

UNIVERSITA' DEGLI STUDI DI NAPOLI
FACOLTA' DI SCIENZE MATEMATICHE FISICHE E
NATURALI
Dottorato di Ricerca in *Matematica Applicata ed Informatica*

SEBASTIANO BATTIATO

Metodi Statistici ed Entropici per
l'Apprendimento e l'Analisi
Automatica di Immagini Digitali

TESI DI DOTTORATO DI RICERCA
XI CICLO

Tutore:
Prof. A. Maugeri

Coordinatore:
Prof. L. M. Ricciardi

Indice

Indice	2
Introduzione	5
Note Bibliografiche	9
1. Entropia e Distribuzioni di Gibbs nell'Elaborazione delle Immagini ...	11
- Introduzione;	
- L'Entropia di Shannon come Misura della Casualità;	
- Concetti di Base;	
- Il Principio della Massima Entropia (MaxEnt);	
- Campi <i>Random</i> di Gibbs;	
- <i>Markov Random Fields</i> (MRF) e Distribuzioni di Gibbs;	
- Il Campionamento di Gibbs;	
- Un Esempio;	
- Le Applicazioni all'Elaborazione delle Immagini;	
- Il Principio dell'Entropia MiniMax;	
- <i>Image Restoration</i> ;	
- <i>Threshold</i> ;	
2. Analisi Automatica di <i>Pattern/Texture</i> Visuali	42
- Concetti Preliminari;	
- <i>Texture e Computer Vision</i> ;	
- Analisi e Sintesi di Tessiture;	

- Il Modello FRAME e l'Insieme di Julesz;
- *Texture* e Multirisoluzione (1): Heeger e Bergen;
- *Texture* e Multirisoluzione (2): De Bonet;
 - Idea di Base;
 - Procedura di De Bonet: Analisi;
 - Procedura di De Bonet: Sintesi;
 - Analisi Computazionale ed Euristiche;
- *Hypertexture*;

3. Analisi e Sintesi di Tessiture: *Clustering* e Multirisoluzione 68

- Introduzione;
- Il *Clustering*;
- Il Nostro Approccio;
 - Analisi;
 - Sintesi;
- Esperimenti e Discussioni;
- Conclusioni e Futuri Sviluppi;

4. Metodi Entropici per la Costruzione di Mappe di Saliienza 81

- Introduzione;
- Costruzione di Mappe di Saliienza;
 - Fase di *Learning*;
 - Calcolo Effettivo della Mappa di Saliienza;
- Risultati Sperimentali;

5. *Watermarking* e Manipolazione dei Colori 94

- Introduzione;
- Nozioni di Base;
- Spazi di Colore;
- Gli Algoritmi;
 - L'Algoritmo di Inserimento;
 - L'Algoritmo di *Detection*;

- Analisi Statistica;
 - Identificazione di Immagini non Manipolate;
 - Identificazione di Immagini Manipolate;
 - Analisi Statistica nello Spazio;
- Risultati Sperimentali;
- Conclusioni e Sviluppi Futuri;

Appendice A:

Un Algoritmo per la Selezione della Mediana Approssimata 121

- L'Algoritmo;
- Analisi dell'Algoritmo: Complessità e *Performance*;
- Conclusioni;

Bibliografia 129

Introduzione

Oggetto di studio della presente tesi è l'applicazione di Metodi Statistici ed Entropici all'Elaborazione delle Immagini e in *Computer Vision*; il tema ed argomento centrale della Tesi è quindi legato a problematiche di visione e grafica che verranno dettagliate nel seguito. Si è ritenuto però utile prevedere di aggiungere, in forma di appendice al corpo centrale di risultati della tesi, una dettagliata relazione anche sui risultati conseguiti nell'ambito della Teoria degli Algoritmi Probabilistici sotto la guida del Prof. M. Hofri della WPI University e del Prof. D. Cantone dell'Università di Catania.

Nella Tesi vengono approfondite tematiche relative all'applicazione di metodi tipici della statistica teorico/applicativa all'Elaborazione delle Immagini. In particolare oggetto di studio e di analisi sono le distribuzioni di probabilità indotte da campionamenti o da una conoscenza *a priori* del fenomeno sotto osservazione, al fine di estrarre metodologie e tecniche di analisi da utilizzare ad esempio negli ambiti applicativi più classici quali Segmentazione, *Pattern Recognition*, Analisi/Sintesi di *texture*, *edge finding*, ecc.. Vengono inoltre brevemente presentati alcuni risultati ottenuti nell'ambito dell'*information hiding* di immagini digitali, tuttora oggetto di studio. Gli studi e le ricerche condotti in questi particolari ambiti applicativi sono stati realizzati insieme con il Prof. G. Gallo dell'Università di Catania.

Il lavoro di Tesi è organizzato secondo i seguenti capitoli:

- *Metodi Statistici ed Entropici Applicati all'Elaborazione delle Immagini.*
Vengono illustrate le potenzialità e gli innumerevoli campi di applicazione legate all'*Entropia dell'Informazione* applicata all'Elaborazione delle Immagini nei suoi vari aspetti e forme. La giusta evidenza viene data ai risultati più significativi presenti in letteratura grazie ad un breve *survey* che ne illustra l'evoluzione storica/scientifica. Particolare enfasi e merito vengono riconosciuti al *Principio della Max Entropia* di E.T. Jaynes e all'originario concetto di *Entropia* dovuto a C. Shannon che, inizialmente utilizzati in ambiti totalmente differenti quali quello della Termodinamica e della Teoria dell'Informazione, assumono oggi notevole rilevanza grazie alle numerose applicazioni in *Computer Graphics* (si notino ad esempio i lavori di D. Mumford della Stanford University);
- *Analisi Automatica di Pattern/Texture Visuali.* Anche in questo caso dopo una breve panoramica storico-scientifica volta a focalizzare gli obiettivi specifici e le problematiche di base del settore, vengono brevemente illustrati i principali risultati presenti in letteratura; in particolare si approfondiscono i risultati del gruppo di ricerca facente capo a J. De Bonet, *Learning & Vision Group* - Cambridge che ha condotto degli studi sull'Analisi/Sintesi di *texture* utilizzando un approccio multirisoluzione non parametrico utilizzato di recente anche nella discriminazione di immagini satellitari di tipo SAR.

Il materiale presente in questi primi capitoli rappresenta la base teorica e scientifica, il punto di partenza dal quale si sono poi condotti gli studi e le ricerche ulteriori che hanno portato ai risultati descritti nei capitoli successivi della Tesi.

- *Analisi e Sintesi di Tessiture.* In questo importante ambito teorico/applicativo sono in fase avanzata di studio e di implementazione alcune tecniche di sintesi che, utilizzando delle procedure di clusterizzazione ed un approccio multirisoluzione permettono di ottenere *pattern* generati dalla stessa legge stocastica di base. Inoltre la metodologia seguita ha permesso di realizzare significativi *improvement*, dal punto di vista computazionale rispetto ai metodi precedenti. Si prefigura inoltre la possibilità di applicare tali tecniche al problema della classificazione di tessiture e nello specifico ambito dell'analisi e sintesi di linguaggi naturali, dove in particolare è possibile attraverso opportuni algoritmi di apprendimento riuscire a selezionare il migliore insieme di *feature* necessario per riuscire a riprodurre le complesse leggi statistiche che fanno capo al fenomeno stesso;
- *Entropia e Mappe di Saliienza.* In questo capitolo vengono illustrati i risultati ottenuti insieme a G. Gallo, dove analizzando l'Entropia di Shannon delle distribuzioni di probabilità indotte dagli istogrammi di una serie ben specifica di *feature* relative ad un determinato *training set* di esempi positivi, si è costruita una mappa di salienza in grado di valutare la probabilità che un dato *pixel* in una data scena appartenesse o meno alla classe di oggetti in

questione. Le potenzialità effettive del metodo proposto vengono validate da un'esaustiva serie di esperimenti. Sono inoltre allo studio diverse estensioni a problematiche più complesse;

