

Alcune tecniche di elaborazione numerica delle immagini

Some of the actual digital image processing techniques

V. CAPPELLINI (*), R. CARLA (**), C. CONESE (**), E. DEL RE (**)

Riassunto - Vengono brevemente passate in rassegna alcune delle principali tecniche di elaborazione numerica delle immagini: operatori locali e puntuali di tipo lineare e non lineare, tecniche di esaltazione dei contrasti ed estrazione dal rumore, estrazione dei contorni, trasformate di Fourier, filtri numerici, ecc. Nella esposizione vengono inoltre presentati alcuni risultati ottenuti mediante l'uso delle tecniche descritte. Tra i vari campi di applicazione, vengono infine descritti alcuni esempi di elaborazione di immagini di telerilevamento per indagini sul territorio.

Abstract - Some main techniques of digital image processing are reviewed: local and point operators of linear and non-linear type, contrast enhancement and noise extraction techniques, edge detection, Fourier transform, digital filtering, etc. Along with the exposition some of the results obtained by means of the described algorithms are also presented. Among the many areas of application, at last some examples of remote sensing image processing for land use are described.

1. Introduzione

Le immagini in bianco-nero e a colori vanno acquistando importanza crescente come sorgenti di informazione per l'analisi, l'interpretazione e la risoluzione dei più vari problemi, dall'analisi del territorio tramite satelliti orbitanti, allo studio di malattie o di strutture organiche in biomedicina, alle applicazioni in astronomia fino alle più recenti tecniche, di grande importanza, di riconoscimento automatico di forme ed oggetti che rientrano nel campo della robotica. Il veloce incremento della utilizzazione dei dati di immagini e la crescente complessità dei problemi da risolvere hanno richiesto lo sviluppo di sistemi numerici automatici, rapidi, accurati ed economici per estrarre ed elaborare l'informazione contenuta nelle immagini. Questo è stato possibile fondamentalmente grazie allo straordinario progresso della microelettronica che ha permesso lo sviluppo di elaboratori a basso costo e di notevole potenza, sempre più facilmente programmabili, disponibili quindi per una più larga fascia di utenti, ed ha portato alla produzione di sistemi perfezionati per trasformare un'immagine in dati numerici e viceversa.

Data l'enorme quantità di tecniche usate nell'elaborazione delle immagini a mezzo calcolatore, saranno qui trattate solo le principali, con qualche esempio di applicazioni pratiche attuate su un tipico sistema di elaborazione.

2. Sistema di elaborazione

Il sistema su cui sono state condotte le esperienze illustrate è stato realizzato (ed è pienamente efficiente) presso l'Istituto di Elettronica della Facoltà di Ingegneria dell'Università degli Studi di Firenze:

(*) Prof. Vito Cappellini, Facoltà di Ingegneria, Università di Firenze.

(**) Dr. ing. Roberto Carla, dr. ing. Claudio Conese, dr. ing. Enrico Del Re, Istituto di Elettronica, Facoltà di Ingegneria, Università di Firenze.

esso si basa su un minicalcolatore Digital PDP 11/34, la cui unità centrale di elaborazione (CPU) possiede 256 kbyte di memoria. Le memorie esterne sono costituite da due dischi rigidi, della capacità di 5 Mbyte ognuno, che vengono usati per l'immagazzinamento dei dati interessati dalle elaborazioni in corso.

La memoria periferica di maggiore capacità è costituita da una unità nastro magnetico a doppia densità.

I terminali di ingresso-uscita (I/O) dati sono due: un monitor con tastiera ed una stampante, sempre con tastiera, che possono quindi funzionare entrambi sia come ingresso sia come uscita; è inoltre possibile l'uso contemporaneo di ambedue.

L'uscita dati più interessante è il sottosistema di visualizzazione delle immagini, che comprende una interfaccia 256 x 256 punti con otto piani di memoria (quindi 256 livelli di grigio) e memorie programmabili per i tre colori fondamentali, con conseguente possibilità di scelta delle miscele di essi. Ad essa fanno capo sia un video (standard televisivo) a colori sia una telecamera per la ripresa di scene dal mondo esterno.

Un ulteriore apparato di ripresa è costituito da una camera AGA Thermovision per acquisizione di immagini nello spettro dell'infrarosso. Completa le possibilità di I/O dati un convertitore A/D e D/A per l'analisi di segnali di varia origine.

