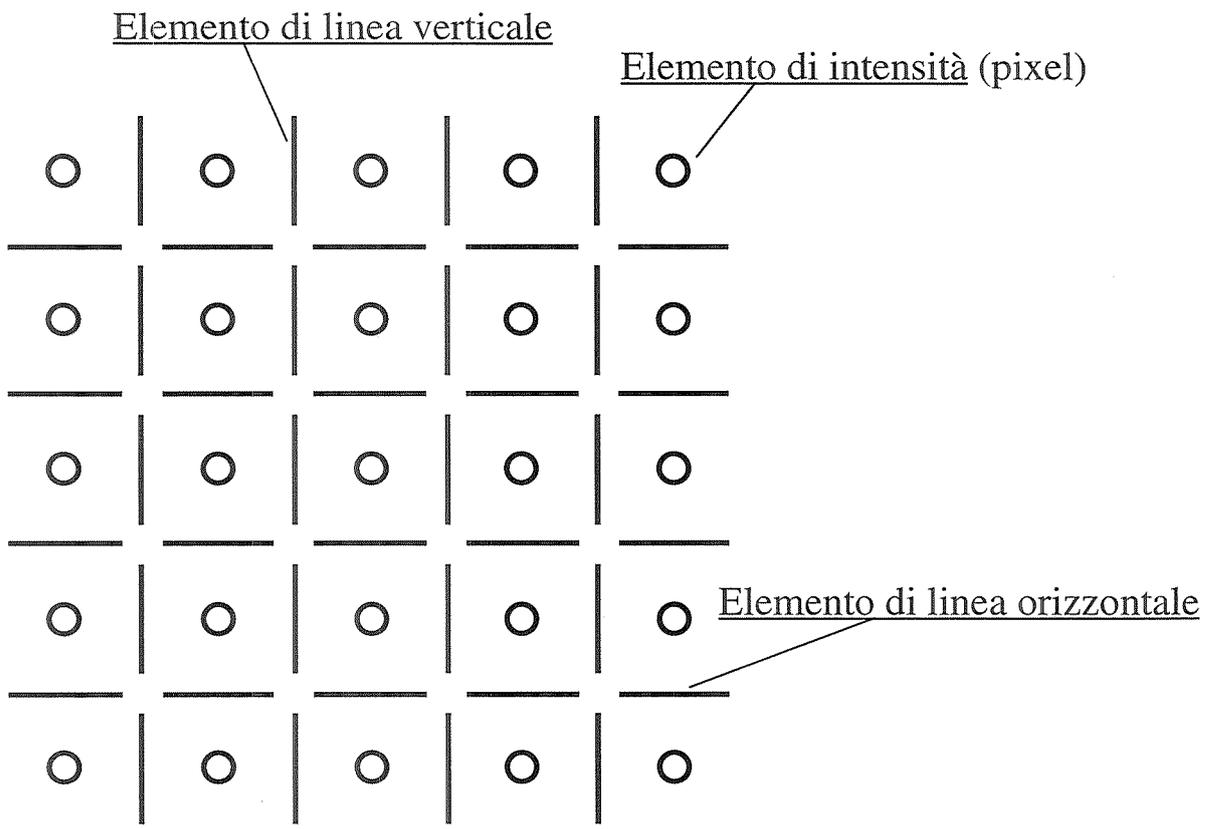
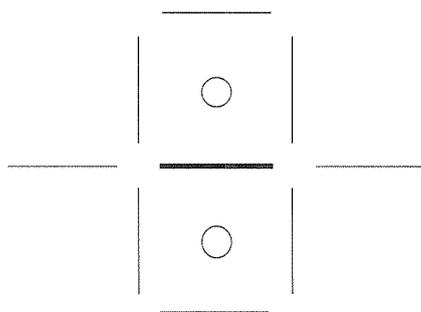
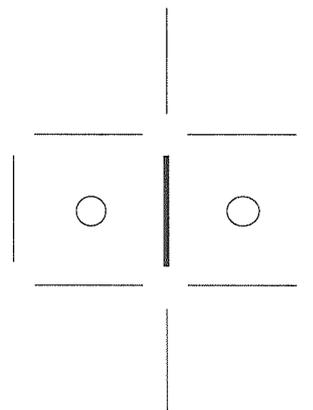


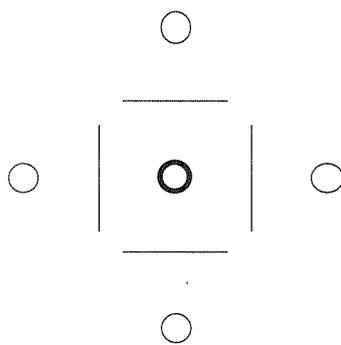
Figura 1



(a)