- *Watermarking*. Questo capitolo è interamente dedicato alla discussione di un nuovo algoritmo di *watermarking* su immagini digitali. Tale algoritmo, tuttora in fase di studio e di ulteriore estensione, permette l'inserimento di un marchio digitale, attraverso un'opportuna tecnica che manipola i *pixel* di un'immagine agendo nello spazio dei colori della *Color Opponency*. In particolare vengono individuate delle regioni, vere e proprie sfere, associate ai singoli *pixel* di un'immagine, dove è possibile eseguire degli opportuni spostamenti impercettibili. Il metodo proposto oltre ad essere robusto rispetto alle più comuni operazioni di *image processing* viene analizzato anche attraverso una rigorosa analisi statistico-teorica che ne evidenzia la resistenza rispetto ad alcuni attacchi intenzionali volti alla rimozione del marchio stesso. Anche in questo caso l'analisi e la descrizione dell'algoritmo è accompagnata da una serie di risultati sperimentali che ne fanno apprezzare la bontà dello stesso.

Appendice:

- Viene descritto un algoritmo per la ricerca della Mediana Approssimata la cui analisi probabilistica è stata realizzata insieme con il Prof. Micha Hofri (WPI University); l'algoritmo in questione, competitivo dal punto di vista della complessità computazionale, si presta bene per la realizzazione di un analogo filtro di *smoothing*.

Note Bibliografiche

Il materiale presente nel Capitolo 3 è largamente tratto da un articolo pubblicato sui *Proceedings di Workshop on Texture Analysis in Machine Vision, TEXTURE'99*, Oulu, Finland, pp.149-156, 1999. Analogamente per il contenuto specifico del Capitolo 4, interamente tratto da un lavoro apparso sui *Proceedings di European Congress on Intelligent Techniques and Soft Computing, EUFIT'98*, Aachen, Germany, Vol.2, pp.1375-1380, 1998. Entrambi i lavori sono stati realizzati insieme con Giovanni Gallo.

Il Capitolo 5 fa invece parte di un filone di ricerca condotto insieme con Dario Catalano e Giovanni Gallo dell'Università di Catania e con Rosario Gennaro dell'IBM T.J. Watson Research Center di New York, tuttora in fase di ulteriore sviluppo; alcuni dei risultati presentati sono stati presentati sui *Proceedings di 3-rd Workshop on Information Hiding*, Dresden 1999, altri invece appariranno sui *Proceedings di IST/SPIE Electronic Imaging 2000*.

L'Appendice A è in parte tratta da un lavoro realizzato con Dario Catalano, Gianluca Cincotti e Domenico Cantone dell'Università di Catania insieme con

Micha Hofri della WPI University. I risultati di tale ricerca sono stati presentati all'annuale *Meeting di Average-Case on the Analysis of Algorithms*, tenutosi a Barcellona, Spagna, nel Giugno 1999. Una versione estesa di tale articolo apparirà sui *Proceedings di CIAC 2000*, Roma.

Si ringraziano i coautori per aver acconsentito alla parziale riproduzione di tali lavori in questa Tesi.

*Non è perchè le cose sono
difficili che non osiamo,
ma è perchè non osiamo
che sono difficili.*
Seneca

Capitolo 1

Entropia e Distribuzioni di Gibbs nell'Elaborazione delle Immagini

In questo capitolo verranno sinteticamente illustrati i concetti di base nonché l'evoluzione storica delle linee di ricerca facenti riferimento alla cosiddetta *Information Entropy* e alle sue applicazioni nella risoluzione di problemi di inferenza statistica nell'elaborazione delle immagini. Saranno anche introdotti i dettagli matematici riguardanti i cosiddetti *Markov Random Field* (MRF) e il relativo campionamento di Gibbs (*Gibbs sampling*). In particolare, attraverso un *excursus* storico dei principali risultati presenti in letteratura, partendo dallo storico lavoro di C. Shannon del 1948 e proseguendo poi con Jaynes (1957) e con Geman & Geman (1984), verrà illustrato l'iter metodologico-scientifico che si è

affermato nel corso degli anni e che sta alla base di diversi risultati teorico-applicativi ottenuti nell'ambito dell'analisi/sintesi di tessiture, nella costruzione di mappe di salienza e nell'*image restoration*.

1. Introduzione

Dopo ormai quasi cinquant'anni dalla sua prima apparizione in letteratura, il concetto di *entropia* è oggi universalmente accettato: diversi risultati vengono continuamente tirati fuori e proposti negli ambiti applicativi più svariati ma facenti comunque capo all'*antica* idea di misura dell'informazione. Come già accennato il concetto di *entropia* fu introdotto nel 1948 da Claude Shannon, in uno storico articolo che è alla base di tutta la moderna *Teoria dell'Informazione*. Non meno importante fu il contributo di E.T. Jaynes, che nel 1957, introducendo il *Principio della Massima Entropia* influenzò in maniera decisa, in un certo senso rivoluzionandolo del tutto, l'approccio fino ad allora seguito nella risoluzione di problematiche di inferenza statistica. L'*entropia* dell'informazione, come noto, misura l'incertezza di una distribuzione di probabilità; l'evidenza sperimentale ne ha messo in luce nel corso degli anni l'utilità pratica nella risoluzione dei cosiddetti problemi *ill-conditioned*, dove cioè la mancanza di informazioni e/o di dati specifici riguardo al problema in oggetto, richiede l'assunzione di ipotesi quanto più generali possibili in riferimento ai parametri del sistema oggetto di studio. Ecco spiegato il motivo per cui anche nell'Elaborazione delle Immagini, l'*entropia* è risultata essere utile, come vedremo, in svariate applicazioni il cui obiettivo specifico è la costruzione di modelli di probabilità basati su assunzioni di tipo statistico riguardanti in particolare la riproduzione stocastica di opportune

caratteristiche (o *features*) di immagini digitali (siano esse legate a scene reali o artificiali).

Nel 1984 S.Geman e D.Geman pubblicarono un articolo per certi versi *storico*, dove attraverso la rielaborazione di differenti concetti teorici già noti nel campo della Fisica Statistica, introdussero nuove tecniche e metodologie di lavoro strettamente legate con le distribuzioni di Gibbs e con i *Markov Random Field* (MRF). In particolare grazie a tali tecniche è oggi possibile stimare, attraverso un opportuno algoritmo di campionamento, detto *Gibbs Sampling*, la legge di distribuzione sottesa all'immagine oggetto di studio anche se il relativo modello probabilistico è troppo complesso da manipolare analiticamente. Seguendo la naturale evoluzione delle tecniche e dei metodi applicativi che si sono succeduti nel corso degli anni e descrivendone le idee di base dei differenti filoni di ricerca ci si propone quindi di analizzare il percorso culturale-scientifico che partendo da C. Shannon ha trovato in Jaynes prima e più tardi in Geman e Geman, i *padri fondatori* dei cosiddetti metodi entropici nell'Elaborazione delle Immagini. In particolare si cercherà di evidenziare come tali risultati ottenuti in ambiti spesso totalmente differenti e con obiettivi specifici anch'essi completamente indipendenti siano invece inevitabilmente interconnessi tra loro e come tali rappresentano la base scientifica delle moderne metodologie di ricerca oggi più in voga.