3. Immagine analogica e numerica

Un elaboratore elettronico può lavorare solamente su dati in forma numerica. Una qualunque immagine non si presenta in genere sotto questa forma. È opportuno chiarire la distinzione tra immagini cosiddette « analogiche » e quelle « digitali » o « numeriche », cioè quelle che sono direttamente utilizzabili dalla macchina.

Un'immagine statica può essere schematizzata come la memorizzazione di informazioni su un piano. Concettualmente, non c'è differenza tra la pagina di un

giornale dove, sul piano x, y delle righe e delle colonne, sono memorizzate lettere dell'alfabeto in un certo ordine in modo da costituire frasi intelligibili, ed una fotografia; in quest'ultima l'informazione è riconoscibile sotto forma di toni di grigio. A differenza delle lettere dell'alfabeto, le tonalità di grigio sono un insieme continuo, cioè un insieme di gradazioni che sfumano impercettibilmente l'una nell'altra tra due estremi (bianco e nero).

All'insieme continuo se ne può sostituire uno discreto (quantizzazione): il campo di variazione della grandezza continua (luminosità) viene suddiviso in un certo numero di fasce o intervalli e tutti i valori che cadono in uno stesso intervallo vengono considerati uguali. Si introduce così un'approssimazione, che può essere migliorata quanto si vuole aumentando il numero delle fasce e diminuendone l'ampiezza. Alla variazione continua si è sostituita una variazione a gradini (livelli di grigio), rappresentabili con numeri.

Più rigorosamente, una qualunque immagine analogica dinamica può essere definita matematicamente da una funzione reale di variabili reali:

$$F(x, y, t) \quad (1)$$

che rappresenta la distribuzione dell'energia spaziale (luminosità) in funzione delle coordinate x, y e del tempo t .

Poiché l'intensità luminosa è una quantità reale e positiva, la funzione risulta reale e non negativa.

I sistemi di registrazione e rappresentazione di immagini impongono alcune restrizioni sulla luminosità massima a causa dei loro limiti fisici intrinseci (fenomeni di saturazione), per cui si suppone che:

$$0 \leq F(x, y, t) \leq L \quad (2)$$

dove L rappresenta la luminosità massima.

Ovviamente, la rappresentazione di un'immagine fisica è anche limitata nelle sue dimensioni; quindi, per semplicità, si suppone che tutte le immagini possano assumere valore diverso da zero solo sulla superficie così definita:

$$\begin{aligned} X_1 &\leq x \leq X_2 \\ Y_1 &\leq y \leq Y_2 \end{aligned} \quad (3)$$

Un'immagine è inoltre praticamente osservabile solo in un intervallo finito di tempo ($T_1 \leq t \leq T_2$), ma in questa sede verranno considerate solo immagini statiche tempo-invarianti e quindi rappresentabili da una funzione indipendente dal tempo:

$$F(x, y) \quad (4)$$

Infine, la funzione F si suppone continua nel suo dominio di definizione.

Il concetto di « continuo » in questo caso è comunque legato al potere risolutivo dell'occhio umano, che vede come tale anche un'immagine discontinua i cui punti però siano tra loro distinti di una quantità inferiore al suo potere discriminante. Tipico esempio

è la fotografia, che è un'immagine costituita da punti (i granuli d'argento), ma talmente vicini da essere indistinguibili all'occhio.

Per come è stata definita, quindi, un'immagine comune comprende teoricamente ∞^2 punti sul piano e presuppone che anche la scala dei livelli di luminosità (o « livelli di grigio ») che ciascun punto del quadro può assumere sia continua e quindi con infiniti valori.

Poiché un elaboratore elettronico è in grado di elaborare un numero di dati che, anche se molto grande, è pur sempre limitato, un trattamento con questo presuppone la conversione in forma numerica dell'immagine analogica.

Questa operazione può essere suddivisa, come è noto, in tre fasi: campionamento, quantizzazione e conversione in forma binaria.

Si può pensare di sovrapporre all'immagine una griglia le cui maglie quadrate sono uguali alla cella di risoluzione dell'immagine numerica; il valore medio di luminosità di ognuna di esse viene quantizzato come già osservato e costituisce il singolo elemento o « pixel » (picture element) della matrice rappresentante l'immagine digitalizzata.

