

Principi ed applicazioni dell'elaborazione digitale delle immagini

di Giuseppe Cardinale Ciccotti

Elaborazione d'immagini può significare molte cose diverse, dipendendo dal tipo di risultato che si vuole ottenere, in ogni caso possiamo generalizzare le operazioni su di un'immagine come una funzione che mappa l'immagine di partenza in una di arrivo.

Questa funzione può essere di vario tipo e nel prosieguo di questi appuntamenti cercheremo di classificare in maniera chiara le principali operazioni che si possono compiere su di un'immagine. Nell'articolo precedente abbiamo analizzato alcuni dei principali concetti su cui poggiare il resto della nostra trattazione, ne servono pochi altri e già da questo numero potremo operare direttamente sulle immagini

La distanza e la sua misura

Dati i pixel p, q e z di coordinate (x, y), (s, t), e (u, v) rispettivamente, la distanza D è definita come una funzione che verifica le seguenti proprietà:

- a) $D(p, q) \geq 0$ con $D(p, q) = 0$ se $p = q$
- b) $D(p, q) = D(q, p)$
- c) $D(p, z) \leq D(p, q) + D(q, z)$

La relazione a) significa che la distanza è un numero positivo eventualmente pari a 0 se i due punti di misura coincidono.

La relazione b) indica che il verso della misura non è significativo, cioè la misura è un valore assoluto.

La relazione c) ci ricorda che la distanza diretta tra due punti è minore della somma delle due distanze tra il punto di partenza e un punto intermedio e da questo a quello di arrivo, essendo uguale soltanto se, ovviamente, il punto intermedio è sul cammino diretto tra p e z.

Esistono vari modi di calcolare la distanza tra due punti; il più comune ed intuitivo è quello indicato come «distanza Euclidea» ed è quello che usiamo comunemente:

$$De(p, q) = ((x - s)^2 + (y - t)^2)^{0.5}$$

Con questo tipo di distanza i pixel che hanno distanza minore o uguale ad r da (x, y) sono tutti i punti contenuti in un cerchio di raggio r centrato in (x, y).

Un altro tipo di distanza detto «distanza a blocchi» è calcolato come:

$$D4(p, q) = |x - s| + |y - t|$$

In questo caso i pixel che hanno una distanza D4 da p minore o uguale a r formano un diamante centrato in (x, y). Per esempio i pixel con distanza $D4 \leq 2$ da (x, y) hanno questo tipo di configurazione:

$$\begin{array}{cccc} & & 2 & \\ & & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ & & 2 & 1 & 2 \\ & & & 2 & \end{array}$$

Notate che i pixel a distanza 1 sono i 4 adiacenti, da qui il simbolo D4.

Un ulteriore tipo di distanza può essere calcolato come il massimo della differenza delle coordinate omologhe dei due punti:

$$D8(p, q) = \max(|x - s|, |y - t|)$$

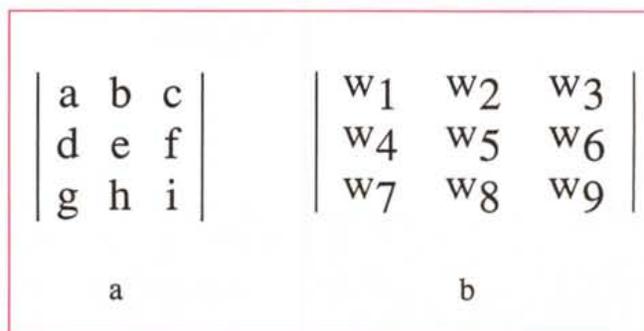


Figura 1 - Schema di maschera o filtro: in a sono rappresentati i pixel, in b la maschera.

I pixel con distanza D8 da p minore o uguale a r sono disposti in questo caso come un quadrato centrato a (x, y). Per esempio, i pixel con distanza D8 <= 2 da (x, y) formano questa figura:

```

2 2 2 2 2
2 1 1 1 2
2 1 0 1 2
2 1 1 1 2
2 2 2 2 2
    
```

Anche in questo caso, notate come i pixel a distanza costante 1 sono gli 8 adiacenti. Un'ultima considerazione sulla distanza è che non si fa, nel calcolo, nessuna ipotesi sul valore dei pixel attraversati, ma soltanto sul valore delle coordinate dei pixel di partenza e di arrivo; tuttavia se si cerca un cammino tra due punti imponendo un criterio di connettività, per esempio soltanto per i pixel di un certo valore, è ovvio che la distanza sarà soltanto il valore minimo che si può ottenere per la distanza: è in definitiva la distanza «in linea d'aria» fra due punti, mentre la distanza vincolata dalla connettività è la distanza «stradale».

Operazioni logiche e aritmetiche

Le operazioni logiche ed aritmetiche sono intensivamente adoperate nell'elaborazione di immagini. Le operazioni aritmetiche fra due pixel p e q sono denotate come segue:

Addizione: p+q
 Sottrazione: p-q
 Moltiplicazione: p*q
 Divisione: p/q

Ovviamente p o q possono essere un valore costante.