2. L'Entropia di Shannon come Misura della Casualità

Il concetto di *entropia* dell'Informazione fu introdotto per la prima volta nel 1948 (Shannon - 1948) ma in ogni caso non si trattava di un'idea completamente

nuova: un analogo concetto era già conosciuto in termodinamica e in meccanica statistica. In particolare Clausius e Boltzmann diedero la prima espressione funzionale per l'*entropia* come misura del grado di disordine esistente in un sistema termodinamico (Zemansky - 1981).

A Shannon va però il merito di aver introdotto in maniera formale e precisa il concetto di *entropia* a seguito dei suoi studi legati al problema della comunicazione di segnali su un canale in presenza di rumore. In particolare i suoi studi si svilupparono rapidamente dando vita a due distinte linee di ricerca:

- La **Teoria dell'Informazione** (*Information Theory*): una teoria probabilistica volta a studiare appunto le caratteristiche statistiche dei sistemi di comunicazione;
- La **Teoria della Codifica** (*Coding Theory*): un insieme di metodi algebrici utili a generare codifiche efficienti per diversi problemi di comunicazione. Un noto algoritmo è per esempio quello di Huffman dove l'*entropia* fornisce una stima della lunghezza media dell'ottimale codifica iniettiva (Goldie e Pinch - 1991; Welsh - 1988).

Ma le applicazioni dell'universale concetto di *entropia* sono innumerevoli e hanno a che fare con gli ambiti applicativi più svariati. Per esempio è possibile applicare l'*entropia* allo studio di criptosistemi: l'*entropia* condizionale è utilizzata per misurare quanta informazione è racchiusa in un qualsiasi criptotesto in relazione alla chiave segreta che lo ha generato (Brassard e Bratley - 1988). Altri campi di applicazione sono l'economia, la biologia, la sociologia, lo studio dei sistemi dinamici, la teoria degli algoritmi, oltre che la teoria dell'informazione e

della codifica e l'inferenza statistica. Per maggiori dettagli si veda l'esaustivo sito Web (Hillmann - 1998).

2.1 Concetti di base

Ancora oggi l'originale articolo di Shannon è per certi versi la migliore introduzione al moderno concetto di *entropia*. Rimandiamo quindi a (Shannon - 1948) per la descrizione analitica delle problematiche teorico-matematiche, limitandoci ad elencare in maniera sintetica alcuni concetti di base.

Supponiamo di avere una variabile casuale X definita su un insieme di eventi mutuamente esclusivi $X=\{x_1, x_2, \dots, x_n\}$ in accordo con la distribuzione di probabilità $p(X)$ e tale che $\sum_i p(x_i) = 1$.

Se consideriamo questa variabile casuale come una sorgente di informazione, diventa lecito cercare di misurarne in maniera consistente il grado di incertezza associato ad ogni singolo evento. In questi casi infatti è disponibile solo la cosiddetta distribuzione di probabilità *a priori*, che se da un lato consente di tirare fuori comunque degli indici statistici di notevole rilevanza statistica (Valor medio, Scarto Quadratico Medio, ecc.) dall'altro invece non permette di stimare in termini probabilistici la casualità associata al fenomeno in oggetto nella sua globalità. Per stimare in maniera numerica il grado di incertezza con un data funzione $H(X)$, quest'ultima deve soddisfare le seguenti proprietà di base:

1. La funzione $H(X)$ deve esistere: deve essere cioè possibile associare un legame di tipo numerico tra l'incertezza di una distribuzione di probabilità ed i numeri reali;

2. $H(X)$ ha una dipendenza funzionale continua su $p(X)$: ciò significa che a piccole variazioni della $p(X)$ corrispondono inevitabilmente altrettante piccole variazioni nella misura di incertezza associata;
3. Se $p(x_i) = 1/n$, $i=1, 2, \dots, n$ allora $H(X)$ non può che essere una funzione monotona crescente in n . In altre parole al crescere del numero di eventi, se questi sono equiprobabili, l'incertezza associata è maggiore;
4. $H(X)$ gode della proprietà additiva: se esiste più di un modo per calcolarne il suo valore questo deve comunque essere lo stesso in ogni caso. Se per esempio un evento può essere visto come l'unione di due o più distribuzioni di probabilità il valore della funzione H per così dire globale deve essere uguale alla media pesata dei valori assunti dalla H nei singoli casi.

E' possibile dimostrare la validità del seguente teorema dovuto a Shannon:

Teorema 2.1 *L'unica funzione H che soddisfa le proprietà 1-4 di cui sopra, è della forma:*

$$H(X) = -k \sum_{i=1}^n p(x_i) \log p(x_i) \quad (1)$$

dove k è una costante positiva (solitamente posta ad 1).

La funzione H è detta *entropia*, *entropia di Shannon* o meglio *entropia dell'informazione*. Ciò per distinguerla dall'analoga espressione funzionale legata all'entropia di un sistema termodinamico; l'*entropia dell'informazione* è una proprietà associata ad una qualunque distribuzione di probabilità mentre la

cosiddetta entropia sperimentale usata in termodinamica è invece una proprietà legata a vere e proprie quantità fisiche (volume, pressione, temperatura, ecc.). Di seguito con il termine *entropia* ci riferiremo all'*entropia dell'informazione*.

Quando nell'espressione (1) $p(x_i)=0$ per qualche i , il relativo termine all'interno della sommatoria è ovviamente non definito. Dato però che $\lim_{x \rightarrow 0^+} x \log x = 0$ si può assumere in maniera naturale che il relativo termine sia nullo o equivalentemente considerare solo i termini diversi da zero. E' possibile generalizzare la funzione *entropia* definendola sulle cosiddette sorgenti di informazione continue. In questo caso la distribuzione della variabile casuale X è espressa in termini di densità di probabilità (*probability density function* o *p.d.f.*) supposta continua per semplicità di notazione.

L'espressione funzionale per la funzione *entropia* diventa in questo caso:

$$H(X) = -k \int_{-\infty}^{+\infty} p(x) \log p(x) dx \quad (2)$$

La funzione *entropia* gode di alcune interessanti proprietà:

1. $H(X) \geq 0$ ed inoltre $H(X)=0$ se e solo se $p(x_i)=1$ per qualche i ed ovviamente $p(x_j)=0$ $\forall j \neq i$. Si tratta cioè di un'effettiva e concreta misura dell'incertezza nascosta nei dati a disposizione;
2. $H(X)$ assume il suo valore massimo pari a $\log n$ quando $p(x_i) = 1/n$, $\forall i = 1, 2, \dots, n$. Ciò sta ad indicare che la massima incertezza si presenta quando tutti gli eventi sono equiprobabili;

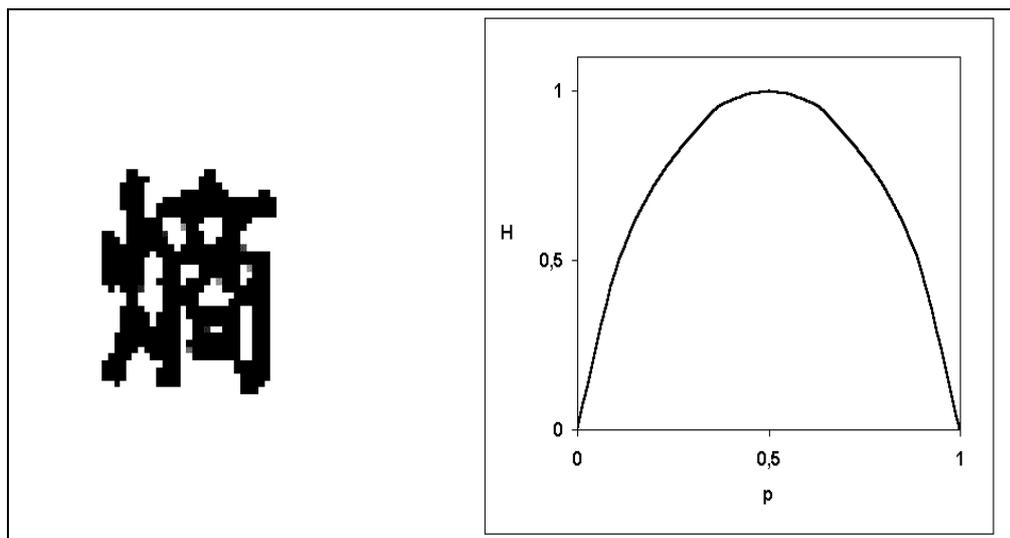


Fig. 1: a) Ideogramma cinese per l'entropia b) L'entropia di due eventi con probabilità p e $(1-p)$ descritta al variare di p .