Un'immagine numerica o digitale è quindi completamente rappresentata da una matrice di interi:

$$\begin{aligned} M(i, j) &, \quad 0 \leq M \leq M_m \\ I_1 &\leq i \leq I_2 \\ J_1 &\leq j \leq J_2 \end{aligned} \quad (5)$$

in cui M_m rappresenta il valore massimo di luminosità.

Minore è l'intervallo di quantizzazione, maggiore sarà il numero dei livelli in cui viene suddivisa la scala dei grigi e quindi dei toni rappresentabili, avvicinandosi così, come effetto visivo, all'originale analogico.

In genere una scala a 64 livelli di grigio offre una restituzione sufficiente per molti scopi pratici.

4. Tecniche di elaborazione numerica delle immagini

L'elaborazione numerica delle immagini è basata su alcuni fondamentali principi del processo visivo umano.

Sebbene questo sia fisiologicamente complesso, il meccanismo base della vista utilizza gli occhi come sensori ed il cervello come sistema di interpretazione automatica delle informazioni; gli occhi ricevono stimoli sotto forma di luce ed il cervello elabora ed interpreta questi ingressi.

Analogamente, un sistema di elaborazione delle immagini usa in genere una unità di acquisizione (ad esempio con scansione elettronica mediante una telecamera per riprese video) per trasferire i dati dell'immagine ad un elaboratore elettronico.

Questo sistema può in definitiva « vedere » le immagini tramite l'unità di acquisizione e, per mezzo della capacità elaborativa dei vari programmi realizzati, può intervenire sulle immagini in modo tale da consentire l'estrazione delle informazioni desiderate.

I tipi di trattamenti possibili vengono generalmente

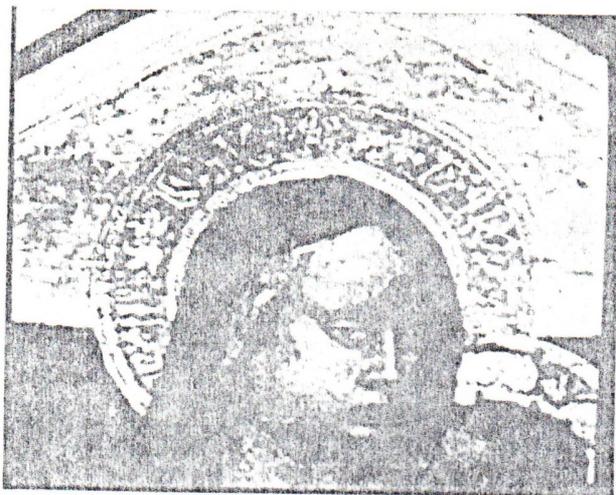


Fig. 1 - Esempio di acquisizione numerica di una parte di un quadro del Masaccio.



Fig. 2 - Ingrandimento 2x2 di una porzione dell'immagine di fig. 1.

suddivisi in pre-elaborazioni ed elaborazioni vere e proprie.

Il campo delle prime comprende tutti i procedimenti volti a trasformare l'informazione contenuta in una immagine in una forma più facilmente estraibile dalle successive elaborazioni ed interpretabile da parte dell'utente.

Le tecniche di pre-elaborazione vengono talora distinte in « enhancement » e « restoration », anche se in realtà molte non sono attribuibili univocamente all'una oppure all'altra classe. Gli algoritmi del primo gruppo permettono di mettere in evidenza o esaltare certi particolari presenti nell'immagine; al secondo gruppo appartengono invece quei procedimenti atti a migliorare o « restaurare » la qualità dell'immagine deteriorata dal rumore introdotto dai circuiti di digitalizzazione e dagli apparecchi di ripresa (ad es.: sfuocamento).

Gli algoritmi di elaborazione più comuni sono basati su criteri puntuali o locali. Un procedimento puntuale elabora ciascun punto dell'immagine indipendentemente da tutti gli altri, mentre, in un procedimento basato su un criterio del tipo locale, ciascun punto assume un valore in funzione anche dei valori dei punti vicini.

Una delle più semplici e nello stesso tempo importanti elaborazioni del tipo puntuale è costituita dall'ingrandimento di una parte dell'immagine digitale, che permette all'osservatore di esaminare dettagli di interesse particolare. Ciò è possibile sia con un opportuno sistema ottico di lenti, sia per mezzo dell'elaboratore con un'operazione « software », come di seguito indicato.

