

Consiglio Nazionale delle Ricerche

**ISTITUTO DI ELABORAZIONE
DELLA INFORMAZIONE**

PISA

**CALCOLO DELLA DISPARITÀ STEREOSCOPICA
IN PRESENZA DI OCCLUSIONI**

Riccardo March, Salvatore Minutoli

Nota Interna B4-10

Aprile 1991

**CALCOLO DELLA DISPARITÀ STEREOSCOPICA
IN PRESENZA DI OCCLUSIONI**

Riccardo March, Salvatore Minutoli

Introduzione.

Il problema affrontato in questo ambito consiste nella simulazione del processo di visione stereoscopica, cioè nella ricostruzione di una rappresentazione tridimensionale a partire da due immagini bidimensionali. L'uomo, infatti, vede una scena tridimensionale da due punti di osservazione distinti : l'occhio destro e quello sinistro. Le immagini proiettate su ciascuna delle due retine sono bidimensionali. Queste due immagini, dopo una prima preelaborazione, arrivano al cervello che provvede a fonderle per darci la sensazione di un'unica immagine tridimensionale. Il criterio adottato è quello di sfruttare la differenza di posizione di ogni punto fra le due immagini per calcolare la sua distanza dall'osservatore.

La difficoltà sta nello stabilire la corrispondenza tra un punto di un'immagine con uno dell'altra. La semplice equivalenza delle intensità non è sufficiente a darci un valido criterio di corrispondenza.

Questo principalmente per due motivi :

- 1) Possono esserci molti punti aventi la stessa intensità luminosa
- 2) Spesso le immagini sono rumorose risultando in un'alterazione più o meno ampia dell'intensità luminosa in alcuni punti.

Questo lavoro si pone l'obiettivo di simulare il processo di fusione stereoscopica per mezzo di un modello matematico basato sul calcolo delle variazioni[1]. Come spesso accade nella visione, i passaggi da una rappresentazione ad un'altra non saranno sempre univoci, ci saranno cioè dei problemi di ambiguità. Questi saranno caratterizzati dal fatto che, dato un problema la sua soluzione in genere non è unica. Per renderla unica si ricorrerà alla teoria della regolarizzazione sviluppata dai matematici russi ed in particolare da Tikhonov[15].

Questa permette di restringere lo spazio delle soluzioni ammissibili imponendo dei vincoli aggiuntivi che dovranno essere

scelti compatibilmente con il tipo di problema da risolvere. Nel caso della visione stereoscopica questi vincoli derivano dal supporre che gli oggetti sono generalmente compatti e opachi dando luogo, nelle immagini proiettate sulle retine, a superfici continue. Cio' e' vero ovunque tranne che sui contorni degli oggetti dove le superfici diventano discontinue a causa della brusca variazione di distanza che si ha passando da un'oggetto ad un altro.

La teoria della regolarizzazione cosi' come originariamente concepita, non si e' dimostrata pero' sufficiente se usata da sola. A questo scopo Terzopoulos[13] ha esteso tale teoria rendendola piu' idonea all'uso nell'ambito della visione.

Infatti con la regolarizzazione si impongono dei vincoli di continuita' globali che distruggerebbero le discontinuita' presenti nella scena. Terzopoulos ha introdotto le funzioni di controllo di continuita', attraverso le quali si possono rilasciare, nelle zone opportune, i vincoli di continuita'. I primi esperimenti fatti si basavano sull'ipotesi che i contorni degli oggetti, e quindi la funzione di controllo di continuita', fossero noti a priori. Questo comunque non e' vero, in quanto i dati di partenza comprendono soltanto le intensita' proiettate su ciascuna delle due immagini. In un secondo momento si e' quindi passati a considerare questa funzione come un ulteriore incognita del problema.

Un modello matematico per la visione stereoscopica basato su questa teoria e' gia' stato formulato in precedenti lavori[4][5].

Le simulazioni su calcolatore di questo modello hanno, difatti, evidenziato un netto miglioramento rispetto all'uso della teoria della regolarizzazione di Tikhonov.

L'obiettivo del presente lavoro e' quello di raffinare ulteriormente il modello, prendendo in considerazione il problema delle occlusioni. Questo e' un aspetto che finora e' stato tralasciato, ma dal quale non si puo' prescindere per una

trattazione piu' rigorosa dell'argomento[8].

Data, infatti, una coppia di immagini stereo, il processo di fusione stereoscopica deve tenere conto che in ciascuna delle due immagini vi sono delle zone che non corrispondono ad alcuna zona nell'altra.

Queste zone si trovano nelle vicinanze delle curve lungo le quali la profondita' e' discontinua, e che, fisicamente, corrispondono alle occlusioni tra oggetti diversi. Nella formulazione del modello matematico si dovra' quindi tenere presente che nelle zone di occlusione la funzione di disparita' stereoscopica, che indica la corrispondenza tra punti delle due immagini, non e' definita.

Il riconoscimento di forma e distanza degli oggetti tramite la visione stereoscopica ha alcune importanti applicazioni. Tra queste possiamo citare il suo uso nella robotica. Uno dei problemi che finora ha impedito di costruire robot che riescano a muoversi con sicurezza in ambienti non noti a priori e' che essi, in genere, non sono dotati di un sistema visivo particolarmente efficiente. Spesso il rimedio consiste nel costruire un ambiente che favorisca le operazioni del robot : si tratta in pratica di adattare l'ambiente al robot. La progettazione di sistemi visivi piu' sofisticati permettera' in un certo senso di invertire questa tendenza : e' piu' ragionevole, infatti, che il robot sia in grado di operare in diversi ambienti, adattandosi di volta in volta alle diverse situazioni.

La visione stereoscopica, in particolare, dara' al robot, dotato di due telecamere, la possibilita' di valutare, ad esempio, la distanza di eventuali ostacoli presenti sul suo cammino. Inoltre semplifichera' notevolmente alcuni compiti come la manipolazione di oggetti : infatti, essendo ora note le loro dimensioni e la distanza a cui si trovano, si hanno a disposizione delle precise informazioni sui movimenti da compiere per afferrarli.

Teoria della regolarizzazione.

Il problema della visione stereoscopica e' un problema matematicamente mal posto. Data una coppia di immagini stereo, il problema di trovare la corrispondenza stereoscopica tra punti delle due immagini non ha, in genere, soluzione unica. Questo e' dovuto alla presenza, in ciascuna delle due immagini, di numerosi punti aventi la stessa intensita' luminosa. Il compito e' reso piu' difficile dalla presenza di rumore.

Normalmente dato il problema

$$Au=f$$

dove A e' un operatore, il problema diretto consiste nel calcolare f una volta dato il valore di u . Noi avremo spesso a che fare con il caso inverso, cioe', conosciamo il valore di f (ad esempio l'immagine proiettata) e vogliamo determinare u (la scena che ha dato luogo all'immagine). In base alle considerazioni appena fatte questo e' un problema mal posto nel senso di Hadamard. Hadamard, infatti, ha definito un problema essere ben posto quando la sua soluzione soddisfa i tre seguenti requisiti :

- 1) esiste
- 2) e' unica
- 3) dipende in modo continuo dai dati iniziali.

Come abbiamo gia' detto, la seconda condizione viene in genere a mancare. Per ovviare a questo inconveniente Poggio ed altri[3][11] hanno proposto l'uso di metodi di risoluzione di problemi mal posti, che impongono dei vincoli aggiuntivi allo scopo di restringere lo spazio delle soluzioni ammissibili. Infatti sono stati sviluppati dai matematici russi, in particolare da Tikhonov[15], dei metodi che si pongono l'obiettivo di rendere unica la soluzione di questi problemi, facendoli diventare cosi'

ben posti. I metodi proposti da Tikhonov vanno sotto il nome di teoria della regolarizzazione. Esaminiamo adesso questa teoria considerando il caso bidimensionale che e' quello che utilizzeremo in seguito.

Sia H uno spazio lineare delle funzioni ammissibili e regolari definite in \mathbb{R}^2 .

Sia

$$\mathcal{J}(v) : H \longrightarrow \mathbb{R}$$

un funzionale definito su H che misura il grado di regolarita' di una soluzione ammissibile $v(x)$ appartenente ad H , con $x=(x_1, x_2)$ appartenente a \mathbb{R}^2 .

Supponiamo inoltre che

$$\mathcal{F}(v) : H \longrightarrow \mathbb{R}$$

sia un funzionale definito su H che misura la discrepanza tra la funzione incognita e i dati del problema.

Consideriamo il seguente principio variazionale:

Determinare u appartenente ad H tale che $\mathcal{E}(u) = \inf \mathcal{E}(v)$ con v in H dove

$$\mathcal{E}(v) = \mathcal{F}(v) + \lambda \mathcal{J}(v) .$$

Il funzionale \mathcal{J} viene chiamato funzionale di stabilizzazione. La sua forma generica, nel caso bidimensionale, e' data dalla seguente seminorma:

$$\mathcal{J}(v) = |v|_p^2 = \sum_{m=0}^p \iint_{\Omega} \sum_{i=0}^m \binom{m}{i} w_{i,m}(x,y) \left(\frac{\partial^m v}{\partial x^i \partial y^{m-i}} \right)^2 dx dy$$

dove si presuppone naturalmente che esistano tutte le derivate parziali della funzione v fino all'ordine p . Da questa definizione si puo' notare che il funzionale \mathcal{J} fornira' i suoi valori piu' bassi in corrispondenza di funzioni che nel dominio di definizione, abbiano un comportamento che non da' luogo a brusche variazioni.

