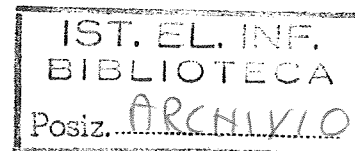


Consiglio Nazionale delle Ricerche

**ISTITUTO DI ELABORAZIONE
DELLA INFORMAZIONE**

PISA



84-27

**METODI ITERATIVI PER LA GENERAZIONE NUMERICA
DI OLOGRAMMI MEDIANTE COMPUTER**

S. MOGLIE, E.A. SALERNO, O. SALVETTI

Nota Interna B4 - 22

Luglio 1991

METODI ITERATIVI PER LA GENERAZIONE NUMERICA DI OLOGRAMMI MEDIANTE COMPUTER

S. Moglie*, E.A. Salerno**, O. Salvetti**

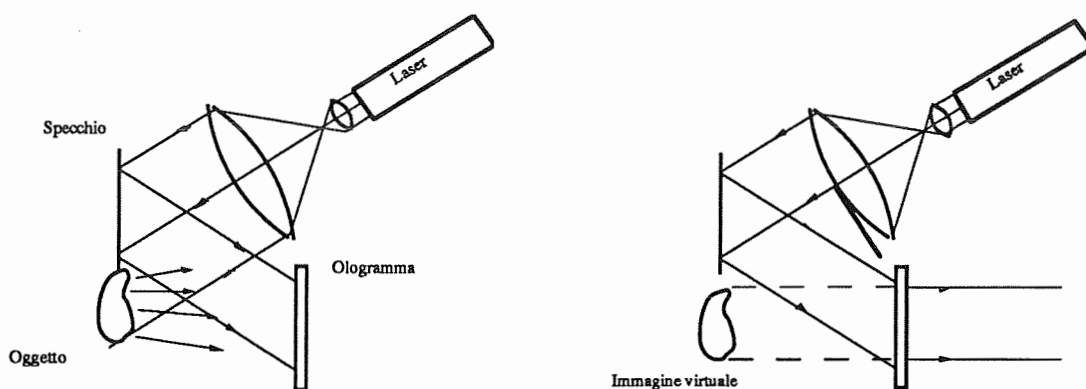
* Corso di Laurea in Scienze dell'Informazione -
Facoltà di Scienze MFN - Università di Pisa
Corso Italia, 40 56100 Pisa

** Istituto di Elaborazione della Informazione - CNR
Via S.Maria, 46 56126 Pisa

1. Introduzione

Con la presente nota, si intende dare una breve introduzione alla tecnica dell'olografia ottica e successivamente affrontare l'argomento della *Computer Holography*, cioè la possibilità di generare ologrammi mediante il calcolo numerico. L'olografia è una tecnica di rappresentazione *totale* di un oggetto, che consente la registrazione di tutte le informazioni ottiche utili alla ricostruzione di una sua immagine tridimensionale. Tale tecnica, basandosi sull'ottica coerente, permette di registrare e ricostruire fronti d'onda. Nella fotografia convenzionale vengono registrate su pellicola solo le intensità dell'onda incidente, che vengono convertite nelle corrispondenti variazioni di opacità dell'emulsione fotografica, e le informazioni relative alla fase vengono perse perché il materiale fotosensibile non reagisce alle variazioni di fase. La tecnica olografica consiste invece nel far interferire l'onda trasmessa o diffusa da un oggetto (onda oggetto), con l'onda *coerente* direttamente emessa dalla sorgente (onda di riferimento), trasformando così l'informazione di fase in una variazione di intensità, registrabile su lastra fotografica. La figura risultante da questa interferenza sarà un insieme di frange le cui irregolarità dipendono dal fronte dell'onda oggetto (in particolare, contrasto e spaziatura delle frange dipendono rispettivamente da ampiezza e fase dell'onda riflessa dall'oggetto). Rimuovendo l'oggetto reale, il fronte d'onda diffratto dall'ologramma illuminato dall'onda di riferimento coinciderà con il fronte d'onda dell'oggetto originale e,

guardando attraverso l'ologramma come fosse una finestra, appare un'immagine virtuale dell'oggetto in forma tridimensionale [12] [13] [19]. L'ideazione dell'olografia ottica risale al 1948 ed è dovuta al fisico ungherese Dennis Gabor [8], il quale venne per questo insignito del premio Nobel nel 1971, quando ormai la sua intuizione, grazie all'introduzione del laser, stava concretizzandosi nelle prime applicazioni pratiche. Uno schema tipico per la formazione dell'ologramma e la ricostruzione dell'immagine è mostrato in Figura 1.



a) Formazione dell'ologramma; b) Ricostruzione dell'immagine

Figura 1

L'olografia ottica si è venuta sviluppando essenzialmente lungo tre filoni [21]: display olografico, ottiche olografiche, interferometria olografica.

Il *Display Olografico* consiste nella visualizzazione tridimensionale dell'immagine registrata. Questo campo si è particolarmente sviluppato in seguito alla possibilità di restituzione dell'ologramma per mezzo di comuni sorgenti luminose, anziché con laser; le relative applicazioni riguardano essenzialmente i campi della creazione artistica, grafica pubblicitaria e anticontraffazione.