3. Se X e Y sono due variabili casuali allora $H(X,Y) \neq H(X) + H(Y)$ dove $H(X,Y)$ denota l'entropia associata alla distribuzione di probabilità congiunta di X e Y . L'eguaglianza sussiste se e solo se X e Y sono tra di loro indipendenti;
4. Se X e Y sono due variabili casuali è possibile considerare la distribuzione di probabilità condizionale data da (X/Y) ; in questo caso si definisce entropia condizionale $H(X/Y)$ la funzione:

$$H(X|Y) = - \sum_x \sum_y p(y) p(x|y) \log p(x|y) \quad (3)$$

L'entropia condizionale è quindi la media pesata con i valori $p(y)$ dell'entropie $H(X/y)$ su tutti i possibili y .

5. $H(X,Y) = H(X/Y) + H(Y)$. Si noti che se X e Y sono tra loro indipendenti allora $H(X/Y) = H(X)$.

La dimostrazione di tali semplici proprietà può essere trovata in (Shannon 1948; Stinson 1995).

3. Il Principio della Massima Entropia (*MaxEnt*)

Il percorso storico-scientifico volto ad ottenere degli strumenti teorico-applicativi nell'applicazione dell'*entropia* all'elaborazione di segnali digitali passa attraverso questa ulteriore tappa: il ***Principio della Massima Entropia*** spesso denotato brevemente con ***MaxEnt***. Tale principio venne introdotto nel 1957 da E.T. Jaynes in un articolo centrato proprio sull'inferenza statistica. In particolare Jaynes riuscì a riformulare risultati di meccanica statistica precedentemente ottenuti da J. Willard Gibbs, in termini di distribuzioni di probabilità, utilizzando appunto questa nuova metodologia di lavoro che oltre a semplificare notevolmente la notazione e i dettagli matematici ne permise un'ulteriore estensione. Fu infatti possibile studiare problemi di meccanica statistica come particolari istanze di problemi inferenziali basati sulla parziale conoscenza dei dati reali (*ill-conditioned problems*). Il **Principio della Massima Entropia** è enunciato nella seguente definizione:

Definizione 3.1 *In presenza di dati e/o evidenze sperimentali riguardanti un ben determinato fenomeno fisico-statistico per stimarne la relativa distribuzione di probabilità è sufficiente scegliere un modello che sia consistente con i dati disponibili ma che altrove abbia la massima entropia;*

Come è evidente dalla stessa definizione, tale principio è applicabile nei problemi inferenziali o di *learning*, legati a problemi basati su dati non completamente noti a priori o viziati da errori di qualche tipo. Ci volle comunque, qualche anno prima che la comunità scientifica ne apprezzasse a pieno la potenza e soprattutto cominciasse a tirarne fuori dei risultati in differenti aree di ricerca (Jaynes - 1990). A testimonianza del fatto che oggi invece la situazione è ben diversa, ogni anno viene organizzato un meeting internazionale dedicato al principio *MaxEnt* in cui vengono presentate le più recenti applicazioni di tale metodologia in svariate aree di ricerca: matematica, fisica, ingegneria, linguaggi naturali, elaborazione delle immagini, ecc.

Massimizzare l'*entropia* di una distribuzione di probabilità, significa appunto generalizzarne il comportamento statistico su tutto ciò che non è conosciuto a priori. Si massimizza cioè l'incertezza assunta sui parametri del sistema non noti ottenendo dei modelli basati solo su dati certi. Non si sta facendo altro che tradurre in termini matematici un antico principio di saggezza: *Nunquam ponenda est pluralitas sine necessitate*, (Occam's Razor). Analizziamone brevemente il legame con in *Principio di Indifferenza di Laplace* che afferma che a due o più eventi debba essere assegnata eguale probabilità a meno che non ci sia un valido motivo per fare altrimenti. Consideriamo allora il seguente problema di ottimizzazione:

$$Max (H (X)) = Max \left(- \sum_{i=1}^n p_i \log p_i \right) \quad (4)$$

con:

$$\sum_{i=1}^n p_i = 1 \quad (5)$$

Applicando il metodo dei moltiplicatori di Lagrange la soluzione di tale problema è proprio $p_i=1/n$, $i=1,2,\dots,n$. In altre parole il principio **MaxEnt** a differenza di altri noti modelli predittivi (per esempio le reti neurali) che inavvertitamente fanno delle assunzioni fittizie circa i dati, si basa esattamente ed esclusivamente sui dati disponibili. In particolare oggi è possibile, grazie all'ausilio di computer sempre più potenti e veloci, riuscire ad applicare il principio **MaxEnt** anche a problemi aventi un gran numero di dati (*high dimension problems*), con un significativo miglioramento rispetto alle tecniche finora utilizzate.

Passiamo ad analizzare adesso i dettagli matematici che stanno alla base dell'applicazione specifica del principio **MaxEnt**.

Sia X una variabile casuale definita su $\{x_1, x_2, \dots, x_n\}$ di cui si vuole stimare la distribuzione di probabilità. Supponiamo inoltre di avere a disposizione m distinte funzioni di x , dette *feature*: $f_k(X)$, $k=1, 2, \dots, m$. Di solito le *feature* sono strettamente legate allo specifico problema d'inferenza e misurano appunto le distribuzioni marginali del modello di probabilità che si vuole esplicitare. Tali funzioni vengono utilizzate per vincolare i valori attesi che il modello assegnerebbe alle corrispondenti *feature*:

$$F_k = \sum_{i=1}^n p(x_i) f_k(x_i) \quad k = 1, \dots, m \quad (6)$$

dove F_k è la distribuzione empirica di X direttamente calcolata sui dati di *training*. L'obiettivo è quindi quello di costruire un modello statistico p della variabile casuale X che sia generato a partire da una dato insieme di *training*.

Applicando il principio *MaxEnt* (3.1), è sufficiente trovare la soluzione del seguente problema:

$$\text{Max} (H (X)) = \text{Max} \left(- \sum_{i=1}^n p_i \log p_i \right) \quad (7)$$

soggetto ai vincoli (5) e (6).

Tale problema può essere risolto, per esempio, utilizzando il metodo dei moltiplicatori di Lagrange noto dalla teoria dell'ottimizzazione vincolata, si veda (Jaynes - 1957) per maggiori dettagli. Per ciascuna *feature* f_i resta individuato un parametro I_i , un moltiplicatore di Lagrange appunto, tale che il problema è ricondotto al calcolo di un massimo libero, non più vincolato, della funzione Lagrangiana corrispondente:

$$\Lambda(p, \mathbf{I}) = H(p_1, p_2, \dots, p_n) + \sum_{k=1}^m I_k \left(\sum_{i=1}^n p_i \cdot f_k(x_i) - F_k \right) \quad (8)$$

la cui soluzione è data da:

$$p_{\Lambda}(x_i) = \frac{1}{Z(\Lambda)} e^{-\sum_{k=1}^m I_k f_k(x_i)} \quad i = 1, 2, \dots, n \quad (9)$$

e $Z(\Lambda)$ è una funzione costante di normalizzazione definita a partire dal vincolo (5):

$$Z(I_1, I_2, \dots, I_m) = \sum_{i=1}^n e^{-\sum_{k=1}^m I_k f_k(x_i)} \quad (10)$$

E' possibile dimostrare che il valore H_{MAX} trovato può anche essere espresso in termini dei valori F_1, F_2, \dots, F_m nella seguente maniera:

$$H_{MAX} = S(F_1, F_2, \dots, F_m) = \log Z(\Lambda) + \sum_{k=1}^m I_k F_k \quad (11)$$

E' possibile allora definire un'esplicita relazione per i I_k :

$$I_k = \frac{\partial S}{\partial F_k} \quad k=1, \dots, m \quad (12)$$

In realtà riuscire a determinare univocamente tali valori non è così facile a causa della sostanziale incalcolabilità della funzione di partizione Z . Questo problema viene risolto, come vedremo, attraverso l'utilizzo estensivo di opportune tecniche numeriche.