Il metodo più semplice consiste nel ripetere lungo le due dimensioni dell'immagine il valore di ogni singolo pixel un numero di volte pari al fattore di ingrandimento. Se supponiamo ad esempio che la zona da ingrandire sia rappresentata dalla seguente matrice 3 x 3:

$$\begin{bmatrix} 0 & 10 & 7 \\ 5 & 6 & 8 \\ 6 & 8 & 9 \end{bmatrix}$$

effettuando un ingrandimento 2 x 2 si ottiene:

$$\begin{bmatrix} 0 & 0 & 10 & 10 & 7 & 7 \\ 0 & 0 & 10 & 10 & 7 & 7 \\ 5 & 5 & 6 & 6 & 8 & 8 \\ 5 & 5 & 6 & 6 & 8 & 8 \\ 6 & 6 & 8 & 8 & 9 & 9 \\ 6 & 6 & 8 & 8 & 9 & 9 \end{bmatrix}$$

in cui ciascun pixel è sostituito da 4 pixel (2 x 2) dello stesso valore.

La fig. 1 rappresenta, come esempio, un'acquisizione effettuata mediante telecamera su un quadro del Masaccio; in fig. 2 è invece riportato il risultato ottenuto dall'elaborazione di ingrandimento 2 x 2 effettuata su una porzione dell'immagine originale. È evidente come al crescere del fattore di ingrandimento risulta sempre più visibile il carattere discreto della digitalizzazione.

È possibile inoltre variare la dimensione dell'immagine di un fattore non intero mediante opportune operazioni di media, come pure eseguire l'operazione inversa mediante un campionamento nelle due dimensioni.

Una delle caratteristiche più importanti di un'immagine è il contrasto, che è determinato dall'ampiezza delle variazioni di luminosità tra punti adiacenti. L'occhio umano, essendo sensibile al contrasto, distingue più facilmente i particolari di una figura fortemente contrastata che di una poco contrastata. L'operazione di esaltazione del contrasto può quindi essere impiegata per rendere più facilmente interpretabili le immagini e per evidenziare dettagli poco chiari. Tra i molti algoritmi sviluppati per questo scopo, uno dei più semplici consiste nel sommare o sottrarre un valore di luminosità costante a ciascun pixel dell'immagine in considerazione. Se l'immagine originale è una fotografia, tale operazione è equivalente ad una variazione del tempo di esposizione oppure ad una variazione della quantità di illuminazione (fig. 3).

Un altro algoritmo per l'esaltazione del contrasto,

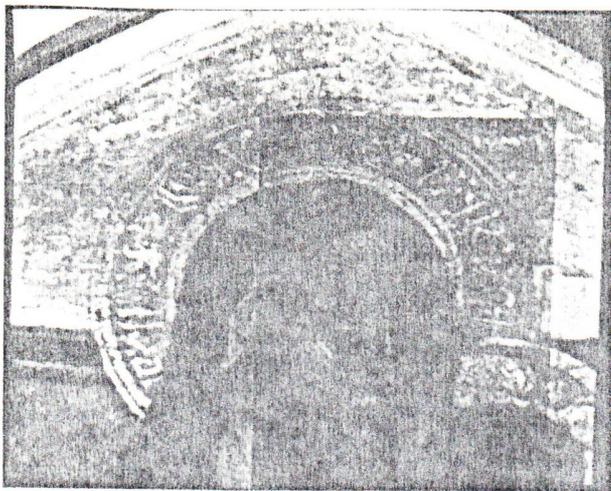


Fig. 3 - Variazione di un fattore costante dei livelli di grigio di una sottoparte dell'immagine di fig. 1.

sempre del tipo puntuale, è detto « stretching » (stiramento) lineare. Supponiamo che i valori di luminosità non occupino uniformemente tutta la gamma dei valori possibili (0-Max), ma la maggior parte sia compresa tra una soglia inferiore (S_1) ed una soglia superiore (S_2).