Le *Ottiche Olografiche* sono elementi ottici che agiscono sul fascio luminoso su cui operano per via diffrattiva anziché rifrattiva come le ottiche tradizionali. Esse servono principalmente a correg-

gere le aberrazioni. La loro realizzazione avviene a partire da una registrazione fisica del sistema di interferenza impiegando fasci laser opportunamente modellati, oppure dalla generazione al computer del reticolo diffrattivo, che si può poi riportare sul supporto fisico mediante varie tecnologie (Computer Holography). I vantaggi che le ottiche olografiche presentano rispetto alle ottiche tradizionali si possono individuare in:

- minor peso;
- minor ingombro;
- maggior compattezza (più funzioni possono essere svolte contemporaneamente dal medesimo componente);
- minor costo
- elevata efficienza di diffrazione (è il rapporto tra l'intensità del fascio di luce emergente dall'ologramma e l'intensità del fascio di luce incidente) [5];
- selettività spettrale.

L'*Interferometria Olografica* viene applicata alla soluzione di problemi di controllo qualità. Essa si fonda sulla possibilità di confrontare sulla medesima lastra fotografica una sequenza di immagini dello stesso oggetto leggermente differenti fra loro. La valutazione dell'andamento e delle entità delle differenze fra le diverse immagini è resa agevole per l'osservatore perché l'immagine olografica dell'oggetto così realizzata risulta solcata da un insieme di bande chiare e scure (frange di interferenza), la cui forma e dimensioni sono diretta espressione delle suddette differenze.

Le applicazioni dell'olografia si sono fatte più vaste anche perché un ologramma, come già accennato, può essere generato da computer. Gli ologrammi generati da computer (*Computer Generated Holograms* = CGH) furono introdotti da Brown e Lohmann nel 1966 [3]. Gli scopi per cui sono nati sono legati alla soluzione di problemi di filtraggio spaziale ottico. Un importante aspetto dei CGH sta nel fatto che non è necessario che l'oggetto esista realmente, se ne è disponibile una descrizione analitica o numerica. Nel seguito, a seconda del tipo di applicazione, con la parola *oggetto* si indicherà indifferentemente l'oggetto fisico da cui

emerge l'onda oggetto incidente sull'ologramma, o il fascio luminoso che si vuole emerga dall'ologramma da sintetizzare quando questo è attraversato da uno specificato fascio di riferimento. Nel primo caso la sintesi dell'ologramma deriverà dalla soluzione di un problema di scattering diretto, mentre nel secondo caso, per esempio nei problemi di sintesi di ottiche olografiche, il problema della computer holography è un problema di Scattering Inverso, e quindi sarà accompagnato dalle note difficoltà legate alla generale malposizione di tale tipo di problema [2]. La ricerca in questo campo ha riguardato i metodi di sintesi dei CGH, gli errori derivanti dalle approssimazioni adottate e le strategie per minimizzarli.

Gli impieghi della computer holography riguardano tutte quelle applicazioni in cui siano necessarie manipolazioni di fronti d'onda luminosi, come, ad esempio, le verifiche su elementi ottici rifrattivi e la sintesi di nuove ottiche di tipo diffrattivo. Nella presente nota ci si occupa soltanto della generazione di ottiche olografiche. Nel prossimo paragrafo sono descritte le caratteristiche di vari tipi di CGH e i problemi relativi alla loro sintesi, mentre al Paragrafo 3 sono riportati due algoritmi iterativi, entrambi derivati dal classico metodo di Gerchberg e Saxton, utilizzati per la soluzione di alcuni di questi problemi. Come esempio, si riporta poi l'applicazione alla minimizzazione dell'errore di quantizzazione. L'estesa bibliografia riportata ha lo scopo di permettere l'approfondimento di singole parti del vasto argomento qui trattato; testi base per la comprensione dei fenomeni fisici legati all'olografia possono essere considerati [12] [13] [19], mentre per quanto riguarda gli aspetti matematici specificamente legati alla computer holography possono essere consultati [14] [16] [22] [23].

2. Sintesi dei CGH

Un ologramma può essere considerato come un sistema ottico formato come una semplice maschera, con una certa funzione di trasmittanza reale e positiva. Se tale maschera è stata realizzata per mezzo di un display digitale la sua trasmittanza avrà anche valori quantizzati; supponiamo che i livelli di quantizzazione siano

solo due e quindi la trasmittanza sia una funzione binaria. In questo caso, il problema della sintesi di un ologramma digitale è di cercare una funzione di trasmittanza tale che, illuminando l'ologramma con il fascio luminoso di riferimento, una porzione del fronte d'onda emergente assuma la forma desiderata ad una certa distanza dall'ologramma stesso. Riportiamo qui di seguito una breve trattazione matematica, seguita dall'esame di alcuni problemi legati alla fabbricazione dei CGH [14].

Con riferimento alla Figura 2, immaginiamo un ologramma posto in un piano di coordinate (u,v) , illuminato da uno specifico fronte d'onda. L'ologramma è formato da un array di $M \times N$ (M e N sono qui considerati pari) celle indirizzabili di dimensioni $R \times S$, ed ha trasmittanza binaria data da:

$$H(u,v) = \sum_{k=-M/2}^{M/2-1} \sum_{l=-N/2}^{N/2-1} H_{kl} \text{rect} \left(\frac{u - kR}{R}, \frac{v - lS}{S} \right) \quad (1)$$

in cui

$$\text{rect}(a,b) = \begin{cases} 1 & \text{se } |a|, |b| < 1/2 \\ 0 & \text{altrimenti} \end{cases} \quad (2)$$

e H_{kl} è la trasmittanza binaria della (k,l) -esima cella.