Uno dei parametri chiave di tale approccio è comunque legato alla scelta delle *feature* in qualche modo più significative, tali cioè da riuscire a racchiudere in sé i dettagli statistici più rilevanti del modello di probabilità che si vuole stimare. Come già accennato, infatti l'applicazione di tale metodologia richiede un grosso sforzo computazionale, spesso al di sopra delle reali disponibilità di calcolo; diventa allora fondamentale riuscire a selezionare un insieme di *feature* sufficientemente espressivo controllandone comunque la cardinalità (Berger, Della Pietra, Della Pietra - 1996; Della Pietra, Della Pietra, Lafferty - 1997; Jain, Zongker - 1995; Zhu, Mumford - 1998). Nella Sezione 5 verranno comunque discussi alcuni di questi approcci. Il principio **MaxEnt** trova applicazione oltre che nell'Elaborazione delle Immagini anche nell'Analisi Spettrale di segnali uni e

bidimensionali dove oggetto di studio è l'elaborazione di segnali radar e di segnali sismici o vocali (Wu - 1997).

4. Campi *Random* di Gibbs

I *Gibbs Random fields* sono particolari campi *random* strettamente correlati con il principio *MaxEnt* (3.1) e con la definizione di *Markov Random Fields*. In questo paragrafo ci si propone di richiamarne alcune nozioni di base, ricorrendo ove necessario ad alcuni risultati classici che ne evidenzino i legami specifici. L'espressione funzionale della funzione Lagrangiana (9) descritta nel paragrafo precedente, descrive una forma funzionale ben nota in fisica statistica e in termodinamica.

In uno storico articolo del 1984 S. Geman e D. Geman, introdussero la possibilità di applicare all'elaborazione delle immagini tecniche analoghe a quelle presenti in tali settori. In questo caso l'idea di base consiste nel trattare i livelli di grigio dei *pixel* come molecole o atomi di sistemi fisici, applicando idee e tecniche usate per lo studio degli stati di equilibrio dei processi chimici a differenti temperature. E' da notare come Geman e Geman ottennero i loro risultati in maniera del tutto indipendente rispetto ai lavori legati al principio *MaxEnt*. Come vedremo successivamente solo recentemente sono stati pubblicati alcuni articoli in cui questi risultati sono stati approfonditi e studiati da un unico punto di vista (Berger, Della Pietra, Della Pietra - 1996; Zhu, Mumford - 1997).

4.1 *Markov Random Fields (MRF) e Distribuzioni di Gibbs*

Sia $S = \{s_1, s_2, \dots, s_N\}$ un insieme di *sites* o *places*, definiti su un reticolo (nel nostro specifico caso una data immagine I) e sia $\Psi = \{\Psi_s, s \in S\}$ un sistema di intorni per S , cioè una collezione di sottoinsiemi di S tali che $\forall s \in S$:

- 1) $s \in \Psi_s$
- 2) $s \in \Psi_r \Leftrightarrow r \in \Psi_s$

Gli elementi in Ψ_s vengono detti i *vicini* di s .

Un sottoinsieme di S è una *clique* C se ogni coppia di distinti *place* in C risulta essere formata da elementi *vicini* nel senso appena definito. Un sistema di intorni omogeneo ben noto in *Image Processing* è il seguente :

$$\Psi = \{\mathfrak{S}_{x,y}, (x,y) \in I\} \quad (13)$$

dove :

$$\mathfrak{S}_{x,y} = \{(k,l) \in I : 0 < (k-x)^2 + (l-y)^2 \leq c\} \quad (14)$$

e il parametro c controlla la dimensione del sistema. Ovviamente il numero delle possibili *clique* cresce rapidamente con c . Un altro sistema di intorni, attualmente applicato al problema della classificazione delle parole nell'elaborazione automatica dei linguaggi naturali, è descritto in dettaglio in (Della Pietra, Della Pietra, Lafferty - 1997): in questo caso le proprietà naturali del linguaggio (una singola lettera minuscola, una singola lettera maiuscola, lettere minuscole adiacenti, ecc.), sono ridefinite utilizzando delle opportune *feature* su grafi.

Def. 4.1 Sia $X = \{X_s, s \in \hat{I} S\}$ una variabile random indicizzata da S e sia G l'insieme dei valori possibili (per esempio il numero dei livelli di grigio in una data immagine) tali che $X_s \in G, \forall s \in \hat{I} S$ e sia \mathbf{W} l'insieme di tutte le possibili configurazioni. X è un **Markov Random Field (MRF)** su \mathbf{Y} se $P(X) = \frac{1}{Z} e^{-U(\mathbf{w})}$, $\forall \mathbf{w} \in \mathbf{W}$ e $P(X_s = x_s / X_r = x_r, r \in \hat{I} s) = P(X_s = x_s / X_r = x_r, r \in \hat{I} Y_s)$.

In pratica in un MRF è possibile valutare la probabilità di avere uno specifico valore in uno stato del sistema, utilizzando solo le informazioni del relativo intorno.

Def. 4.2 Dato un insieme di place S e un sistema di intorni \mathbf{Y} , una **distribuzione di Gibbs**, (Gibbs Distribution), è una misura di probabilità p su \mathbf{W} tale che:

$$p(\mathbf{w}) = \frac{1}{Z} e^{\frac{-U(\mathbf{w})}{T}} \quad (15)$$

dove Z è la costante di normalizzazione o di partizione definita da :

$$Z = \sum_{\mathbf{w} \in \Omega} e^{\frac{-U(\mathbf{w})}{T}} \quad (16)$$

T è la costante relativa alla temperatura e U è l'energia data da :

$$U(\mathbf{w}) = \sum_C V_C(\mathbf{w}) \quad (17)$$

dove le V^C sono chiamate potenziali o funzioni potenziale e si riferiscono a specifiche clique C .

Le funzioni potenziale dipendono dalle *clique*, cioè il loro valore dipende solo dai valori assunti dagli stati in una data *clique* C .

Il parametro globale T viene usato per simulare gradualmente un processo di *annealing* che converge ad uno stato di equilibrio del sistema. Nell'originale articolo di Geman e Geman, tale tecnica è stata utilizzata per ottenere una stima *MAP - Maximum a Posteriori*, di un'immagine degradata. La scelta del sistema di intorno è cruciale, ed è strettamente correlata con la conoscenza reale del problema esaminato. Si veda al riguardo (Leviton, Chan, Herman - 1995) dove viene descritto, in un particolare caso applicativo, il *tuning* dei parametri di una distribuzione di Gibbs e un metodo statistico per la validazione della distribuzione ottenuta rispetto alla distribuzione da modellare.

4.2 Il Campionamento di Gibbs

Quando si ha a che fare con le distribuzioni di Gibbs è spesso impossibile calcolare il valore della costante Z in maniera esplicita poiché essa coinvolge un gran numero di configurazioni. In pratica è spesso necessario ricorrere ad un'efficace procedura di campionamento o *sampling*.