Sottraendo dal livello di grigio S di ciascun pixel il valore della soglia inferiore e moltiplicando per un opportuno fattore di espansione, tutti i valori tra zero ed S_1 assumono valore zero, quelli tra S_2 e Max assumono valore Max ed i rimanenti vengono distribuiti uniformemente tra 0 e Max secondo l'espressione:

$$S' = (S - S_1) \frac{Max}{(S_2 - S_1)} \quad (6)$$

Si ottiene quindi un aumento lineare della variazione di luminosità da punto a punto e quindi del contrasto. Per potere stabilire le soglie più opportune, di grande utilità risulta la conoscenza dell'istogramma dei livelli di grigio dell'immagine, che rappresenta la densità di probabilità di ciascun livello di grigio. La fig. 4 rappresenta l'istogramma dell'immagine originale (fig. 1). L'asse delle ascisse rappresenta da sinistra a destra la scala dei livelli di grigio (0, 255), mentre sulle ordinate è riportato quante volte ciascun livello appare nell'immagine. La fig. 5 rappresenta il risultato ottenuto dall'applicazione dell'algoritmo di stretching lineare.

Un'altra semplice tecnica puntuale, che risulta in alcuni casi di notevole utilità ed è comunemente impiegata nell'elaborazione delle immagini, è la « quantizzazione », detta anche « slicing », che letteralmente significa « affettamento ». Il campo dei valori assunti dalla grandezza « livelli di grigio » viene diviso in un certo numero di fasce contigue ed uguali tra loro, o isointervalli; tutti i valori che cadono in una stessa fascia si considerano uguali tra loro ed assumono uno stesso valore (isolivello). Nel caso più generale ed anche più comune queste fasce contigue sono regolabili in ampiezza e posizione. In fig. 6 è riportata un'immagine della Toscana ripresa dal satellite Land-

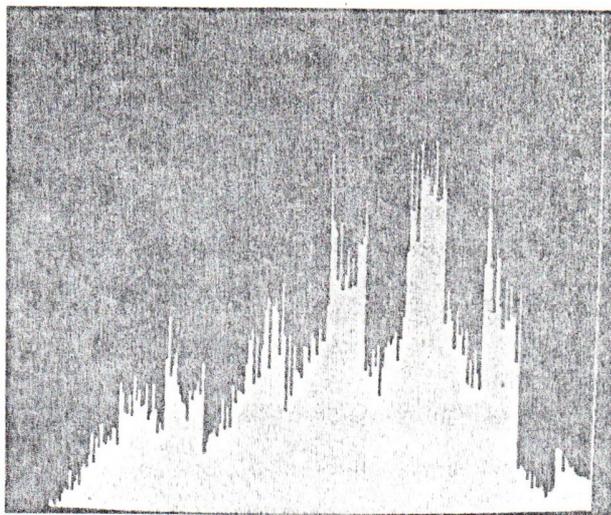


Fig. 4 - Iistogramma (distribuzione di ampiezza) dei livelli di grigio dell'immagine di fig. 1: ascisse - scala dei livelli di grigio (0 a sinistra, 255 a destra); ordinate - numero di volte che ciascun livello appare nell'immagine.



Fig. 5 - Elaborazione mediante l'algoritmo di « stretching » lineare.

sat; in particolare i due fiumi visibili sono l'Arno ed il Serchio. In quest'immagine la scala dei grigi ha 256 livelli. In fig. 7 è riportata invece la stessa immagine, ma quantizzata su solo 6 livelli di grigio. Si può facilmente notare come in quest'ultima vengano evidenziate zone particolari. In fig. 8 infine sono presenti solo 2 livelli (bianco, nero). Si ottiene così una prima separazione tra superficie terrestre e acque. In quest'ultima infatti sono particolarmente evidenti le coste, i fiumi.

È da notare che l'immagine è complementata. L'operazione di complementazione è un'altra preelaborazione del tipo puntuale, che consiste nell'invertire la scala dei grigi, assegnando il valore minimo al livello di luminosità più alto e viceversa. Si ottiene così il negativo dell'immagine su cui possono evidenziarsi particolari non chiari dell'originale.