Ad una distanza z dall'ologramma, in un piano $x-y$ (detto *piano di osservazione*), il fronte d'onda diffratto ha ampiezza complessa $h(x,y)$. Supponiamo, seguendo [14], che il piano di osservazione si trovi in zona di Fraunhofer [12] (o di campo lontano) rispetto all'ologramma; in questo caso la propagazione dell'onda tra il piano $u-v$ e il piano $x-y$ è descritta dall'antitrasformata di Fourier analogica bidimensionale di $H(u,v)$:

$$h(x,y) = \iint H(u,v) \exp \left\{ -i2\pi(ux + vy) \right\} du dv =$$

$$= RS \operatorname{sinc}(Rx, Sy) \sum_{k=-M/2}^{M/2-1} \sum_{l=-N/2}^{N/2-1} H_{kl} \exp[i2\pi(kRx + lSy)] \quad (3)$$

in cui

$$\operatorname{sinc}(x, y) = \frac{\sin(\pi x) \sin(\pi y)}{\pi^2 x y} \quad (4)$$

Le frequenze spaziali u e v saranno legate alle effettive coordinate sul piano dell'ologramma da una costante di proporzionalità dipendente dalla distanza z e dalla lunghezza d'onda della radiazione luminosa utilizzata [12] [18] [23].

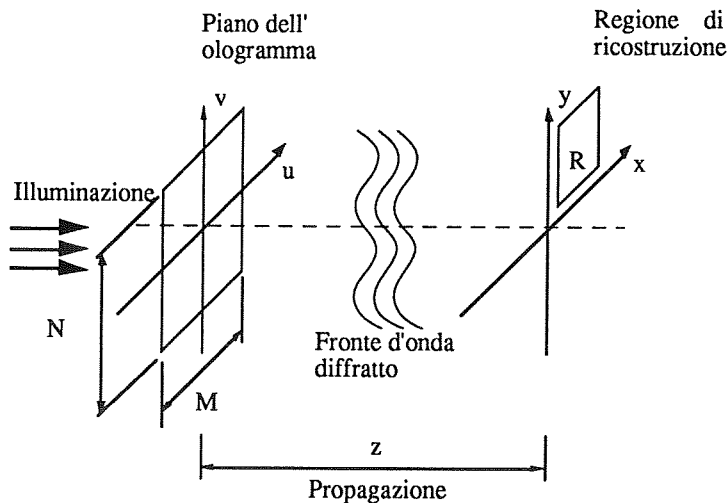


Figura 2

Campionando $h(x, y)$ nei punti $(x, y) = (mX, nY)$, con $X=1/MR$, $Y=1/NS$ (intervallo di campionamento di Nyquist) si ottiene:

$$h(mX, nY) = \sqrt{\frac{RS}{XY}} \operatorname{sinc}\left(\frac{m}{M}, \frac{n}{N}\right) h_{mn} \quad (5)$$

dove h_{mn} è la trasformata di Fourier discreta (DFT) inversa della sequenza doppia H_{kl} :

$$h_{mn} = \frac{1}{\sqrt{MN}} \sum_{k=-M/2}^{M/2-1} \sum_{l=-N/2}^{N/2-1} H_{kl} \exp \left[i2\pi \left(\frac{mk}{M} + \frac{nl}{N} \right) \right] \quad (6)$$

Il nostro obiettivo è quello di trovare una h_{mn} che uguagli una funzione specificata, \tilde{f}_{mn} , in una regione rettangolare R contenente $A \times B$ campioni. La notazione adottata (con l'uso della tilde) sta ad indicare che \tilde{f}_{mn} è stata scalata in modo tale che la sua DFT \tilde{F}_{kl} abbia picco di ampiezza unitaria. Ricordando che \tilde{f}_{mn} è l'oggetto e h_{mn} la sua immagine ricostruita, consideriamo l'errore quadratico medio:

$$e = \frac{1}{AB} \sum_{(m,n) \in R} \sum | \tilde{f}_{mn} - \lambda h_{mn} |^2 \quad (7)$$

dove la costante λ è scelta in modo da minimizzare l'errore tra h_{mn} ed un fissato \tilde{f}_{mn} . Si ottiene [14]:

$$\lambda = \frac{\sum_{(m,n) \in R} \sum \tilde{f}_{mn} h_{mn}^*}{\sum_{(m,n) \in R} \sum | h_{mn} |^2}$$

in cui l'asterisco indica l'operazione di complesso coniugato.

Vogliamo inoltre che l'immagine ricostruita contenga la più grande frazione possibile dell'energia incidente sull'ologramma, cioè che l'efficienza di diffrazione sia la massima possibile. Consideriamo adesso l'efficienza di diffrazione di un ologramma ideale che riproduce esattamente la \tilde{f} nella regione R , convogliandovi tutta l'energia luminosa da cui è attraversato, e la corrispondente efficienza dell'ologramma binario che produce su R l'immagine h_{mn} . L'ologramma ideale dovrebbe avere trasmittanza

complessa $H_{kl} = \tilde{F}_{kl}$. Illuminandolo con un'onda piana di ampiezza unitaria la frazione di energia da questo trasmessa è:

$$\eta_{\text{obj}} = \frac{1}{MN} \sum_{(m,n) \in R} |\tilde{f}_{mn}|^2 \quad (8)$$