Tali procedure utilizzano i cosiddetti metodi di *Monte Carlo Markov Chain* (MCMC) che operano solo su campioni estratti da una distribuzione teorica ottenuta generando una sequenza campione la cui evoluzione viene controllata dalle opportune variazioni di energia. Per evitare di cadere in minimi/massimi locali, si deve necessariamente operare in maniera non deterministica, utilizzando opportunamente il parametro di controllo relativo alla temperatura del sistema. In

questo modo è possibile ottenere campioni di distribuzione, i cosiddetti *true samples*, con bassa temperatura che sono più reali ed espressivi.

I MRF e le distribuzioni di Gibbs sono legati tra loro dal *teorema di Hammersly-Clifford* :

Teorema 4.1 *Sia \mathbf{Y} un sistema di intorni. Allora X è un MRF su \mathbf{Y} se e soltanto se $p(\mathbf{w}) = P(X = \mathbf{w})$ è una distribuzione di Gibbs su \mathbf{Y} .*

Si veda (Besag - 1973) per la relativa dimostrazione. Grazie a quest'importante risultato teorico è possibile combinare insieme MRF e distribuzioni di Gibbs per estrarre correttamente un campione da una popolazione oggetto di studio.

Osserviamo che la probabilità condizionale rispetto alle caratteristiche locali di una distribuzione di Gibbs è espressa da:

$$p(x_s | x_r, r = s) = \frac{p(\mathbf{w})}{\sum_{x_s \in G} p(\mathbf{w})} \quad s \in S, \mathbf{w} \in \Omega \quad (18)$$

E' adesso possibile esprimere questa distribuzione in un modo più efficace usando il Teorema (4.1).

Infatti fissati $\mathbf{w} \in \mathbf{W}$ e $s \in S$, se $p(\mathbf{w}) = P(X = \mathbf{w})$ è Gibbsiana e X è un MRF allora :

$$P(X_s = x_s | X_r = x_r, r = s) = \frac{1}{Z_s} e^{-\frac{1}{T} \sum_{C: s \in C} V_C(\mathbf{w})} \quad (19)$$

dove

$$Z_s = \sum_{x \in G} e^{-\frac{1}{T} \sum_{C: x \in C} V_C(\mathbf{w}^x)} \quad (20)$$

Il simbolo \mathbf{w}^x denota la configurazione il cui valore è x al *place* s e concorda con \mathbf{w} in qualunque altro *place*.

L'idea fondamentale del *Gibbs Sampling* è dunque quella di realizzare un algoritmo di campionamento che attraverso una sequenza di “sostituzioni” di *place* converga ad un campione che segua la distribuzione teorica in oggetto.

Questo campionamento genera una sequenza di configurazioni $X(t) = (X_{s_1}(t), X_{s_2}(t), \dots)$ per $t = 1, 2, \dots$, ottenute dalle probabilità condizionali di cui sopra nella seguente maniera: tutti i *place* vengono ripetutamente visitati e ad ogni passo ne viene rimpiazzato uno in accordo con la (18) e (19), realizzando quello che viene tecnicamente chiamato uno *sweep*.

Dato il sito s_t e conoscendo l'espressione funzionale del modello generale di probabilità P , è possibile campionare per esso un nuovo valore $x \in G$, dove $s = s_t$ e $\mathbf{w} = X(t-1)$. In altre parole otteniamo una nuova configurazione dove $X_{s_t}(t) = x$ e $X_s(t) = X_s(t-1), s \neq s_t$.

Il *Gibbs Sampling* è quindi un processo iterativo capace di valutare ad ogni passo la distribuzione di probabilità relativa al cambiamento di un dato *place*, osservandone semplicemente i suoi vicini.

Questo semplice algoritmo converge alla distribuzione P senza tenere conto della configurazione iniziale. Inoltre la convergenza è garantita anche usando una sequenza decrescente di temperature.

La dimostrazione della correttezza di questi algoritmi chiamati *Relaxation* e *Annealing*, può essere trovata in (Besag - 1973; Geman, Geman - 1984).

4.3 Un esempio

Per meglio comprendere il *Gibbs Sampling* e la sua utilità pratica nei problemi statistici viene presentato e discusso un semplice esempio.

Partendo da una stringa *random* di caratteri ASCII, di qualsiasi lunghezza, è possibile generare stringhe che obbediscono ad uno specifico modello probabilistico.

Per esempio se vogliamo modellare una stringa della forma *bab*, è possibile applicare il *Gibbs Sampling* usando la seguente funzione esponenziale :

$$p(\mathbf{w}) = \frac{1}{Z} e^{\sum_I f(\mathbf{w})} \quad (21)$$

dove

$$Z = \sum_{\mathbf{w}} e^{\sum_I f(\mathbf{w})} \quad (22)$$

e la *feature* f esprime il *peso* associato con la caratteristica che vogliamo modellare. Definiamo tale *feature* come segue: f assume il suo massimo valore K quando, nel dato intorno I , rappresentato in questo caso dalle 3 lettere più vicine, c'è esattamente la configurazione *bab*, zero altrimenti. In questo modo, usando il *Gibbs Sampling*, è possibile applicare immediatamente la probabilità condizionale (18) e (19) di cui sopra. Restringendo per semplicità il *range* dei possibili caratteri ASCII ad $\mathbf{A}=\{a,b,c\}$, e supponendo di estrarre ad un certo punto della

computazione un *place* che corrisponde al centro x della sottostringa **cbcb** a , la procedura di *sampling* computa :

$$p(x=a | \mathbf{cb_ba}) = \frac{e^{f_{MAX}}}{e^{f_{MAX}} + e^0 + e^0} = \frac{e^K}{e^K + 2}$$

$$p(x=b | \mathbf{cb_ba}) = \frac{e^0}{e^{f_{MAX}} + e^0 + e^0} = \frac{1}{e^K + 2}$$

$$p(x=c | \mathbf{cb_ba}) = \frac{e^0}{e^{f_{MAX}} + e^0 + e^0} = \frac{1}{e^K + 2}$$

la lettera c viene sostituita semplicemente campionando il giusto valore/lettera in accordo con la distribuzione di cui sopra.

Ovviamente nel dato esempio sarà più probabile scegliere la lettera a che riproduce esattamente la distribuzione in oggetto piuttosto che le altre due.

5. Le Applicazioni all'Elaborazione delle Immagini

In questa Sezione descriveremo come quanto visto finora abbia trovato nel campo delle Elaborazione delle Immagini una diretta quanto efficace risposta a differenti e specifiche problematiche di ricerca. In particolare descriveremo alcune applicazioni dei cosiddetti metodi *entropici* nella modellizzazione di immagini digitali, nell'*image restoration*, nell'*edge finding* e in un algoritmo di *threshold*. Si rimanda invece al Capitolo 4 per la discussione dettagliata di alcuni risultati preliminari da noi ottenuti nella costruzione di Mappe di Salienza.

5.1 Il Principio dell'Entropia MiniMax

L'utilizzo dei *Markov Random Field* (MRF) in *Computer Vision* è oggi uno dei metodi di modellizzazione più diffusa. Partendo dall'idea di Geman e Geman si sono sviluppate infatti nel corso degli anni innumerevoli variazioni sul tema, di cui si può trovare traccia sfogliando una qualsiasi rivista scientifica di *Image Processing*; ciò ne conferma, a dispetto della relativa semplicità di base, le effettive potenzialità di tale metodologia che permette di modellare in maniera quasi del tutto naturale una variegata classe di fenomeni visuali. Per definire un *Markov Random Field* è necessario definire opportunamente un dato sistema di intorni e una serie di *feature*, con i relativi pesi, che ne caratterizzano il fenomeno sotto osservazione.