La proprieta' per cui una funzione abbia delle variazioni lente e graduali viene talvolta chiamata "smoothness", e le funzioni che abbiano tali caratteristiche sono dette funzioni "smooth".

Le funzioni $w_{i,m}(x)$ sono funzioni peso note a priori, continue e non negative.

Alcuni esempi di funzionali di stabilizzazione che sono spesso usati sono :

$$a) \quad |v|_1^2 = \int \int_{\Omega} (v_x^2 + v_y^2) dx dy$$

dove $p=1$ e

$$w_{0,0}(x,y)=0 \quad , \quad w_{0,1}(x,y)=w_{1,1}(x,y)=1$$

$$b) \quad |v|_2^2 = \int \int (v_{xx}^2 + 2v_{xy}^2 + v_{yy}^2) dx dy$$

dove $p=2$ e

$$w_{0,0}(x,y)=w_{0,1}(x,y)=w_{1,1}(x,y)=0$$

$$w_{0,2}(x,y)=w_{1,2}(x,y)=w_{2,2}(x,y)=1$$

La scelta della formulazione del funzionale \mathcal{J} andra' ovviamente fatta in base a considerazioni fisiche riguardanti il particolare problema da risolvere. In caso contrario si otterrebbe una soluzione poco significativa.

Nel caso oggetto di questo studio saremo interessati a determinare la forma tridimensionale degli oggetti presenti in una scena in base alle informazioni contenute nei valori dell'intensita' luminosa delle immagini proiettate. L'obiettivo sara', essenzialmente, quello di calcolare la distanza dei vari punti degli oggetti dall'osservatore. In questo caso la giustificazione del vincolo aggiuntivo di regolarita' consegue dal fatto che, essendo gli oggetti usualmente compatti ed opachi, le superfici generate avranno un andamento con variazioni graduali senza sbalzi improvvisi, ad eccezione delle regioni situate in prossimita' dei bordi degli oggetti.

Il funzionale \mathcal{F} , detto funzionale di penalizzazione, puo' essere espresso per esempio come :

$$\mathcal{F}(v) = \iint_{\Omega} (Au - f)^2 dx dy$$

Questo esprime una misura di quanto la funzione incognita v si discosta dai dati f del problema. Se l'operatore A e' continuo e lineare, si dimostra che vengono soddisfatte le tre condizioni di Hadamard, e quindi possiamo concludere che il problema e' ben posto. Il funzionale \mathcal{E} viene espresso in termini sia di \mathcal{F} che di \mathcal{S} : in questo caso l'obiettivo sara' quello di cercare il miglior compromesso tra la precisione con cui la funzione incognita approssima i dati del problema e il grado di regolarita' della funzione stessa. Il parametro λ ha il compito di determinare l'importanza relativa da attribuire a ciascuno dei due funzionali. La scelta di λ , in genere, si rivela critica: valori troppo bassi renderanno la soluzione molto sensibile al rumore ; valori troppo elevati saranno causa di un forte predominio del vincolo di smoothness, a scapito della precisione dei dati.

Il calcolo della funzione che minimizza il funzionale \mathcal{E} definisce un problema di approssimazione ottimale. Questo consiste nel

trovare la funzione ammissibile in H che rappresenta il miglior compromesso tra la "smoothness" e la vicinanza ai dati.

Una condizione necessaria ma, in generale, non sufficiente affinché u sia la funzione che minimizza il funzionale \mathfrak{E} e' data dall'annullarsi della variazione prima del funzionale. Questa condizione viene espressa dall'equazione di Eulero-Lagrange

$$\delta\mathfrak{E}(u) = \delta\mathcal{J}(u) + \lambda \delta\mathcal{F}(u) = 0$$

La non sufficienza della condizione dipende dal fatto che essa puo' essere soddisfatta anche da massimi e minimi locali. Si dimostra, comunque, che, se il funzionale e' convesso, se vale cioè che

$$\mathfrak{E}(a+kb) \leq \mathfrak{E}(a) + k\mathfrak{E}(b), \text{ dove } a < b \text{ e } 0 \leq k \leq 1,$$

la condizione e' anche sufficiente.

Nel caso che A sia un operatore lineare e continuo, l'equazione di Eulero-Lagrange e' una equazione differenziale lineare, ed ammetterá un'unica soluzione. In questo caso, quindi, vengono soddisfatti i tre requisiti formulati da Hadamard.

E' stato comunque dimostrato che il problema risulta essere ben posto anche nel caso in cui l'operatore A e' sottoposto a condizioni piu' deboli. Questo e' un risultato molto importante, come vedremo in un successivo capitolo quando sara' discusso piu' in dettaglio il problema della visione stereoscopica. Per ora ci limitiamo a dire che nel nostro caso l'operatore A non sara' lineare. Malgrado il rilascio del vincolo di linearita' la dimostrazione sopra citata ci consentira' di affermare che il problema sara' ugualmente ben posto.

Teoria della regolarizzazione in presenza di discontinuita'.

La teoria della regolarizzazione di Tikhonov si basa sull'assunzione che le funzioni da determinare siano continue su tutta la regione in cui vengono calcolate.

A questo riguardo, consideriamo il nostro problema di ricostruzione tridimensionale di una scena visiva. Pensiamo a come varia la distanza dall'osservatore man mano che ci spostiamo gradualmente da un punto all'altro dell'immagine : se ci muoviamo tra due punti vicini appartenenti allo stesso oggetto, la distanza varierà con continuità'.

Questa affermazione, però, non è più vera nelle vicinanze dei punti situati sui bordi degli oggetti. Qui accade, infatti, che due punti vicini possono appartenere ad oggetti, e quindi a superfici, diversi. La distanza in questo caso, passando da un punto all'altro, subisce una brusca variazione. Siamo perciò in presenza di una discontinuita' nella funzione che associa ad ogni punto dell'immagine la sua distanza dall'osservatore.

In questa circostanza l'applicazione dei metodi di regolarizzazione di Tikhonov, imponendo dei vincoli di regolarità globali, non determina un risultato corretto. Si avrà cioè una ricostruzione della scena nella quale le superfici saranno tutte raccordate tra di loro. Questo succede perché il funzionale di stabilizzazione penalizza le brusche variazioni della funzione incognita. Di conseguenza la funzione, punto per punto, dovrà assumere dei valori abbastanza vicini a quelli assunti dai punti nell'intorno. Questo dà luogo ad un effetto di interpolazione che tende ad eliminare i salti tra diverse superfici, impedendo così di ottenere una ricostruzione corretta che preservi le discontinuita' presenti nelle usuali scene visive.

La proposta di Terzopoulos consiste nell'introduzione delle

cosiddette funzioni di controllo di continuita'. Nella teoria della regolarizzazione di Tikhonov il funzionale di stabilizzazione e'

$$\mathcal{J}(v) = |v|_p^2 = \sum_{m=0}^p \iint_{\Omega} \sum_{i=0}^m \binom{m}{i} w_{i,m}(x,y) \left(\frac{\partial^m v}{\partial x^i \partial y^{m-i}} \right)^2 dx dy$$

dove le funzioni peso si suppongono essere non negative e continue.

Terzopoulos, invece, utilizza delle funzioni peso che possono avere delle discontinuita', permettendo a queste funzioni di avere dei salti da un valore positivo allo zero. Il funzionale cosi' ottenuto viene chiamato stabilizzatore con controllo di continuita'. Queste funzioni peso, cosi' definite, permettono di esercitare un controllo spaziale sulle caratteristiche di regolarita' della soluzione.

Assegnando lungo un determinato contorno il valore zero alla funzione peso, si annullera' il contributo della derivata nei punti contenuti in quel contorno. Questo fa si' che, lungo questi contorni, il gradiente della funzione incognita possa assumere valori arbitrariamente elevati, tali che, passando da una superficie all'altra, si possono verificare dei salti. Questo, infatti, e' quello che normalmente avviene nelle scene visive reali, ogniqualevolta un oggetto e' posto davanti ad un altro in modo da occluderlo parzialmente.

L'importante problema che si presenta e' quello di come definire le funzioni di controllo di continuita'. In alcuni casi possiamo supporre che le funzioni di controllo di continuita' siano gia' specificate, e quindi possiamo inserirle nella formula. Possiamo definire ad esempio $W(x,y)=1$ per tutti i punti (x,y) in cui la

funzione incognita deve risultare continua, e $W(x,y)=0$ nei punti in cui invece la funzione ammette discontinuita'. A questo punto rimane soltanto da calcolare la funzione u che minimizza il funzionale.

Vi sono, comunque, dei casi in cui le funzioni peso non sono note a priori, ma rappresentano un'ulteriore incognita del problema. Anche in questo caso e' possibile esprimere la soluzione del problema mediante un funzionale da minimizzare. In questo caso, in realta', il funzionale dipende da due funzioni incognite, in quanto bisogna considerare adesso come tale anche la funzione di controllo di continuita'[13][14]. Questo, pero', introduce alcune difficolta' nella risoluzione delle equazioni differenziali, per cui e' stato deciso di adottare un metodo alternativo. Si parte dall'osservazione che in assenza della funzione di controllo di continuita' la soluzione ottenuta, seppure non corretta, fornisce utili informazioni per il miglioramento della stessa. Il caso tipico infatti e' il seguente. Nelle zone sufficientemente distanti dai contorni di occlusione la ricostruzione sara' generalmente buona. In prossimita' dei contorni degli oggetti invece di avere due superfici nettamente separate tra di loro, si otterra' un'unica superficie dove il brusco salto e' stato praticamente smussato.