Poiché tutta l'energia trasmessa sarebbe in questo caso diffratta nella regione di ricostruzione, l'efficienza di diffrazione coinciderebbe con η_{obj} . Con un ologramma binario solo una parte dell'energia trasmessa verrà diffratta nella regione di ricostruzione R; la frazione dell'energia incidente sull'ologramma e diffratta in R è:

$$\eta_{\text{tot}} = \frac{1}{MN} \sum_{(m,n) \in R} |h_{mn}|^2 \quad (9)$$

Per misurare l'efficienza dell'ologramma binario rispetto all'ologramma reale si considera l'*efficienza di binarizzazione*:

$$\eta_{\text{bin}} = \eta_{\text{tot}} / \eta_{\text{obj}} \quad (10)$$

A questo punto la formalizzazione del problema della sintesi di CGH può considerarsi completa. L'obiettivo enunciato all'inizio del paragrafo, cioè la ricostruzione secondo la forma voluta di un fronte d'onda diffratto, si traduce così nella minimizzazione dell'errore quadratico medio e la massimizzazione dell'efficienza di binarizzazione η_{bin} .

I metodi tradizionali usati per la sintesi di CGH si basano sui seguenti due passi, schematizzati anche nel diagramma in Figura 3.

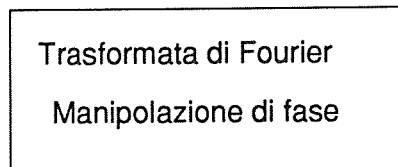
- *Simulazione* del fronte d'onda in un piano immediatamente a valle dell'ologramma, che in molti casi, come in quello precedentemente esposto, viene calcolata semplicemente attraverso la DFT. Per evitare errori dovuti al fatto che l'ampiezza complessa nel piano dell'ologramma non è soddisfacentemen-

te rappresentabile sul supporto fisico (per esempio per limitazioni tecnologiche) si ricorre spesso a manipolazioni della funzione di fase.

- *Codifica* della distribuzione di ampiezza complessa per ottenere valori reali e non negativi della trasmittenza. Spesso l'ologramma è codificato a valori binari poiché molti dispositivi hardware (come plotter e dispositivi che usano fotoresist) producono tipicamente uscite binarie.

Sintesi di CGH Metodi classici

1 - Simulazione della propagazione



2 - Codifica

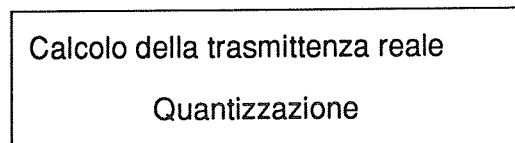


Figura 3

La qualità dell'immagine ricostruita potrebbe essere migliorata aumentando il numero di livelli di quantizzazione. Tuttavia, oltre un certo limite, tale aumento non migliora ulteriormente la qualità dell'immagine ricostruita. Sia nel caso binario che nel caso di quantizzazione a più livelli, il fascio luminoso sintetizzato per mezzo del CGH differirà da quello desiderato per l'*errore di quantizzazione*.

Accanto ai metodi tradizionali, esiste una classe di approcci diretti che, invece di ricorrere alla codifica dello spettro complesso,

manipolano direttamente la funzione di trasmittenza reale per minimizzare l'errore quadratico medio tra l'immagine ricostruita e quella desiderata. Come esempio di tecnica appartenente a questa classe, si espone qui di seguito uno di questi approcci: il metodo iterativo di *ricerca binaria diretta* (DBS) [14], [22]. Esso parte assegnando una configurazione binaria casuale all'array cartesiano $M \times N$ su cui si campiona l'ologramma, che viene poi scandito per righe o per colonne. Per ognuno dei punti dell'array, si modifica il valore della trasmittenza e si calcola l'errore quadratico medio tra il fascio così generato e quello desiderato. Se si ha una riduzione dell'errore rispetto alla situazione precedente, la modifica viene mantenuta, altrimenti si ripristina il valore iniziale. Si procede così finché tutto l'ologramma è scandito senza alcun cambio di valore. Questo algoritmo evita il passo della manipolazione dello spettro, ma introduce una elevata complessità computazionale, dovuta principalmente al calcolo dell'errore quadratico medio ripetuto $M \times N$ volte ad ogni iterazione. Ciò limita la sua applicabilità alla generazione di ologrammi di ridotte dimensioni, a meno di non avere a disposizione risorse di supercalcolo o macchine specializzate.

In base al metodo di codifica dell'ampiezza complessa i CGH possono essere raggruppati in categorie. Una classificazione può essere fatta a seconda che la tecnica di codifica segua oppure no il concetto di *detour-phase* [23].

Detour Phase Hologram

E' il classico metodo di registrazione introdotto da Brown e Lohmann. L'ologramma è formato calcolando la trasformata di Fourier dell'immagine e rappresentandone i valori in una maschera che presenta aperture trasparenti in uno schermo opaco. Tale schermo è suddiviso in un array di celle indirizzabili di ugual misura, su ognuna delle quali è praticata un'apertura la cui altezza codifica l'ampiezza della trasformata, mentre la fase dipende dalla sua posizione rispetto al centro della cella. Malgrado le approssimazioni il *detour-phase* è abbastanza utile in imaging e filtraggio

perché le distorsioni possono essere ridotte riducendo l'ampiezza delle celle, cioè infittendo il campionamento.

Non Detour Phase Hologram

Il metodo fu inizialmente sviluppato da Lee [9]. L'ologramma calcolato viene considerato come un interferogramma. La tecnica comporta il calcolo della figura di diffrazione che risulterebbe dall'interferenza del fronte d'onda desiderato con un'onda piana. Tale figura viene poi convertita in forma binaria in maniera tale da poter codificare sia l'ampiezza che la fase, agendo sulla posizione e la larghezza delle frange. Questa tecnica ha avuto riscontri sperimentali nella ricostruzione di fronti d'onda sferici, conici ed elicoidali [4], [15].