Solo per citare un esempio anche uno dei problemi classici quali quello della ricerca dei lati in un'immagine è stato affrontato con l'utilizzo dei MRF, attraverso un algoritmo di *simulated annealing* volto alla minimizzazione di un'opportuna funzione costo (Tan, Gelfand, Delp - 1991).

In questo paragrafo descriveremo brevemente una ricerca tuttora in corso, dovuta principalmente a S.C. Zhu e D. Mumford, i quali in una serie di lavori pubblicati nel corso degli ultimi anni, hanno proposto una nuova teoria per la costruzione di modelli statistici di immagini digitali e/o *texture* attraverso un utilizzo combinato del principio **MaxEnt** e dei *Markov Random Field* nonché della cosiddetta *filtering theory*.

L'apprendimento automatico proposto utilizza infatti una ben specificata serie di *feature*, provenienti da quella che in gergo tecnico viene definita una *bank*

filter, dette anche statistiche, dalla quale andranno poi scelte quelle più significative in accordo con un opportuno criterio di scelta basato sul *gain* che ciascuna *feature* apporta al modello incrementale che si sta via via costruendo.

Siano allora assegnate tali *feature* $F_{\mathbf{a}}(\mathbf{I})$ $\mathbf{a}=1,2, \dots,k$. Data un'immagine \mathbf{I} di cui si vuole stimare il modello di probabilità $f(\mathbf{I})$ è possibile considerarne il valore atteso $E_f[F_{\mathbf{a}}(\mathbf{I})]$ rispetto alle $F_{\mathbf{a}}(\mathbf{I})$ $\mathbf{a}=1,2, \dots,k$. Tali valori (solo teorici, in quanto la f è sconosciuta) riassumono il comportamento statistico del modello, rispetto alle singole *feature* e prendono il nome di *distribuzioni marginali*. L'ipotesi alla base del modello in questione si basa sul fatto che il modello di probabilità $f(\mathbf{I})$ è univocamente determinato dalle sue distribuzioni marginali (Zhu, Whu, Mumford - 1997).

Il modello viene via via costruito, attraverso un approccio strettamente incrementale, aggiungendo di volta in volta quella *feature* che trasporta più informazione. In particolare si procede stimando empiricamente, attraverso cioè un vero e proprio calcolo effettivo, il valor medio delle singole *feature* sui dati di *training* disponibili, vincolandone di conseguenza il relativo problema di inferenza associato.

Ci si pone allora l'obiettivo di trovare un modello $p(\mathbf{I})$, tale che:

$$E_p[F_{\mathbf{a}}(\mathbf{I})] = E_f[F_{\mathbf{a}}(\mathbf{I})] \quad \mathbf{a} = 1, 2, \dots, k \quad (23)$$

Tale modello si costruisce attraverso l'utilizzo combinato di due apparentemente opposti criteri di scelta:

- **Il Principio della Massima Entropia** o *MaxEnt* (3.1). Tra tutti i modelli $p(\mathbf{I})$ che soddisfano la (23), questo principio sceglie il più generale, massimizzando l'*entropia* (e quindi l'incertezza associata) su tutte le distribuzioni di probabilità che riproducono le particolari *feature* statistiche $F_a(\mathbf{I})$ $a = 1, 2, \dots, k$;
- **Il Principio della Minima Entropia.** Nella fase di estrazione/scelta delle *feature*, tra tutti i possibili insiemi di *feature* statistiche utili a caratterizzare il modello, si sceglie quello avente la minima distanza di Kullback-Leibler tra $f(\mathbf{I})$ e $p(\mathbf{I})$. E' possibile dimostrare che tale valore è proprio uguale all'*entropia* del modello $p(\mathbf{I})$ (a meno di una costante).

La stima della distribuzione di probabilità $f(\mathbf{I})$, di una o più immagini digitali, gioca un ruolo decisivo in diverse aree dell'Elaborazione delle Immagini. In questo caso il modello ottenuto prende il nome di *FRAME (Filters, Random fields, And Maximum Entropy)* ed è stato applicato con promettenti risultati a diversi fenomeni visuali. In particolare, come vedremo nel Capitolo successivo, tale approccio è stato utilizzato con ottimi risultati nell'Analisi e Sintesi di *texture* (Zhu, Wu, Mumford - 1997; Zhu, Liu, Whu - 1999).

Zhu e Mumford propongono di utilizzare come *feature* del loro modello gli istogrammi dell'immagine di riferimento \mathbf{I} , filtrata con un opportuno insieme di filtri (*bank filter*). Sono appunto questi istogrammi ad essere utilizzati come base, come stima delle distribuzioni marginali della funzione $f(\mathbf{I})$. Evidenze sperimentali mostrano come siano sufficienti un numero limitato di *feature*/filtri; l'algoritmo proposto, di tipo *greedy*, fa un uso estensivo del campionamento di

Gibbs, sia nella fase di estrazione/selezione delle *feature* sia nella fase di costruzione vera e propria del modello $p(\mathbf{I})$. L'espressione funzionale finale del modello *FRAME* è la seguente:

$$p_{\Lambda_k}(\mathbf{I}) = \frac{1}{Z(\Lambda_k)} e^{-\sum_{a=1}^k I_a(I_a)} \quad (24)$$

dove k sta ad indicare il numero dei filtri utilizzati, $\mathbf{L}_k = (\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_k)$ e I_a altro non è che l'immagine I filtrata con il filtro \mathbf{a} .

La stima dei parametri I_a $a=1, 2, \dots, k$ avviene mediante un algoritmo iterativo, facente uso di opportune tecniche numeriche. La potenza espressiva dei modelli d'inferenza statistica esistenti è limitata spesso dall'esponenziale numero di parametri che debbono essere considerati per ottenere risultati significativi. Nell'approccio considerato il costo computazionale è comunque rilevante, anche se si è in grado di inferire complesse strutture visuali. I precedenti modelli MRF sono infatti fortemente condizionati dalla relativa dimensione dello spazio degli intorni (il cosiddetto *neighborhood size*): il numero dei parametri cresce troppo anche utilizzando un modesto spazio degli intorni (per esempio 13x13). Il modello *FRAME* invece risolve tale problema, riuscendo ad incrementare la potenza espressiva del modello stesso, caratterizzando le interazioni locali come risposta a filtri digitali (lineari e non). Ulteriori riflessioni in merito all'utilizzo effettivo di tale metodo saranno presenti nel Capitolo successivo e verteranno principalmente sui risultati che tale metodologia di lavoro ha permesso di ottenere nel campo dell'Analisi e Sintesi di *texture*.

5.2 *Image Restoration*

La qualità visiva delle immagini digitali, siano esse relative a scene reali o "sintetiche" è spesso fortemente condizionata dagli effetti combinati di molteplici fattori *ambientali* (problemi legati per esempio all'acquisizione o alla trasmissione dei dati stessi) che introducono *rumore* nell'immagine stessa. Per eliminare o quanto meno limitare tali difetti visivi si rende necessario invertire, nei limiti del possibile, tale processo di degradazione dell'immagine. Tale operazione è resa ancora più difficile dal fatto che nella maggior parte dei casi non si è a conoscenza dell'esatta natura e/o tipo di *rumore* in oggetto. Si procede allora utilizzando delle opportune ipotesi di lavoro.

Supponiamo ad esempio di avere del cosiddetto *rumore additivo* ed un'immagine degradata I ottenuta come:

$$I = O * H + N \quad (25)$$

dove N è il rumore, H è un filtro (lineare o meno) applicato all'immagine iniziale O e $*$ è il classico operatore di convoluzione. I cosiddetti algoritmi di *restoration* provano ad invertire la relazione di cui sopra allo scopo di ottenere un'immagine quanto più "vicina" possibile, almeno nelle sue caratteristiche percettive, all'immagine originale.