Si puo' allora tentare di costruire una approssimazione della funzione di controllo di continuita' in base a considerazioni sul gradiente della soluzione ottenuta. Siccome le discontinuita' originarie vengono trasformate in porzioni di superficie con gradiente piuttosto elevato, si puo' ritenere con una certa probabilita' che un gradiente abbastanza alto in un determinato punto sia indice della presenza di una discontinuita'. Possiamo quindi fissare una soglia T per cui se il modulo del gradiente supera, in un determinato punto, tale valore, se ne deduce la presenza di una discontinuita'.

Nel caso di superfici visive, inoltre, possiamo sfruttare il fatto che nei punti di discontinuita' saranno presenti degli edges[5], che possono essere identificati nelle immagini. Il problema di determinare un valore corretto per la soglia T e' piuttosto difficile. Non e' pensabile di determinare la funzione di controllo di continuita' in un'unica volta.

Si rendera' necessario organizzare il procedimento in piu' stadi[5]. Nel primo stadio si calcolera' la soluzione completamente "smooth" in quanto non avremo ancora una valutazione della funzione di controllo di continuita'. Successivamente, ad ogni stadio, si calcola dapprima una approssimazione della funzione di controllo di continuita' in base alla ricostruzione ottenuta nello stadio precedente. Poi si ricalcola la ricostruzione delle superfici della scena che sara' quella piu' "smooth" compatibilmente con i vincoli imposti dalla nuova funzione di controllo di continuita'. Queste superfici saranno quindi forzate ad assumere un andamento con variazioni graduali in tutti i punti esclusi quelli in cui la funzione di controllo di continuita' ha valore nullo. In questi ultimi punti, come sappiamo, le superfici sono libere di formare dei dislivelli di ampiezza arbitraria. Ad ogni stadio il valore della soglia usata per il gradiente sara' generalmente diverso in quanto man mano che si va avanti le caratteristiche della ricostruzione saranno diverse.

Il metodo che, in linea di massima, si e' rivelato piu' opportuno e' quello di partire con un valore di soglia sufficientemente alto e diminuirlo ad ogni stadio successivo. Questo corrisponde, a partire dalla soluzione completamente "smooth", a rilevare prima i contorni in cui il salto e' piu' brusco. Negli stadi successivi la soluzione viene continuamente raffinata rilevando di volta in volta salti sempre piu' piccoli.

La visione stereoscopica.

La visione stereoscopica e' quel processo che si occupa di ricostruire una scena tridimensionale a partire dalle due immagini bidimensionali provenienti dalle retine.

Inizialmente si pensava che il fatto di percepire le scene in modo tridimensionale fosse dovuto alla capacita' del cervello di sfruttare le conoscenze gia' acquisite per determinare la distanza e la forma degli oggetti. In pratica se noi conosciamo, ad esempio, le dimensioni di un oggetto, confrontandole con quelle che lo stesso ha nell'immagine, possiamo stimare la distanza a cui si trova.

Un esperimento eseguito da Bela Julesz ha pero' dimostrato che il processo di fusione stereoscopica avviene correttamente anche quando nelle due immagini stereo sono presenti oggetti di cui non abbiamo alcuna conoscenza. L'esperimento e' stato svolto nel modo seguente(vedi figura 1)[10].



Figura 1.

E' stata dapprima costruita una immagine contenente dei punti bianchi e neri distribuiti casualmente. In questo modo e' stata ottenuta l'immagine sinistra. L'immagine destra e' stata formata traslando i punti contenuti in una figura geometrica, in questo caso un quadrato, al centro dell'immagine verso sinistra e

riempiendo le parti rimaste vuote con altri punti casuali. Chiaramente le due immagini così ottenute viste singolarmente non contengono alcuna figura di cui conosciamo qualche caratteristica.

La cosa sorprendente, e che costituisce un risultato fondamentale per lo studio della visione stereoscopica, è che se guardiamo le due immagini mediante uno stereoscopio, abbiamo la sensazione di percepire un quadrato che si stacca nettamente dal fondo (vedi figura 2).

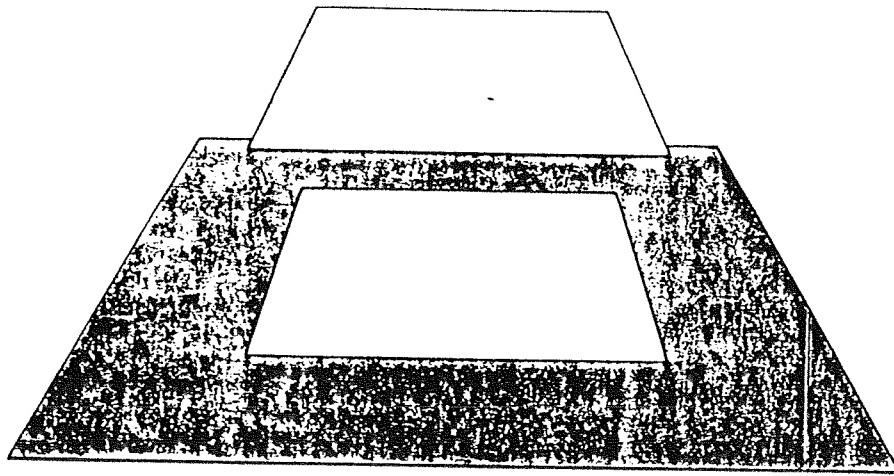


Figura 2.

Vediamo adesso più in dettaglio il procedimento che utilizzeremo per determinare la distanza dei punti della scena.

- 1) Per ogni punto dell'immagine sinistra determiniamo il punto dell'immagine destra che corrisponde allo stesso punto della scena. Fra i due punti delle due immagini diremo che si ha corrispondenza stereoscopica. Si noti che non è detto che, preso un punto in un'immagine, esista nell'altra immagine un punto corrispondente, cioè proiettato dallo stesso punto dello spazio. Questo è dovuto a due motivi. Uno è il fatto che i due campi visivi non sono mai completamente sovrapposti. Inoltre c'è anche il problema delle occlusioni. Se

consideriamo un oggetto situato davanti ad uno sfondo, esso nascondera' a ciascun occhio una regione dello sfondo stesso. Essendo pero' le due immagini prese da posizioni differenti la regione nascosta non sara' uguale nelle due viste. Ci saranno quindi alcune zone che in un'immagine si vedranno e nell'altra no. Per i punti di queste zone non esistera' quindi corrispondenza stereoscopica. Il processo che determina la corrispondenza stereoscopica tra i punti delle due immagini viene chiamato fusione stereoscopica.

2) Mediante l'uso di semplici formule geometriche determiniamo la distanza e quindi la precisa posizione del punto nello spazio. Illustriamo ora il modello che verra' utilizzato nel nostro algoritmo per calcolare la distanza date le coordinate di due punti fra i quali si ha corrispondenza stereoscopica.

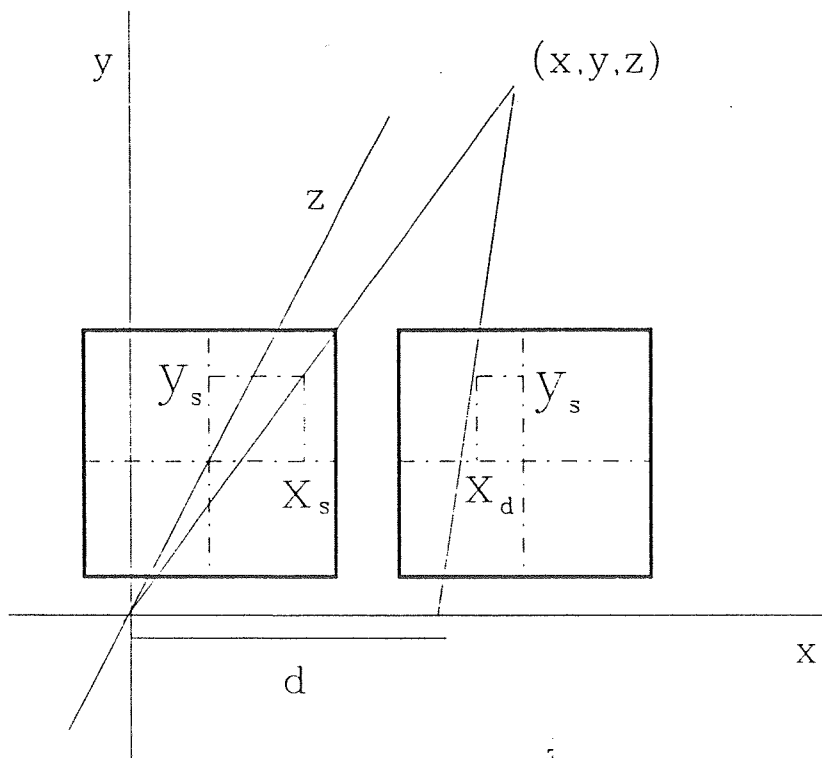


Figura 3.

Dati i tre assi x , y e z supponiamo, senza perdere in generalita', che i due schermi su cui verranno proiettate le immagini siano dei rettangoli paralleli al piano xy . Supponiamo inoltre che il centro dello schermo sinistro si trovi sull'asse z e che lo schermo destro risulti traslato di una distanza d sull'asse x rispetto a quello sinistro. Per quanto riguarda la proiezione di un punto sullo schermo, supponiamo che essa sia data dall'intersezione tra lo schermo stesso e la retta passante per il punto da proiettare e un punto posto dietro lo schermo (vedi figura 3). Se calcoliamo le proiezioni di un punto (x,y,z) su ciascuno dei due schermi notiamo che $y_s = y_d$ in quanto secondo il modello da noi adottato non vi e' alcuna differenza tra le due proiezioni rispetto l'asse y . I valori di x_s e x_d sono invece diversi .