3. Due algoritmi iterativi per la sintesi di CGH

In questo paragrafo, vedremo due versioni di un metodo iterativo, applicabile alla sintesi di CGH con ridotti errori di aliasing e di quantizzazione [7]. Una delle due versioni è ottenuta dall'altra nel tentativo di accelerarne la convergenza e renderne più generale l'applicabilità; entrambe derivano dal metodo iterativo di Gerchberg e Saxton [10], [11], originariamente usato per ricostruire la fase nei domini dell'immagine e della trasformata di Fourier conoscendo il modulo dell'immagine e della sua trasformata. In generale si tratta di ricavare una funzione che soddisfi un certo numero di vincoli nel suo dominio di definizione e nello spazio della sua trasformata di Fourier. I due algoritmi sono qui designati con i nomi *error-reduction approach* e *input-output approach*.

3.1 Error-Reduction Approach

Come detto sopra l'*error-reduction approach* si usa per cercare una coppia di Fourier $f(x) \leftrightarrow F(u)$ che soddisfi determinati vincoli, posti nei domini oggetto e trasformato. Come è noto f e F sono legate dalla relazione:

$$F(u) = |F(u)| \exp[i\theta(u)] = \mathcal{F}[f(x)] = \int_{-\infty}^{\infty} f(x) \exp(i2\pi u \cdot x) dx \quad (11)$$

dove \mathcal{F} rappresenta l'operatore trasformata di Fourier, \mathbf{x} è il vettore delle coordinate (1, 2 o 3 dimensionali) nello spazio oggetto e \mathbf{u} è il vettore delle corrispondenti frequenze nello spazio trasformato.

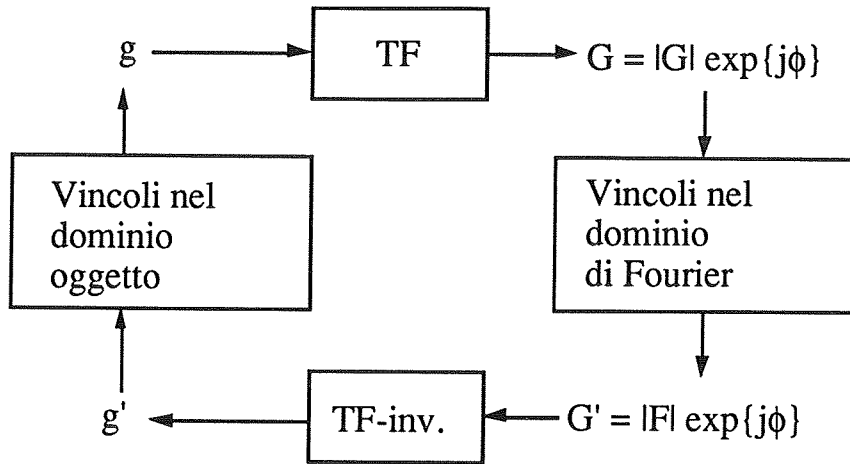


Figura 4

In Figura 4 è schematizzato l'algoritmo iterativo. Esso si sviluppa, con ripetute trasformazioni tra gli spazi oggetto e trasformato, forzando le funzioni a soddisfare i vincoli dati in entrambi i domini. La k -esima iterazione dell'algoritmo si sviluppa come segue:

- 1) Una stima dell'oggetto $g_k(\mathbf{x})$ viene trasformata secondo Fourier e produce

$$G_k(\mathbf{u}) = |G_k(\mathbf{u})| \exp[i\phi_k(\mathbf{u})] = \mathcal{F}[g_k(\mathbf{x})] \quad (12)$$

- 2) Ora, imponendo che $G_k(\mathbf{u})$ soddisfi certi vincoli nel dominio di Fourier, si ottiene una nuova funzione $G'_k(\mathbf{u})$ che è ottenuta da $G(\mathbf{u})$ facendo i più piccoli cambiamenti possibili. Per esempio, se, come assunto in Figura 4, il vincolo nel dominio di Fourier è che la trasformata abbia modulo uguale a $|F(\mathbf{u})|$, allora $G'_k(\mathbf{u})$ sarà data da:

$$G'_k(\mathbf{u}) = |F(\mathbf{u})| \exp[i\phi_k(\mathbf{u})] \quad (13)$$

Cioè, il modulo $|F(u)|$ è sostituito al modulo di $G_k(u)$ mentre la sua fase è lasciata invariata.

- 3) Il risultante $G'_k(u)$, che soddisfa i vincoli nel dominio di Fourier, è antitrasformato e produce una funzione $g'_k(x)$ nel dominio dell'oggetto.
- 4) L'iterazione è completata fornendo una nuova funzione $g_{k+1}(x)$ ottenuta imponendo che $g'_k(x)$ soddisfi i vincoli nel dominio dell'oggetto.

La prima iterazione può iniziare in più modi, per esempio, ponendo $g_1(x)$ o $\phi_1(u)$ uguale ad un array di numeri casuali.