In letteratura è possibile trovare una miriade di lavori di ricerca in cui tale problema è stato oggetto di studio; il già citato lavoro di Geman e Geman del 1984, era rivolto in questa direzione. In questa sezione ci limiteremo a considerare

quei metodi che siano in qualche modo legati ai concetti di *information entropy* e/o *MaxEnt*. In questo senso le prime applicazioni a tale problematica di ricerca del principio *MaxEnt* si ebbero con B.R. Frieden (1972; 1980). In questi lavori egli considera una serie di metodi di *restoration*, basati appunto sull'utilizzo dell'*entropia*, che in maniera del tutto naturale portano a risolvere un classico problema di massimizzazione. Se si considera infatti il processo di formazione di un'immagine come un processo *random* in cui una serie di palline colorate o *silver grains* come li chiama lo stesso Frieden, che rappresentano gli elementi costitutivi dell'immagine stessa, vanno a finire in una data posizione obbedendo ad una specifica legge di probabilità, è possibile esprimere il numero totale di modi possibili W nella seguente maniera:

$$W(o_1, o_2, \dots, o_N) = \frac{\left(\sum_{i=1}^N o_i \right)}{o_1! o_2! \dots o_N!} \quad (26)$$

Facendo uso di semplici riformulazioni (Jaynes - 1957; Wu - 1997) è possibile dimostrare che massimizzare l'espressione (26) è lo stesso che massimizzare l'*entropia* associata con i valori osservati $\{o_i, i=1,2,\dots,N\}$ soggetti ai vincoli dovuti al rumore N . Allora l'espressione finale, stimata, per la matrice O , diventa:

$$o_i = e^{\left(-1 - \sum_m I_m h_{i,m} \right)} \quad i = 1, 2, \dots, N \quad (27)$$

dove l'indice m indica il numero di dati disponibili osservati, i I_i sono i parametri liberi che definiscono l'oggetto e gli $h_{i,m}$ sono i corrispondenti valori nella matrice H di cui sopra. In questo modo è possibile ottenere significativi risultati, soprattutto in campo astronomico. Altri lavori in questo campo, con i

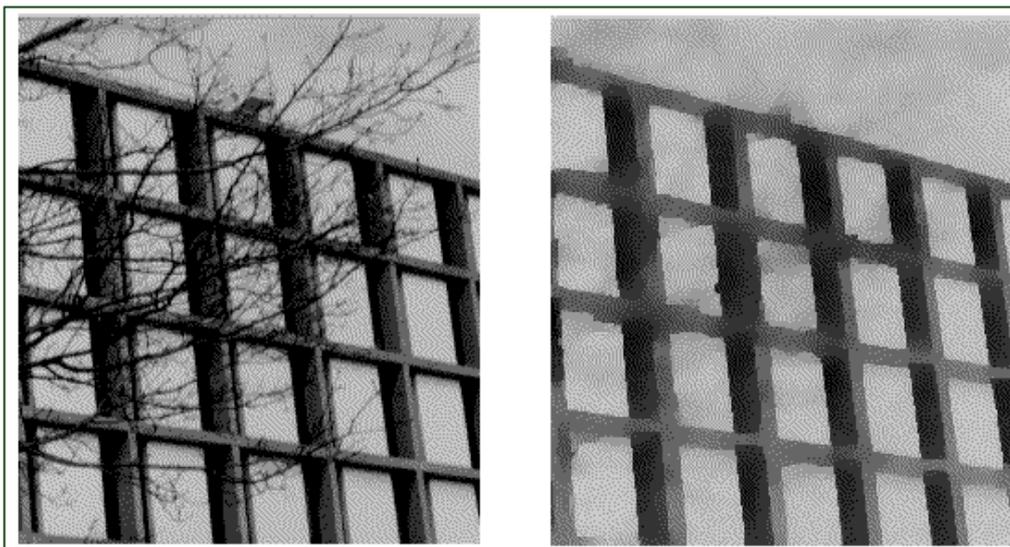


Fig. 3 - Rimozione di rumore ottenuta con il modello GRADE

relativi algoritmi di *restoration* possono essere trovati in (N.Wu - 1997). La maggior parte di questi risultati sono oggi utilizzati ed applicati in astronomia, in medicina, nello studio dei raggi X, nell'analisi di immagini SAR e così via.

Anche i lavori di Zhu e Mumford descritti in precedenza trovano applicazioni in questo campo. Infatti è possibile classificare le funzioni potenziale apprese dall'algoritmo *MiniMax* in due differenti categorie/classi: i termini di *diffusione* e i termini di *reazione*. Studi sperimentali hanno mostrato come i primi agiscano come componenti di *denoising* mentre i termini di *reazione* sono responsabili della formazione dei *pattern* nonché dell'esplicita caratterizzazione delle *feature* coinvolte. Seguendo quest'idea Zhu e Mumford nel 1987 hanno quindi generalizzato il loro modello di riferimento ottenendo il modello *GRADE* (*Gibbs Reaction and Diffusion Equations*) che ha permesso loro di ottenere una funzione energia U , che a seconda dei casi può essere utilizzata per la formazione di specifici *pattern*, per l'eliminazione del rumore (vedasi Fig.3) o in taluni casi per realizzare l'*enhancement*.

5.3 *Threshold*

La selezione o la ricerca di un giusto valore di soglia o *Thresholding* per una data immagine è sicuramente una delle tecniche più utilizzate nella segmentazione delle immagini digitali. In particolare tale approccio si basa sull'assunzione che in una qualunque scena gli oggetti in essa contenuti possono essere nettamente distinti dallo sfondo, dal *background*, semplicemente analizzandone i valori relativi all'intensità. In questo paragrafo presentiamo brevemente un algoritmo per la selezione del giusto valore di soglia che fa uso del principio *MaxEnt* e della logica *fuzzy* (Cheng, Chen, Li - 1998).

La scelta del giusto valore di soglia è una delle componenti più critiche e sensibile nelle applicazioni specifiche in cui si rende necessaria soprattutto perché è molto legata al particolare tipo di immagine che si sta processando.

L'algoritmo determina attraverso un opportuno algoritmo di *simulated annealing* i due parametri critici del sistema, utili per l'individuazione delle relative funzioni di *membership* grazie alle quali si ottiene una partizione *fuzzy* dell'immagine di input. La giusta stima di tali parametri è realizzata grazie all'applicazione del principio *MaxEnt* applicato all'*entropia* delle partizioni *fuzzy* via via considerate. Grazie a tali partizioni è possibile poi determinare i rispettivi valori di soglia sia nel caso del *bi-level thresholding* sia nel caso *multiple thresholding*.

L'algoritmo in questione è applicato ad immagini a toni di grigio ed è in grado di ottenere soddisfacenti risultati riducendo notevolmente le risorse

computazionali utilizzate. Ovviamente si rimanda direttamente a (Cheng, Chen, Li - 1998) per i dettagli del caso.

L'utilizzo combinato dell'*entropia* di Shannon e della teoria *fuzzy* è sicuramente un settore importante della *Computer Vision*. D'altra parte la natura stessa delle problematiche legate all'Elaborazione delle Immagini nasconde in sé una sorta di *fuzziness*, si pensi per esempio ai concetti sicuramente non formali di lato, di tessitura, di bordi, ecc. oppure alla perdita di informazione intrinseca esistente nel rappresentare scene 3D in uno spazio bidimensionale. In quest'ambito diverse applicazioni sono state realizzate, molti dei risultati classici sono stati raffinati utilizzando tecniche *fuzzy* ed in molti di questi approcci l'*entropia* di Shannon vi ha sempre fatto capolino (Chi, Yan, Pham - 1996).