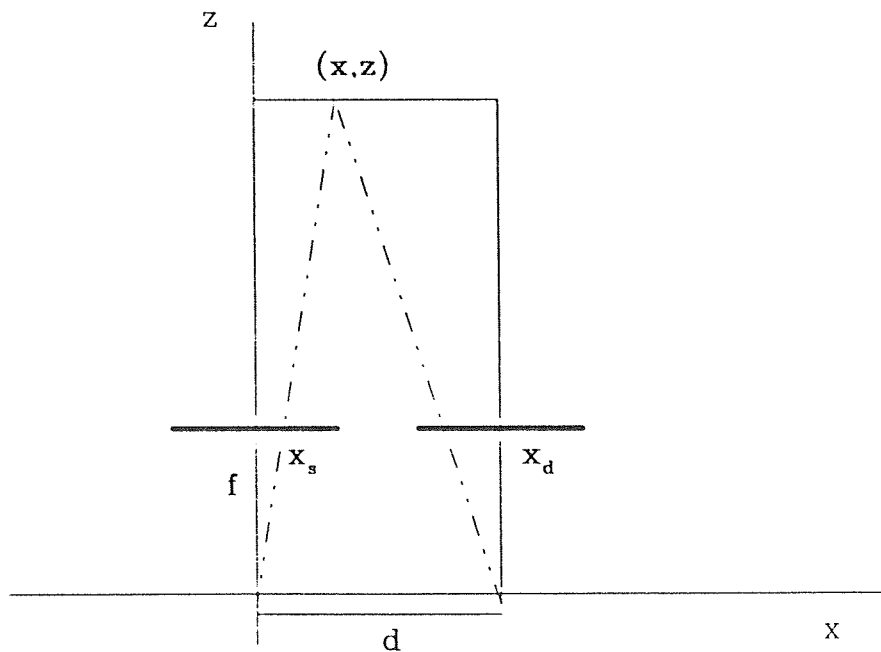


Figura 4.

A questo proposito osserviamo la figura 4, dove per motivi di migliore comprensibilita', ci limitiamo a considerare un piano parallelo al piano $x-z$. Mediante semplici formule che sfruttano la

similitudine fra triangoli, ricaviamo le seguenti espressioni.

Per la proiezione sullo schermo sinistro si ha

$$x_s = \frac{f}{z} x$$

Analogamente per lo schermo destro si ha

$$x_d = \frac{f}{z} (x-d)$$

dove d e' la distanza fra i centri dei due schermi.

Se denotiamo con u la differenza dei due valori ricavati otteniamo

$$u = x_s - x_d = \frac{f}{z} d$$

Il valore u viene detto disparita' binoculare (o disparita' stereoscopica), e come risulta dalla formula e' inversamente proporzionale alla distanza z dall'osservatore.

Nel determinare la corrispondenza stereoscopica tra due punti, si incontrano molte difficolta'. Il criterio utilizzato si basa sul fatto che due punti corrispondenti, sotto ipotesi non molto restrittive, hanno uguale intensita' luminosa. Questo pero' non e' sufficiente in quanto in ciascuna delle due immagini, vi sono, in genere, numerosi punti aventi la stessa intensita' luminosa. Inoltre c'e' il problema del rumore che modifica i valori di intensita'. La presenza delle zone di occlusione complica ulteriormente le cose.

Descriviamo ora come viene espresso matematicamente il problema della stereopsi, ignorando momentaneamente il problema delle occlusioni, e supponendo nota la funzione di controllo di continuita'. L'obiettivo e' di minimizzare il seguente funzionale

$$\mathfrak{E}(v) = \mathcal{P}(v) + \lambda \mathcal{J}(v).$$

La funzione incognita del nostro problema e' la funzione di disparita' stereoscopica[4][9]. Essa indica per ogni punto proiettato sull'immagine sinistra la posizione dello stesso punto nell'immagine destra. Piu' precisamente, detta u la funzione di disparita' stereoscopica, un punto della scena che nell'immagine sinistra viene proiettato in (x_1, y_1) nell'immagine destra si trovera' in $(x_1 + u(x_1, y_1), y_1)$. Utilizzando queste informazioni possiamo poi facilmente calcolare l'esatta posizione del punto nello spazio. Il funzionale di penalizzazione \mathcal{F} deve esprimere la condizione di corrispondenza stereoscopica. Un criterio ragionevole per stabilire la condizione di corrispondenza stereoscopica e' quello di imporre che due punti proiettati da uno stesso punto della scena abbiano la stessa intensita' luminosa. Questo non e' sempre vero, ma possiamo supporre che in certe condizioni non molto restrittive sia valido. Il funzionale \mathcal{F} verra' quindi espresso da

$$\mathcal{F}(u) = \iint_{\Omega} \left(L(x, y) - R(x + U(x, y), y) \right)^2 dx dy$$

che, appunto, penalizzera' i valori della u quando questa cerca di associare punti di diversa intensita' luminosa. Per quanto concerne il funzionale di stabilizzazione utilizzeremo il seguente

$$\mathcal{J}(u) = \iint_{\Omega} W(x, y) \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right) dx dy$$

Questo funzionale favorisce le funzioni che variano in modo lento

ed e' giustificato dall'osservazione che nelle scene reali sono in genere presenti oggetti compatti ed opachi che danno luogo a superfici che, tranne nelle vicinanze dei contorni, non variano in modo brusco. Il parametro λ controllera' l'importanza da attribuire a ciascuno dei due funzionali.

Funzione di controllo di continuita'.

Il passo seguente consiste nel tenere conto della funzione di controllo di continuita'. Essa, infatti, non e' nota a priori e deve essere calcolata insieme alla funzione di disparita' stereoscopica. Tuttavia, come vedremo, le immagini bidimensionali contengono utili informazioni che possono favorire il compito di determinarla.

Bisogna tenere conto che i salti di distanza tra superfici si trovano lungo i contorni che separano un oggetto da un altro. Salvo casi particolari che raramente si presentano nella realta' lungo questi contorni si verificano dei salti di valore di intensita' luminosa. Mediante un algoritmo di edge detection che permette di rivelare la presenza di questi contorni e' possibile determinare i punti in cui la funzione W deve essere posta a 0. Bisogna pero' prestare attenzione al fatto che un algoritmo di edge detection non restituira' soltanto i punti in cui la distanza subisce una brusca variazione ma fornisce tutti i punti in cui vi e' uno sbalzo di intensita' luminosa indipendentemente dal fattore che lo ha causato.

Quindi, una volta determinato l'insieme di tutti gli edge, sara' necessario utilizzare qualche altro metodo per estrarre quelli dovuti a discontinuita' nella distanza. Esaminiamo adesso un metodo per determinare gli edge di un'immagine[7][12].

Supponiamo di avere una zona di un'immagine in cui vi sia una brusca variazione di intensita' luminosa. Se calcoliamo la derivata prima rispetto ad una opportuna direzione, nel punto di massima variazione essa assumerà un valore relativamente elevato. In particolare la derivata prima in quel punto avrà un massimo locale. Questo implica che calcolando in quel punto la derivata seconda lungo la stessa direzione essa assumerà il valore 0 passando da un valore positivo ad uno negativo o viceversa. Abbiamo quindi da parte della derivata seconda un "passaggio attraverso lo zero". Questo termine e' spesso usato con la sua

traduzione in inglese "zero crossing".

Gli edges vengono quindi determinati localizzando gli "zero crossing" di

$$\nabla^2(G * I)$$

dove I e' l'immagine di cui si vogliono trovare gli "zero crossing";

G e' una funzione Gaussiana definita da

$$G(r) = \left(\frac{1}{2\pi\sigma^2} \right) \exp(-r^2/2\sigma^2)$$

$$\text{dove } r^2 = x^2 + y^2$$

utilizzata per filtrare l'immagine ed attenuare cosi' il rumore.

∇^2 e' l'operatore Laplaciano definito da

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

che viene utilizzato al posto della derivata direzionale seconda, in quanto, essendo un operatore isotropo, permette di evitare la dipendenza da una particolare direzione di derivazione. Esiste, infatti, un teorema, formulato da Marr e Hildreth[7], in cui si dimostra che, sotto opportune ipotesi non molto restrittive, l'uso del Laplaciano permette di localizzare correttamente gli "zero crossing".

Ora che abbiamo trovato tutti gli edge dell'immagine, si pone il problema di selezionare quelli che derivano da una discontinuita' nella distanza delle superfici dall'osservatore. Questo sottoinsieme di edge verra' poi usato per definire la funzione di

controllo di continuita'[5].

A questo scopo osserviamo che, se cerchiamo di calcolare la soluzione del problema ponendo la funzione di controllo di continuita' al valore 1 ovunque, otterremo una soluzione completamente "smooth". In questo caso, infatti, nelle zone in cui si sarebbero dovute ottenere delle discontinuita', si avra' un effetto di interpolazione, dovuto al funzionale di stabilizzazione, che raccordera' tra di loro le superfici.