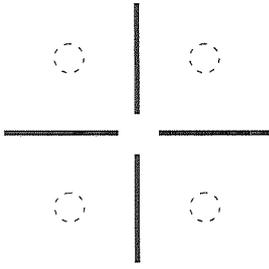


(b)



(c)

Figura 2



(a)



(b)



(c)

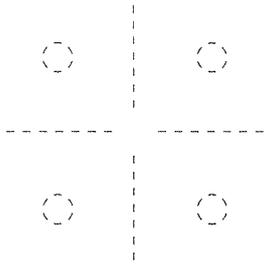


(d)

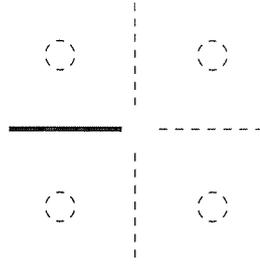


(e)

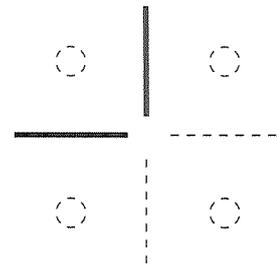
Figura 3



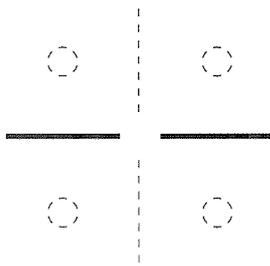
(a)



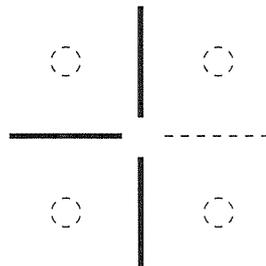
(b)



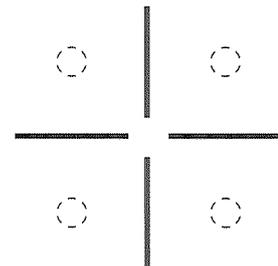
(c)



(d)



(e)



(f)



(g)



(h)



(i)