Il processo iterativo si arresta quando trova una coppia di Fourier che soddisfi i vincoli in entrambi i domini. Una misura del miglioramento da un'iterazione alla successiva, sulla base della quale stabilire il criterio di arresto, è l'errore quadratico medio, definito come segue:

$$E_F^2 = \frac{\int_{-\infty}^{\infty} |G_k(u) - G'_k(u)|^2 du}{\int_{-\infty}^{\infty} |G'_k(u)|^2 du} \quad \text{nel dominio di Fourier} \quad (14)$$

$$E_O^2 = \frac{\int_{-\infty}^{\infty} |g_{k+1}(x) - g'_k(x)|^2 dx}{\int_{-\infty}^{\infty} |g'_k(x)|^2 dx} \quad \text{nel dominio oggetto} \quad (15)$$

Quando l'errore quadratico medio è nullo allora l'oggetto e la sua trasformata di Fourier soddisfano tutti i vincoli ed è stata trovata una soluzione.

E' stato dimostrato per un particolare problema [10] (e forse è vero anche in generale), che l'errore quadratico medio può solo decrescere dopo ogni iterazione; da ciò deriva il nome dato a

questo algoritmo. Per alcune applicazioni l'error-reduction approach è molto efficiente nel trovare una soluzione con un numero ragionevole di passi. Tuttavia, per altre applicazioni, dopo le prime poche iterazioni l'errore quadratico medio decresce molto lentamente, richiedendo così un tempo eccessivo per la convergenza.

3.2 Input-Output Approach

Questo metodo differisce dal precedente solo riguardo le operazioni nel dominio dell'oggetto. Se si raggruppano come in Figura 5 i primi tre passi di ogni iterazione, che sono gli stessi per entrambi i metodi, si ottiene un sistema non lineare che ha input $g(x)$ e output $g'(x)$.

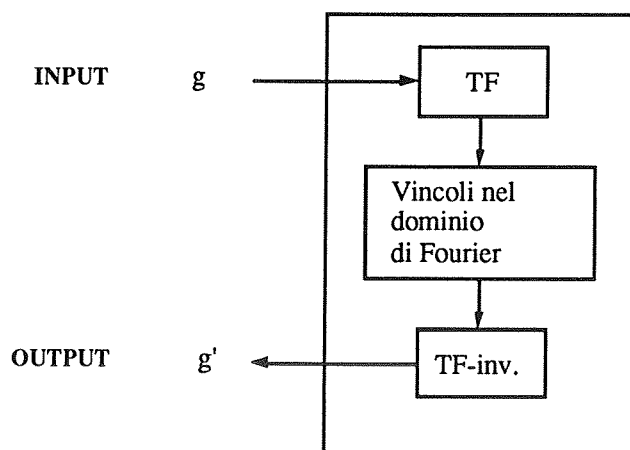


Figura 5

Proprietà di questo sistema è che l'output è sempre una funzione con trasformata di Fourier che soddisfa i vincoli assegnati, perciò se l'output soddisfa anche i vincoli del dominio dell'oggetto si è trovata una soluzione del problema.

Vediamo come bisogna manipolare l'input in modo tale che l'output soddisfi i vincoli del dominio dell'oggetto. Nell'error-reduction approach l'input è scelto come la migliore stima che soddisfi i vincoli nel dominio oggetto, mentre nell'input-output approach esso non è necessariamente una stima dell'oggetto o una modifica dell'output, e neppure deve soddisfare i vincoli; esso è semplicemente la *driving function* per il prossimo output. Il modo

in cui l'input deve essere cambiato affinché l'output soddisfi i vincoli, dipende dalla situazione. Vediamo cosa accade quando viene fatto un arbitrario cambiamento nell'input. Supponiamo che alla k -esima iterazione l'input $g_k(\mathbf{x})$ porti all'output $g'_k(\mathbf{x})$ e che l'input sia cambiato aggiungendo $\Delta g(\mathbf{x})$:

$$g_{k+1}(\mathbf{x}) = g_k(\mathbf{x}) + \Delta g(\mathbf{x}) \quad (16)$$

L'output $g'_{k+1}(\mathbf{x})$ sarà:

$$g'_{k+1}(\mathbf{x}) = g'_k(\mathbf{x}) + \alpha \Delta g(\mathbf{x}) + \textit{rumore addizionale} \quad (17),$$

cioè il valore aspettato del cambiamento nell'output, dovuto al cambiamento $\Delta g(\mathbf{x})$ nell'input, è proporzionale al cambio dell'input. Nonostante questo sia un sistema non lineare, cambiamenti nell'input tendono a portare cambiamenti simili nell'output. La mancata predicibilità dell'effettiva funzione output è rappresentata dal termine *rumore addizionale*. Il parametro α dipende dalle statistiche di $G_k(\mathbf{u})$ e $F(\mathbf{u})$ e dai vincoli nel dominio di Fourier. Se l'output $g'_k(\mathbf{x})$ non soddisfa i vincoli nel dominio oggetto ed esiste $\Delta g_d(\mathbf{x})$ tale che la funzione

$$g'_k(\mathbf{x}) + \Delta g_d(\mathbf{x}) \quad (18),$$

soddisfi tali vincoli, allora avremmo potuto cercare di ottenere un output che soddisfacesse i vincoli cambiando l'input in modo tale da causare nell'output un cambiamento $\Delta g_d(\mathbf{x})$. In accordo con le equazioni (16) e (17) il cambiamento dell'input causa un cambiamento nell'output legati dalla relazione:

$$\Delta g(\mathbf{x}) = \alpha^{-1} \Delta g_d(\mathbf{x}) \quad (19)$$

Una possibile scelta per il nuovo input è quindi:

$$g_{k+1}(x) = g_k(x) + \beta \Delta g_d(x) \quad (20)$$

con $\beta = \alpha^{-1}$. La funzione $\Delta g_d(x)$ è tale che $g'_k(x) + \Delta g_d(x)$ soddisfa i vincoli del dominio oggetto. Se α non è noto allora β potrà essere valutato solo approssimativamente. La scelta di β non influenza il successo dello schema iterativo. L'unico inconveniente sarà una convergenza più lenta se si adottano valori ridotti di β . Il termine di rumore nell'eq. (17) è minimo quando si minimizza $|\Delta g(x)|$. Nell'input-output approach, poiché non importa se $g_k(x)$ soddisfa i vincoli del dominio di Fourier, allora il valore di E_F^2 non è importante, mentre lo è quello di E_O^2 . Se l'output $g'(x)$ è usato come input allora il suo output sarà lo stesso $g'(x)$ poiché tale funzione soddisfa già i vincoli del dominio di Fourier. Da questo punto di vista, un'altra logica scelta per il successivo input è:

$$g_{k+1}(x) = g'_{k+1}(x) + \beta \Delta g_d(x) \quad (21)$$

Si noti che, se in (21) si fa $\beta=1$, i due approcci descritti vengono a coincidere. Poiché il valore ottimo di β è di solito diverso dall'unità, allora l'error-reduction approach può essere considerato una versione sub-ottima dell'input-output approach.

3.3 Esempio di applicazione: riduzione dell'errore di quantizzazione

In [7] sono riportati i risultati dell'applicazione dei due algoritmi al problema della riduzione del rumore di quantizzazione, così come si presenta nella computer holography. In questo paragrafo si mostra quali sono i vincoli da porre nei due domini per questo particolare problema e, due possibili strategie di aggiornamento della funzione in ingresso. Sia f la funzione a valori complessi che descrive l'immagine richiesta e $F = \mathcal{F}[f]$ la sua trasformata di Fourier, che dovrà essere codificata sull'ologramma. Il rumore di quantizzazione è dovuto alle limitazioni del display e dei materiali usati per sintetizzare l'ologramma che impediscono di rappresentare con continuità i coefficienti complessi di Fourier. Per esempio, gli ologrammi di tipo detour phase, quantizzano sia

l'ampiezza che la fase. La finezza di tale quantizzazione dipende dal numero di elementi di risoluzione contenuti in ogni cella. Altri tipi di ologrammi [17] [18], codificati in modo da permettere un controllo quasi continuo della fase, quantizzano tutte le ampiezze ad un singolo livello (se poi si usa un display a livelli discreti di grigio, allora anche la fase risulta quantizzata). In entrambi i casi il coefficiente complesso F è solo approssimato dal suo valore quantizzato

$$Q = F + N \quad (22)$$

dove N è il rumore di quantizzazione nel piano dell'ologramma, cioè nel dominio di F . L'immagine risultante è:

$$f' = \mathcal{F}^{-1}[Q] = \mathcal{F}^{-1}[F] + \mathcal{F}^{-1}[N] = f + n \quad (23)$$

dove n è l'errore di quantizzazione nel dominio dell'immagine.

Gli approcci visti nei due sottoparagrafi precedenti possono essere usati per la codifica di CGH quantizzati in modo da ridurre al minimo il valore di n . In questo caso i vincoli nel dominio di Fourier impongono che i valori della trasformata siano quantizzati su un numero finito di livelli prestabiliti, mentre nel dominio dell'oggetto si richiede che l'ampiezza dell'immagine ricostruita uguagli l'ampiezza ideale $|f(x)|$. Poiché nel dominio oggetto interessa solo l'ampiezza, all'immagine si può aggiungere qualunque funzione di fase in modo da soddisfare al meglio i vincoli nel dominio di Fourier.

Una scelta plausibile per il cambiamento dell'output è il più piccolo cambiamento $\Delta g_d(x)$ tale che

$$g'(x) + \Delta g_d(x) = |f(x)| \quad (24)$$

cioè

$$\Delta g_d(\mathbf{x}) = |f(\mathbf{x})| \frac{g'(\mathbf{x})}{|g'(\mathbf{x})|} - g'(\mathbf{x}) \quad (25)$$

Si è notato [7] che la differenza di fase tra $g'(\mathbf{x})$ e $g(\mathbf{x})$ tende ad avere lo stesso segno del cambio di fase di $g'(\mathbf{x})$ in iterate successive, quindi è desiderabile scegliere $\Delta g_d(\mathbf{x})$ che tende a ruotare l'angolo di fase del nuovo input verso quello dell'ultimo output. Per questo motivo un'altra buona scelta per il cambio dell'output è la seguente:

$$\Delta g_d(\mathbf{x}) = \left[|f(\mathbf{x})| \frac{g'(\mathbf{x})}{|g'(\mathbf{x})|} - g'(\mathbf{x}) \right] + \left[|f(\mathbf{x})| \frac{g'(\mathbf{x})}{|g'(\mathbf{x})|} - |f(\mathbf{x})| \frac{g(\mathbf{x})}{|g(\mathbf{x})|} \right] \quad (26)$$

in cui il primo componente modifica l'ampiezza dell'output al livello desiderato e il secondo ruota la sua fase verso l'angolo di output. Il nuovo input è dato inserendo l'eq. (26) nell'eq. (20) o (21). In [7] sono riportati i risultati di alcune simulazioni riguardanti ologrammi sia detour-phase che non detour-phase. Le tecniche error-reduction e input-output sono utilizzate per ridurre l'errore di quantizzazione in ologrammi sintetizzati con una tecnica tradizionale con codifica di fase di tipo casuale. Entrambi gli approcci consentono di raggiungere significativi miglioramenti nelle immagini ricostruite con un ridotto numero di iterazioni.