Ciononostante, in queste zone la funzione sara' caratterizzata da un andamento abbastanza ripido. Allora, ripetendo il ragionamento in senso contrario, possiamo aspettarci che la presenza di gradienti sufficientemente elevati, sia un segno della presenza di una discontinuita' nella funzione che minimizza il funzionale. Cosi' facendo si corre il rischio di supporre la presenza di una discontinuita' anche nei punti in cui la soluzione ha intrinsecamente una superficie piuttosto inclinata. Comunque la funzione di controllo di continuita', quando e' posta a 0, permette la ricostruzione di discontinuita' ma non la impone. Quindi l'algoritmo che calcola la soluzione ha la possibilita' di correggere questi eventuali errori.

L'insieme degli zero-crossings dell'immagine ha l'importante compito di rendere questo procedimento piu' preciso ed affidabile. Infatti, in questo modo, abbiamo due condizioni che devono essere verificate per assumere in un certo punto la presenza di una discontinuita' : un gradiente sufficientemente elevato e la presenza di un edge. Questo fa diminuire la probabilita' di errori simili a quello descritto in precedenza, anche se non la annulla. Fatte queste premesse siamo adesso in grado di descrivere il metodo che utilizzeremo per determinare, iterativamente, sia la funzione incognita che descrive le superfici della scena da ricostruire sia la funzione di controllo di continuita'.

Il procedimento utilizzato e' il seguente. Definiamo prima la

funzione E che avra' il compito di indicare la presenza o meno di un edge in un determinato punto.

Porremo

$$E(i,j)=\begin{cases} 1 & \text{se nel punto } (i,j) \text{ e' presente un edge} \\ 0 & \text{altrimenti} \end{cases}$$

Calcoliamo dapprima le superfici degli oggetti ponendo la funzione $W=1$ ovunque. Come gia' detto, otterremo una funzione "smooth". A questo punto viene stabilito un valore di soglia T tale per cui se

$$|\nabla u(i,j)| > T$$

e

$$E(i,j)=1$$

viene posto

$$W(i,j)=0$$

Questo, in altre parole, equivale a selezionare i punti in cui la funzione U abbia una pendenza superiore ad un certo valore e, contemporaneamente, vi e' la presenza di un edge . In questi punti, secondo il ragionamento precedentemente descritto, vi e' una probabilita' abbastanza elevata che sia presente una discontinuita' nelle superfici visibili. Adesso tenendo fissa la W, calcoliamo le nuove superfici corrispondenti. Fatto questo fissiamo un altro valore di soglia e ripetiamo il procedimento di prima.

Tutto questo viene iterato per un opportuno numero di volte sufficiente a determinare una buona ricostruzione finale. La scelta del valore di soglia, durante le simulazioni eseguite si e' rivelata molto critica. Valori di soglia troppo piccoli rendono

l'operazione di selezione piuttosto sensibile al rumore .Con valori troppo alti, invece, si corre il rischio di ignorare punti in cui vi sono delle discontinuita'.

La strategia che in linea di massima e' stata adottata e' stata quella di iniziare con una soglia un po' alta e di abbassarla mano mano che si procedeva nei seguenti stadi di iterazioni.

Discretizzazione dell'equazione di Eulero-Lagrange.

Vediamo ora come viene discretizzata l'equazione di Eulero-Lagrange, per calcolare la soluzione iterativamente con metodi numerici[2].

Nelle regioni in cui le superfici sono continue, e si ha quindi $W(x,y)=1$, applicando il calcolo delle variazioni otteniamo l'equazione di Eulero-Lagrange :

$$\left[L(x,y) - R(x+U(x,y), y) \right] R_x(x+U(x,y), y) + \lambda \Delta U(x,y) = 0$$

dove

$$\Delta U(x,y) = \frac{\partial^2 U(x,y)}{\partial x^2} + \frac{\partial^2 U(x,y)}{\partial y^2}$$

Definendo, per ogni funzione presente nell'equazione,

$$\hat{f}(i,j) = f(x_i, y_j)$$

ed approssimando

$$\hat{f}_i(i,j) = \frac{1}{h} \left(f(i+\frac{1}{2}, j) - f(i-\frac{1}{2}, j) \right)$$

$$\hat{f}_{ij}(i,j) = \frac{1}{h^2} \left(f(i+\frac{1}{2}, j+\frac{1}{2}) - f(i+\frac{1}{2}, j-\frac{1}{2}) - f(i-\frac{1}{2}, j+\frac{1}{2}) + f(i-\frac{1}{2}, j-\frac{1}{2}) \right).$$

(analogamente si ottengono \hat{f}_j , \hat{f}_{ii} ed \hat{f}_{jj}) otterremo l'equazione alle differenze finite

$$\left[L(i,j) - R(i+U(i,j), j) \right] R_x(i+U(i,j), j) + 4\lambda \left[U^*(i,j) - U(i,j) \right] = 0$$

da cui

$$U^{n+1}(i,j) = U^{*n}(i,j) + \frac{1}{\lambda} \left[L(i,j) - R(i+U^{*n}(i,j),j) \right] R_x(i+U^{*n}(i,j),j)$$

dove $U^n(i,j)$ indica il valore di $U(i,j)$ alla n-esima iterazione e

$$U^{*n}(i,j) = U^n(i+1,j) + U^n(i-1,j) + U^n(i,j+1) + U^n(i,j-1).$$

Si noti che per motivi di robustezza al rumore e' stato sostituito, nel calcolo di R ed R_x , U con U^* . Si e' inoltre posto $h=1$ per poterla eliminare dalle formule, senza comunque perdere in generalita'.

Le espressioni precedenti sono valide nelle regioni in cui $W=1$. Per dare una formulazione generale valida in qualsiasi punto poniamo

$$U^*(i,j) = \frac{1}{\alpha} \left[U(i+1,j)W(i+1,j) + U(i-1,j)W(i-1,j) + U(i,j+1)W(i,j+1) + U(i,j-1)W(i,j-1) \right]$$

a condizione che $\alpha \neq 0$ dove

$$\alpha = W(i+1,j) + W(i-1,j) + W(i,j+1) + W(i,j-1)$$

A queste condizioni e' facile vedere che i valori di disparita' relativi a punti separati da un contorno sul quale $W(x,y)=0$, non hanno alcuna influenza uno sull'altro.

In questo modo sara' quindi possibile ricostruire le superfici della scena in modo corretto anche nelle zone dove sono presenti discontinuita'.

Analisi delle occlusioni.

Per rendere piu' complete le simulazioni del processo di fusione stereoscopica e' stato a questo punto affrontato il problema delle occlusioni. Bisogna tenere presente, infatti, che in ciascuna delle due immagini sono generalmente presenti alcune zone che non compaiono nell'altra immagine. L'esistenza di queste zone e' dovuta alla presenza, nella scena osservata, di forti variazioni di distanza dall'osservatore da parte delle superfici degli oggetti[8].

Questo fatto si verifica principalmente nelle regioni adiacenti ai bordi degli oggetti dove la distanza delle superfici dall'osservatore subisce una brusca variazione. Non e' comunque necessario che sia presente una discontinuita' nelle superfici, affinche' si formi una zona di occlusione: e' sufficiente una variazione piuttosto accentuata.

Un altro caso di zone di occlusione si verifica in corrispondenza dei bordi laterali di ciascuna delle due immagini stereo. La limitazione del campo visivo, infatti, fa si che, ad esempio, dal punto di osservazione destro non si vedano le regioni poste all'estremita' sinistra dell'altra immagine.

La conseguenza di tutto cio' e' che ci saranno punti, in ciascuna delle due immagini, per cui non ci sara' corrispondenza stereoscopica con nessun punto dell'altra immagine. In termini matematici questo vuol dire che esistono alcune regioni delle immagini in cui la funzione di disparita' stereoscopica non e' definita.

Vediamo allora la relazione esistente tra zone di occlusione e variazioni di distanza delle superfici rispetto all'osservatore. Per semplicita' ci limitiamo a trattare gli effetti dovuti alle discontinuita', assumendo che l'ampiezza delle zone di occlusione in corrispondenza di superfici continue ma molto ripide sia

trascurabile. Questo e' vero se si suppone che l'osservatore si trovi abbastanza distante dalle superfici.

Facciamo l'ipotesi che i due punti di osservazione si trovino sul piano x-z e che i due relativi schermi siano paralleli al piano x-y. In tali condizioni, la disparita' stereoscopica tra punti delle due immagini proiettate sui due schermi sara' calcolata soltanto rispetto l'asse x.

Supponiamo di avere due punti $p_1=(x_1,y,z_1)$ e $p_2=(x_2,y,z_2)$, e che sullo schermo sinistro vengano proiettati rispettivamente nei punti (x_1^s,y) e (x_2^s,y) .

Per come e' definita la disparita' stereoscopica le coordinate dei due punti proiettati sullo schermo destro saranno rispettivamente

(x_1^d,y) e (x_2^d,y) , dove $x_1^d=x_1^s-u_1$ e $x_2^d=x_2^s-u_2$.

Un caso che ci interessa particolarmente e' quello rappresentato dalla seguente figura.

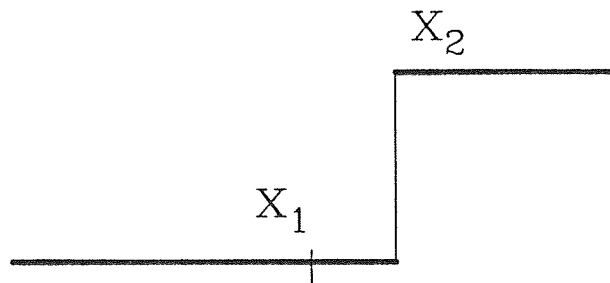


Figura 5.