Figura 4

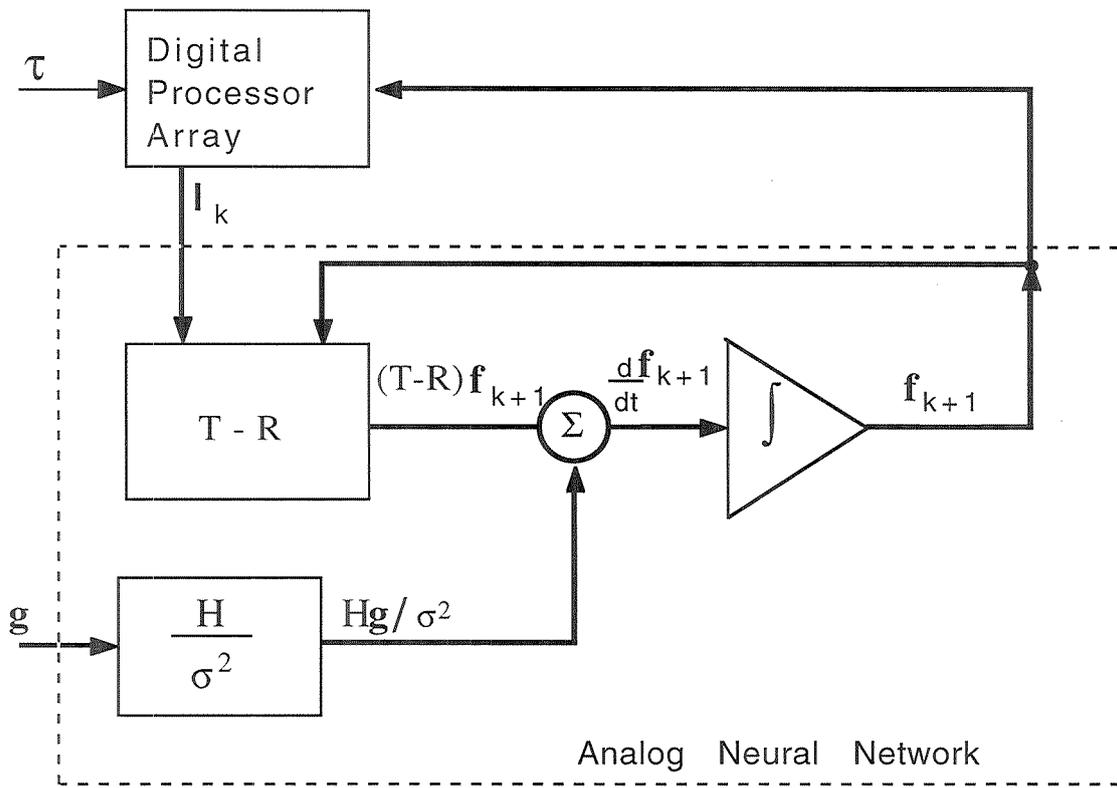


Figura 5

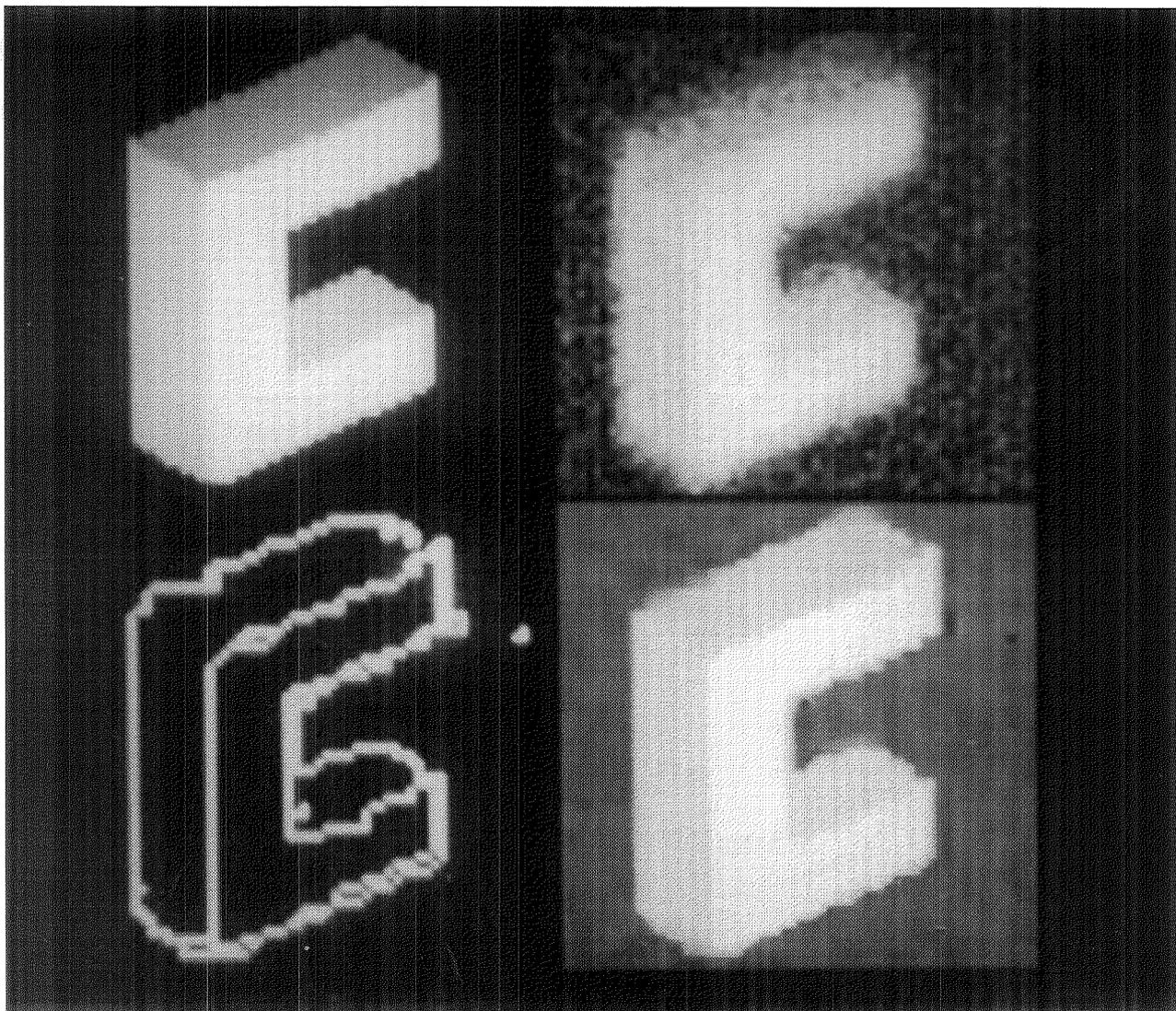


Figura 6



Figura 7

Configurazione di linea	Potenziale V Restauro	Potenziale V Tomografia
nessuna linea (a)	0	0
terminazione (b)	0.25	3
angolo (c)	0.1	0
continuazione (d)	0.05	0
"T" (e)	0.25	1.25
croce (f)	0.4	1.75
nessuna linea (g)	0	0
linea semplice (h)	0	0
linea doppia (i)	0	2.25

Tabella I - Potenziali V adottati per le clique di linea; le configurazioni sono identificate come in Fig.4.