4. Conclusioni

La possibilità offerta dall'olografia di ricostruire immagini tridimensionali può essere combinata con la flessibilità del computer. Abbiamo visto come la sintesi di un ologramma digitale può essere realizzata per mezzo di un computer, adottando due tipi fondamentali di codifica (detour phase e non detour phase). Basandoci sul fatto che la funzione di trasmittenza dell'ologramma e il fronte d'onda diffratto sono una coppia di Fourier, abbiamo presentato due schemi iterativi, derivanti dall'algoritmo di Gerchberg e Saxton, finalizzati a trovare una coppia di Fourier che

soddisfi i vincoli in entrambi i domini. Nel nostro caso i vincoli imposti nel dominio di Fourier sono relativi ai livelli di quantizzazione adottati e quelli nel dominio oggetto impongono un andamento prestabilito all'ampiezza della funzione immagine. Essendo vincolata solo l'ampiezza, il metodo può scegliere liberamente la fase in modo da ridurre il rumore di quantizzazione. Lo stesso obiettivo è perseguito dalle tecniche di sintesi tradizionali per mezzo della manipolazione di fase, deterministica o casuale. Il confronto tra le immagini ricostruite mostra però un netto vantaggio di questi metodi iterativi rispetto a quelli tradizionali.

5. Bibliografia

- [1] E. Barnard, P. Vermeulen, D.P. Casasent: "Optical Correlation CGH with Modulated Error Diffusion", *Applied Optics*, Vol. 28, pag. 5358, (1989).
- [2] M. Bertero: "Linear inverse and ill-posed problems", INFN/TC 88/2, (1988).
- [3] B.R. Brown and A.W. Lohmann: "Complex Spatial Filtering with Binary Masks", *Applied Optics*, Vol. 5, pag. 1739, (1967).
- [4] M. Cairoli, M. Pagani, P. Vannotti: "Generazione di Fronti d'Onda Mediante Olografia Sintetica", Atti del 1^o Convegno Nazionale *Strumentazione e metodi di misura elettroottici*, pag.163, Milano (1990).
- [5] N. Collings: *Optical Pattern Recognition Using Holographic Techniques*, Addison-Wesley, (1988).
- [6] T. Eismann, A.M. Tai, J.N. Cederquist: "Iterative Design of a Holographic Beamformer", *Applied Optics*, Vol.28, pag. 2641, (1989).
- [7] J. R. Fienup: "Iterative Method Applied to Image Reconstruction and to Computer-Generated Holograms", *Opt. Eng.*, Vol. 19, pag. 297, (1980).
- [8] D. Gabor: "A New Microscopic Principle", *Nature London*, Vol. 161, pag. 777, (1948).
- [9] N.C. Gallagher, jr., J.A. Bucklew: "Non-Detour Phase Digital Holograms", *Applied Optics*, Vol. 19, pag. 4266, (1980).
- [10] W. Gerchberg: W.O. Saxton: "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures", *Optik*, Vol. 35, pag. 237, (1972).
- [11] W. Gerchberg: "Super-Resolution through Error Energy Reduction", *Optica Acta*, Vol. 21, pag. 709, (1974).

- [12] J.W.Goodman: *Introduction to Fourier Optics*, McGraw-Hill, (1968).
- [13] J.W.Goodman: "An Introduction to the Principles and Application of Holography", *IEEE Proc.*, Vol. 59, No. 9, pag. 1292 (1971).
- [14] B.K. Jennison, J.P. Allebach, D. Sweeney: "Iterative Approaches to Computer Generated Holography", *Opt. Eng.*, Vol 28, pag. 629, (1989).
- [15] D. Leseberg: "Computer Generated Hologram: Cylindrical, Conical, and Helical Waves", *Applied Optics*, Vol. 26, pag. 4385, (1987).
- [16] D.Leseberg, C.Frère: "Computer-Generated Holograms of 3-D Objects Composed of Tilted Planar Segments", *Applied Optics*, Vol. 27, pag. 3020 (1988)
- [17] L.B. Lesem, P.M. Hirsch, J. Jordan, jr.: "Kinoforms", *IBM J. Res. Dev.*, Vol. 13, pag. 150, (1969).
- [18] A.W. Lohmann, D.P. Paris: "Binary Fraunhofer Holograms, Generated by Computer", *Applied Optics*, Vol. 6, pag. 1739, (1967)
- [19] Ju.I. Ostrovskij: *Olografia e Sue Applicazioni*, Edizioni Mir, Mosca (1982).
- [20] J. C. Pateau et al.:"Incoherent Filtering Using Kinoforms", *IBM J. Res. Dev.*, Vol. 14, pag. 485, (1970).
- [21] L. Rizzi, P. Delvò, A. Menegazzo: "L'Olografia come Tecnica di Misura: Panoramica delle Applicazioni", *Atti del 1^o Convegno Nazionale Strumentazione e metodi di misura elettroottici*, pag.151, Milano (1990).
- [22] M.A. Seldowitz, J.P. Allebach, D.W. Sweeney: "Synthesis of Digital Holograms by Direct Binary Search", *Applied Optics*, Vol. 26, pag. 2788, (1987).
- [23] G. Tricoles: "Computer Generated Holograms: an Historical Review", *Applied Optics*, Vol 19, pag. 4266, (1987).