Abbiamo qui mostrato una porzione della sezione della scena su un piano parallelo alla direzione dell'osservatore. In questo caso p_2 si trova sul bordo dell'oggetto posto davanti ad un piano, mentre p_1 e' un punto del piano stesso vicino all'oggetto.

Possiamo facilmente ricavare che

$$x_2^d - x_1^d = x_2^s - x_1^s - (u_2 - u_1) \quad (1)$$

Essendo p_2 piu' vicino all'osservatore la sua disparita' binoculare sara' maggiore di quella di p_1 , in quanto essa e' inversamente proporzionale alla distanza.

Da questo se ne deduce che

$$(u_2 - u_1) > 0,$$

e quindi

$$x_2^d - x_1^d < x_2^s - x_1^s$$

cioe' passando dall'immagine sinistra all'immagine destra, le proiezioni dei due punti si avvicinano tra di loro.

Vogliamo ora determinare le condizioni per cui

$$x_2^d - x_1^d < 0.$$

La situazione in cui $x_2^s > x_1^s$ e $x_2^d < x_1^d$, implica che nel passaggio dall'immagine sinistra all'immagine destra il punto p_1 e' stato "scavalcato" dal punto p_2 , come pure da altri punti nell'intorno destro del punto p_2 . La conseguenza e' che dal punto di osservazione destro il punto p_1 non sara' piu' visibile, in quanto tra esso e lo schermo si e' inserita una parte dell'oggetto antistante.

In particolare fissato p_2 vogliamo trovare tutti i punti p_1 per cui $x_2^d - x_1^d < 0$.

Supporremo per semplicita' che i punti p_1 siano su un piano perpendicolare alla direzione dell'osservatore. In queste ipotesi

essi avranno tutti la medesima disparita' stereoscopica u_1 .

Dalla (1) vediamo che questo si verifica quando $x_2^s - x_1^s < (u_2 - u_1)$.
Tenendo conto che stiamo considerando gli x_1^s alla sinistra di x_2^s otteniamo infine che

$$x_2^d - x_1^d < 0 \quad \text{per } x_1^s \in (x_2^s - (u_2 - u_1), x_2^s).$$

Questo significa che immediatamente alla sinistra di x_2^s c'e' un segmento di ampiezza $(u_2 - u_1)$ che viene occluso dall'oggetto posto davanti ad esso.

Si noti che se $x_2^s - x_1^s > 0$ e $x_2^d - x_1^d < 0$ e trovandosi quindi p_1 in una zona di occlusione, il valore u_1 non puo' piu' definirsi come disparita' binoculare del punto p_1 , in quanto qui questa funzione non e' definita.

Nell'implementazione dell'algoritmo queste regole vengono applicate nel modo seguente.

Si definisce una matrice C che chiameremo matrice indicatrice di occlusione che conterra' informazioni sul fatto che un dato punto dell'immagine si trovi o no in una zona di occlusione. In particolare essa sara' definita nel modo seguente:

$$C(i,j) = \begin{cases} 0 & \text{se } (i,j) \text{ e' un punto occluso} \\ 1 & \text{altrimenti} \end{cases}$$

Ogniqualevolta si ha un punto in cui la funzione di controllo di continuita' e' posta ad 0, con $u_2 > u_1$, si pone ad 0 la matrice indicatrice di occlusione nei punti immediatamente a sinistra che coprono una striscia di ampiezza $(u_2 - u_1)$.

Si fa poi in modo che i punti di queste zone non diano alcun

contributo nel calcolo delle derivate e nel calcolo dei valori del funzionale.

Per ottenere questo si ridefinisce l'espressione di U^* precedentemente descritta. In particolare si pone

$$U^*(i, j) = \frac{1}{\alpha} \left[\begin{aligned} &U(i+1, j)W(i+1, j)C(i+1, j) + \\ &+ U(i-1, j)W(i-1, j)C(i-1, j) + \\ &+ U(i, j+1)W(i, j+1)C(i, j+1) + \\ &+ U(i, j-1)W(i, j-1)C(i, j-1) \end{aligned} \right]$$

a condizione che $\alpha \neq 0$ dove

$$\alpha = W(i+1, j)C(i+1, j) + W(i-1, j)C(i-1, j) + \\ + W(i, j+1)C(i, j+1) + W(i, j-1)C(i, j-1)$$

Questa e' una ulteriore estensione analoga a quella fatta quando e' stata introdotta la matrice di controllo di continuita'. In questo caso, oltre ai punti di discontinuita', si annulla il contributo dei punti situati nelle zone di occlusione.

Vediamo ora i passi seguiti dall'algoritmo che tiene conto delle occlusioni.

Dopo aver calcolato una funzione di controllo di continuita' a partire dalla soluzione completamente "smooth" si eseguono delle iterazioni per determinare la corrispondente funzione di disparita'. A questo punto, utilizzando la funzione di disparita' e la funzione di controllo di continuita' si determinano le zone di occlusione corrispondenti. Adesso, tenendo conto delle occlusioni determinate si calcola una nuova funzione di disparita'. Anche qui, come nel caso in cui si determinava la sola funzione di controllo di continuita', il procedimento viene iterato per un sufficiente numero di stadi nel corso dei quali si calcolano le tre funzioni.

Simulazioni del modello.

Gli algoritmi fin qui descritti sono stati implementati e sono stati utilizzati per eseguire delle simulazioni su calcolatore. La coppia di immagini stereo utilizzate e' rappresentata in figura 6.

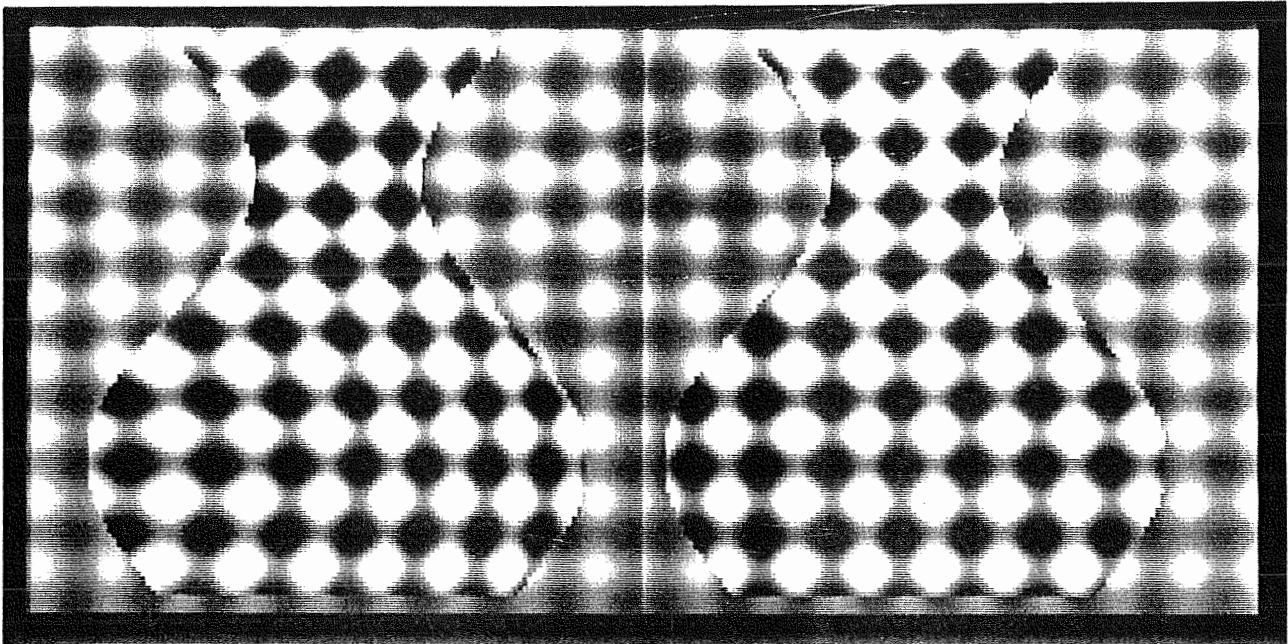


Figura 6.

Nel grafico in figura 7 vediamo la ricostruzione completamente "smooth" ottenuta nel primo stadio quando non vi e' ancora alcuna conoscenza ne' sulla matrice di controllo di continuita', ne' sulla matrice indicatrice di occlusione. Si vede chiaramente come le discontinuita' vengano distrutte dall'effetto di interpolazione dovuto al vincolo di "smoothness".