Un'informazione di notevole importanza è data dalla conoscenza dei contorni degli oggetti e delle zone raffigurate. Numerosi sono gli algoritmi per la rivela-

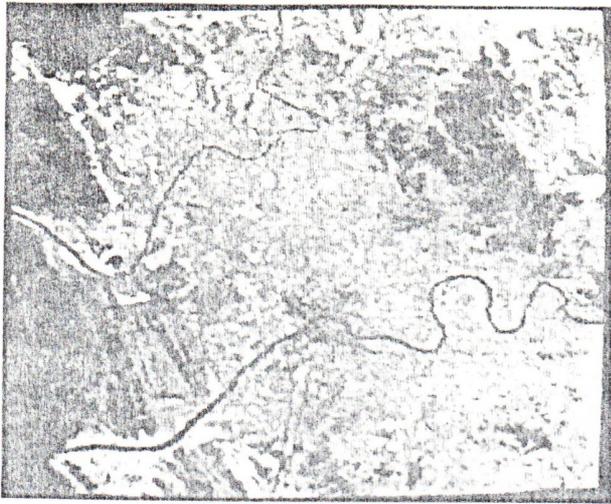


Fig. 6 - Immagine della Toscana ripresa dal satellite Landsat (256 livelli di grigio): in basso è visibile il fiume Arno, più in alto il fiume Serchio.



Fig. 7 - Presentazione con solo 6 livelli di grigio di una parte dell'immagine di fig. 6.

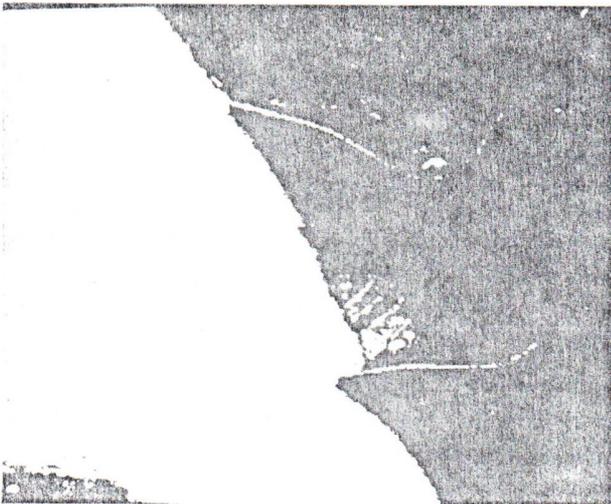


Fig. 8 - Presentazione con solo 2 livelli di grigio dell'immagine di fig. 6.

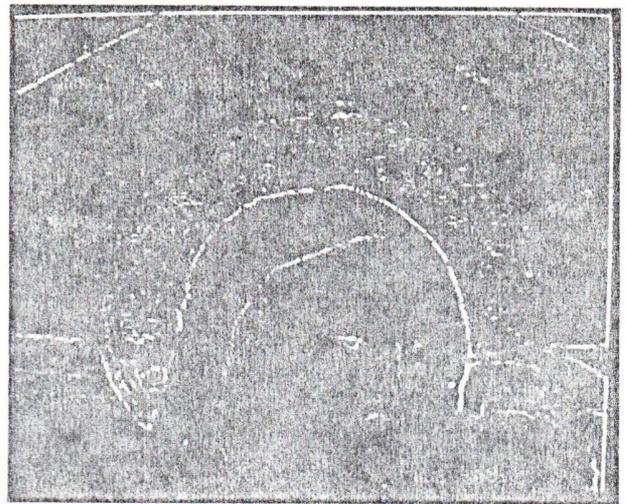


Fig. 9 - Estrazione dei contorni con algoritmo di Roberts sulla immagine di fig. 1.



Fig. 10 - Estrazione dei contorni con algoritmo di Sobel sull'immagine di fig. 1.

zione od estrazione dei contorni già sviluppati. Alcuni di questi fanno uso di particolari filtri numerici, mentre la maggior parte impiega operatori locali 2×2 o 3×3 per cercare di visualizzare in ogni punto il gradiente massimo e cioè le linee di maggiore disomogeneità tra zone sufficientemente uniformi. I risultati ottenuti dall'applicazione di ciascun algoritmo risultano più o meno buoni in funzione del tipo e delle caratteristiche dell'immagine elaborata; in genere, la scelta dell'algoritmo più adatto andrà eseguita caso per caso. Data la vastità dell'argomento, per una trattazione più approfondita si rimanda alla bibliografia allegata. A titolo di esempio, in fig. 9 e fig. 10 sono riportati i risultati ottenuti dall'applicazione di due tra i più noti algoritmi per l'estrazione dei contorni, dovuti rispettivamente a Roberts ed a Sobel.

Un altro esempio di elaborazione del tipo locale che può risultare di notevole utilità, effettua un'opera-