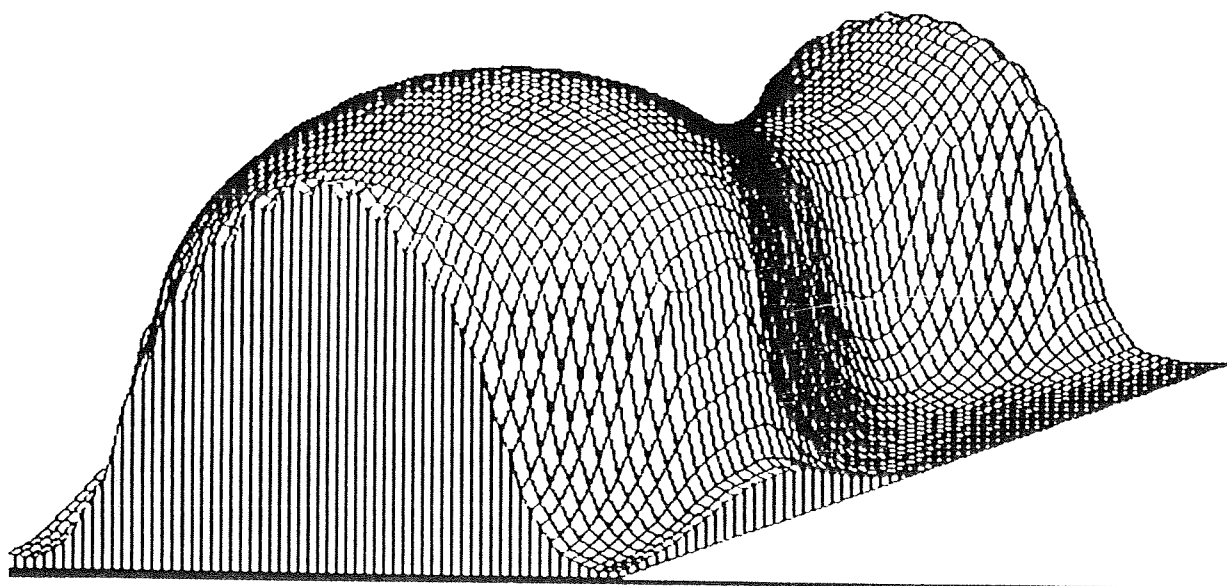


Figura 7.

Quella che segue, invece, e' la ricostruzione ottenuta calcolando nel corso delle iterazioni la matrice di controllo di continuita'. Manca pero' il controllo sulle zone di occlusione, e, come si vede dalla figura 8, la ricostruzione ne risente risultando cosi' un andamento non corretto.

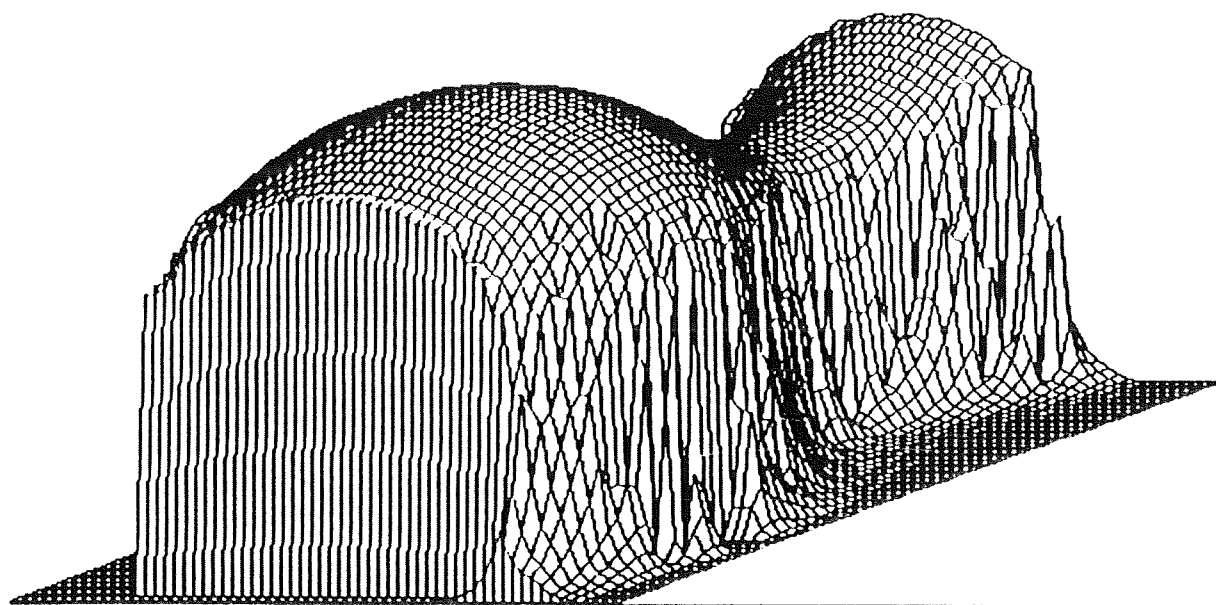


Figura 8.

Abbiamo allora applicato il criterio che prevede il trattamento delle zone di occlusione. Il risultato, come possiamo vedere nella figura 9, e' adesso considerabilmente migliorato.

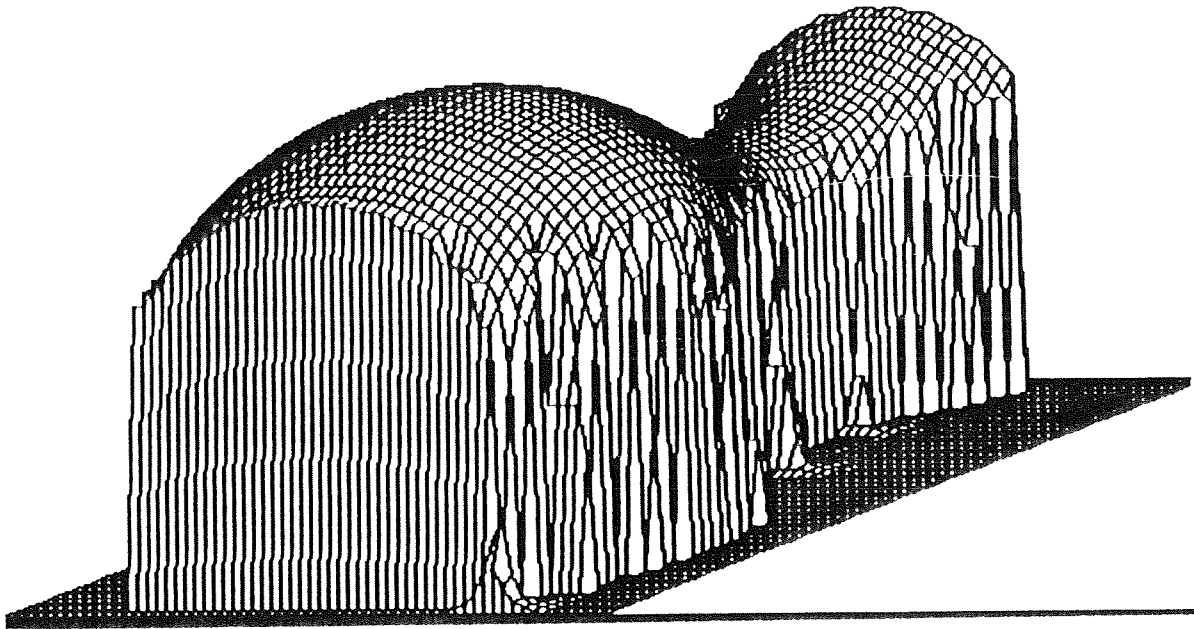


Figura 9.

Nelle immagini di figura 10 e 11 vengono illustrate rispettivamente la matrice di controllo di continuita' e la matrice indicatrice di occlusione.

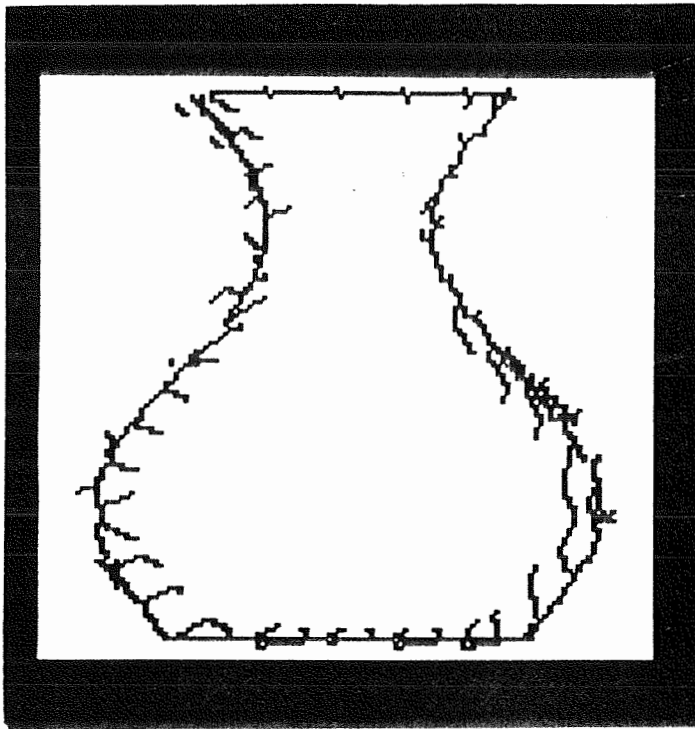


Figura 10.

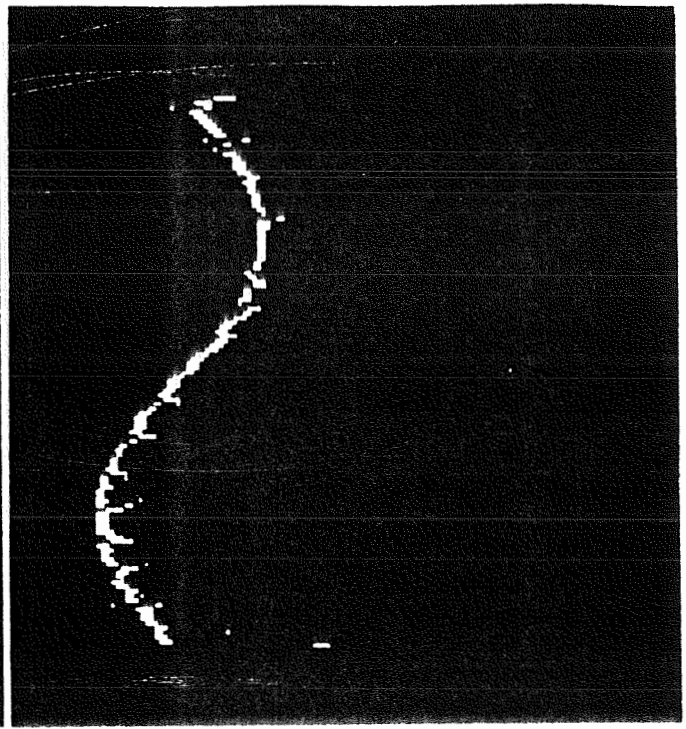


Figura 11.

Risulta evidente la difficoltà che si incontra nel determinare l'esatta funzione di controllo di continuità. La particolare scena in esame, d'altronde, ha contribuito a rendere più complicato il procedimento. In prossimità dei bordi dell'oggetto la superficie dello stesso ha un gradiente molto elevato, tendendo all'infinito man mano che ci si avvicina ai bordi.

Questo rende molto arduo trovare un valore di soglia che permetta di distinguere le discontinuità dalle intrinseche forti variazioni di orientamento delle superfici. Guardando la coppia di immagini stereo, inoltre, ci si rende facilmente conto che, oltre a quelle relative ai contorni dell'oggetto, sono presenti una elevata quantità di curve di zero-crossings. Diverse di queste curve sono posizionate anche vicino ai bordi dove la superficie possiede valori del gradiente molto elevati. Se si considera il criterio utilizzato per determinare le discontinuità, è facile a

questo punto capire la difficoltà' del problema ed il conseguente risultato.

La qualità' della matrice di controllo di continuità' ha influito anche sulla matrice indicatrice di occlusioni, essendoci tra loro una stretta dipendenza. Anche in questo caso però' le caratteristiche fondamentali delle zone di occlusione sono state determinate.

Il risultato complessivo è' interessante, considerando che la ricostruzione finale è' molto corretta. Bisogna tenere conto, d'altra parte, che quando nella funzione di controllo di continuità' è' presente il valore nullo, la funzione di disparità' stereoscopica può' avere un salto di ampiezza arbitraria, ma ciò' non è' obbligatorio. L'algoritmo può' quindi costruire superfici "smooth" anche in prossimità' di punti in cui la funzione di controllo di continuità' è' 0, se il vincolo di corrispondenza dei dati lo richiede. Ciò' spiega il motivo per cui spesso, malgrado la funzione di controllo di continuità' abbia qualche difetto, la ricostruzione della scena è' risultata invece molto più' soddisfacente.

Conclusioni.

E' stato qui studiato il processo di visione stereoscopica esprimendolo per mezzo di un modello matematico. Cio' ha richiesto l'uso di metodi di regolarizzazione[15] in quanto i problemi di visione sono, in genere, matematicamente mal posti.

I risultati ottenuti dalle simulazioni hanno dimostrato che la teoria della regolarizzazione rappresenta un valido approccio nella ricostruzione di immagini tridimensionali a partire da immagini stereo[3][4][9].

In particolare l'estensione della teoria della regolarizzazione standard da parte di Terzopoulos[13], che consente di trattare anche le discontinuita', ha rappresentato un passo in avanti. Cio' ha consentito la ricostruzione di superfici discontinue, eliminando cosi' il pesante vincolo che avrebbe, altrimenti, costretto a ricostruire correttamente solo scene in cui fosse presente una singola superficie.

Nel caso in cui la funzione di controllo di continuita' non e' nota a priori, la formulazione del funzionale richiede l'espressione di ulteriori vincoli che caratterizzino le sue proprieta'. La conseguenza e' che si sarebbe dovuto fare ricorso a tecniche di risoluzione piu' sofisticate per le equazioni differenziali associate al funzionale. Queste tecniche richiedono, in genere, algoritmi di una non indifferente complessita' computazionale. Si pone, inoltre, il problema, tutt'altro che banale, di determinare correttamente i vincoli opportuni sulla funzione di controllo di continuita'.

In base a queste considerazioni e' stato ritenuto opportuno utilizzare un metodo alternativo di piu' semplice attuazione, ottenendo cosi' un buon compromesso tra semplicita', prestazioni e precisione della ricostruzione[5].

E' stato quindi usato un metodo per calcolare la funzione di controllo di continuita', senza doverla cosi' supporre nota a priori. I risultati sono stati soddisfacenti e potranno essere migliorati utilizzando algoritmi di edge detection piu' sofisticati e metodi di approssimazione piu' precisi.

Lo scopo di questo lavoro era quello di affrontare il problema delle occlusioni. E' stato mostrato come anche queste, analogamente alla funzione di controllo di continuita', possono essere determinate nel corso delle iterazioni.

E' stata cosi' ottenuta una formulazione piu' accurata del modello matematico relativo al problema della visione stereoscopica.

In assenza della funzione indicatrice di occlusioni, infatti, l'algoritmo tenta di trovare, per ogni punto di un'immagine, un punto ad esso corrispondente nell'altra immagine. E' chiaro che, per i punti delle zone di occlusione, il valore della funzione di disparita' stereoscopica cosi' calcolato non e' significativo. Inoltre puo', a causa della propagazione imposta dal vincolo di smoothness, influire negativamente sulla qualita' della ricostruzione nelle regioni vicine.

Dalle simulazioni eseguite e' emerso che il ricorso ad una funzione indicatrice di occlusioni migliora considerabilmente la qualita' della scena tridimensionale ricostruita.

Il contesto in cui si pone questo lavoro e' quello di effettuare uno studio delle possibilita' offerte dalla teoria della regolarizzazione nell'ambito della visione stereoscopica, e piu' in generale nel settore della Computer Vision.

Vi sono ancora, come in parte gia' accennato, diversi aspetti da prendere in considerazione per permettere ulteriori sviluppi. Alcuni di questi, come ad esempio la formulazione di un funzionale

che prevede la funzione di controllo di continuita' come incognita, e l'uso di metodi numerici piu' sofisticati, sono gia' stati affrontati[6].

Anche alla luce dei risultati ottenuti da March[6], gli esperimenti eseguiti sulle occlusioni possono assumere un ruolo di una certa importanza.

Seguendo questo esempio, infatti, anche il problema delle occlusioni puo' essere impostato in maniera piu' precisa, determinando un funzionale che preveda la funzione indicatrice di occlusione come incognita. Va considerato, ovviamente, il problema non banale di esprimere dei vincoli consistenti sulle caratteristiche della funzione indicatrice di occlusione. In questo modo il modello matematico, e di conseguenza la ricostruzione da esso ottenuta, potrebbe beneficiare di considerevoli miglioramenti. Chiaramente con una tale impostazione, e' possibile che si renda necessario l'uso e lo sviluppo di tecniche matematiche piuttosto sofisticate. I risultati che si otterrebbero, comunque, sarebbero rilevanti e potrebbero giustificare il ricorso a tali metodi.

BIBLIOGRAFIA

- [1] Giusti
"Analisi Matematica 2"
Boringhieri, Torino, 1983

- [2] Hildebrand
"Finite difference equations and simulation"
Prentice Hall, Englewood Cliffs, 1968

- [3] Koch C. , Poggio T. , Torre V.
"Computational vision and regularization theory"
Nature , vol. 317 , pg. 314-319 , Settembre 1985

- [4] March R.
"Computation of stereo disparity using regularization"
Pattern Recognition Letters , vol. 8 , no. 3,
pg. 181-187 , 1988

- [5] March R.
"A regularization model for stereo vision with
controlled continuity"
Pattern Recognition Letters , Vol. 10, no. 4,
pg. 259-263, 1989

- [6] March R.
"Variational problems arising in early vision:
Approximation by Γ -convergence"
Nota Interna B4-55 dell'Istituto di Elaborazione
delle Informazioni del CNR di Pisa, Dicembre 1990,
in corso di pubblicazione su Image and Vision Computing.

- [7] Marr D. , Hildreth E.
"Theory of edge detection"
Proc. R. Soc. Lond. B 207 , pg. 187-217 , 1980
- [8] Minutoli S.
"Teoria della regolarizzazione e visione stereoscopica"
Tesi di Laurea in Scienze dell'Informazione, 1990
Universita' degli Studi di Pisa
- [9] Mori M.
"Ricostruzione automatica di superfici discontinue da
immagini stereo"
Tesi di Laurea in Scienze dell'Informazione , 1988
Universita' degli Studi di Pisa
- [10] Poggio T.
"La visione nell'uomo e nella macchina"
Quaderni delle scienze , no. 26 , 1985
- [11] Poggio T.
"Early vision : from computational structure to
algoritms and parallel hardware"
Computer Vision, Graphics, and Image Processing 31,
pg. 139-155 , 1985
- [12] Poggio T. , Torre V.
"On edge detection"
IEEE Trans. on PAMI , vol. 8 , no. 2 , pg. 147-163
Marzo 1986

- [13] Terzopoulos D.
"Regularization of inverse problems involving
discontinuities"
IEEE Trans. on PAMI , vol. 8 , no. 4 , Luglio 1986
- [14] Terzopoulos D.
"The computation of visible-surface representation"
IEEE Trans. on PAMI , vol. 10 , no. 4 , Luglio 1986
- [15] Tikhonov A.N., Arsenin V.Y.
"Solutions of ill-posed problems"
Winston, Washington, 1977