Le principali operazioni logiche usate nell'elaborazione d'immagini sono l'AND, OR, e il COMPLEMENTO indicate con:

AND: p AND q
 OR: p OR q
 COMPLEMENTO: NOT p

È noto come questi tre operatori costituiscono un insieme funzionalmente completo, cioè sono sufficienti a completare qualsiasi tipo di operazione logica tramite la loro successiva composi-

$$\begin{bmatrix} 0 & 0 & 1 & X_0 \\ 0 & 0 & 1 & Y_0 \\ 0 & 0 & 1 & Z_0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix}$$

Figura 2 - Traslazione, la matrice è 3x4. Bisogna aggiungere una componente fittizia per eseguire il calcolo.

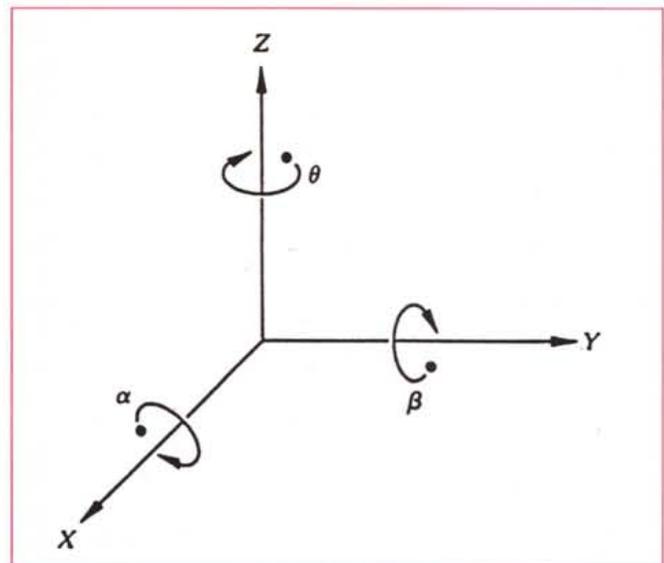
$$\begin{bmatrix} 0 & 0 & 1 & X_0 \\ 0 & 0 & 1 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ 1 \end{bmatrix}$$

Figura 3 - Matrice di traslazione normalizzata a 4x4.

$$\begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Figura 4 - Matrice per la trasformazione di scala.

Figura 5 - Convenzioni per le rotazioni rispetto ai tre assi. Gli angoli sono positivi in senso orario guardando l'asse.



$$R_{\theta} = \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Figura 6 - Matrice di rotazione per l'asse z. Notate come la componente z non è interessata.

$$R_{\alpha} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Figura 7 - Matrice di rotazione per l'asse x.

$$R_{\beta} = \begin{bmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Figura 8 - Matrice di rotazione per l'asse y.

zione. Chiaramente gli operatori matematici possono essere applicati direttamente al valore dei pixel nelle immagini multilivello mentre per gli operatori logici si dovrà operare bit a bit.

Le operazioni appena illustrate possono essere adoperate in due modi distinti: pixel per pixel oppure operando sui pixel adiacenti.

Per esempio la somma di due immagini è compiuta addizionando ciascun pixel con l'omologo dell'altra immagine. Alternativamente posso definire delle operazioni in cui il valore da assegnare ad un pixel è funzione del valore degli adiacenti. Questa operazione è spesso detta «mascheratura» o «filtraggio».

Illustriamo con un esempio questo importante concetto; supponiamo di voler rimpiazzare un pixel con la media dei suoi 8 pixel adiacenti. Riferendoci a figura 1a, vogliamo sostituire al valore di e il valore p così calcolato:

$$p = 1/9*(a+b+c+d+e+f+g+h+i)$$

Definendo un filtro come in figura 1b possiamo «sovrapporlo» ai 9 pixel di partenza e moltiplicare ciascun coefficiente per il valore del pixel omologo (cioè quello che si trova «sotto») e sommare tutti i valori insieme indi assegnare il risultato a e, in formule:

$$p = w1*a + w2*b + w3*c + w4*d + w5*e + w6*f + w7*g + w8*h + w9*i$$

Per far sì che le due formule precedenti corrispondano è necessario che i valori di tutti i wi con i=1..9 siano pari a 1/9.

Quest'operazione così semplice è fondamentale nell'elaborazione di immagini, ne faremo perciò largo uso. Scegliendo opportunamente i valori del filtro è possibile compiere una gran varietà di operazioni come la riduzione del rumore e l'evidenziazione dei contorni soltanto per citarne due fra le più note. Bisogna tuttavia puntualizzare come l'applicazione del filtro a ciascun pixel di un'immagine è un compito computazionalmente gravoso. Come potete rendervi conto filtrare un'immagine 512x512 con un filtro 3x3 richiede 9 moltiplicazioni e 8 addizioni per ciascun pixel, in totale 2359296 moltiplicazioni e 2097152 addizioni.

Traslazione

Supponiamo di voler spostare un punto di coordinate (X, Y, Z) in un nuova locazione che dista X0 in X, Y0 in Y e Z0 in Z.

Semplicemente basterà sommare tali valori:

$$\begin{aligned} X1 &= X + X0 \\ Y1 &= Y + Y0 \\ Z1 &= Z + Z0 \end{aligned}$$

Le coordinate (X1, Y1, Z1) rappresentano il nuovo punto.

Queste equazioni possono essere espresse in forma matriciale come in figura 2. Tuttavia come vedete la matrice è 3x4 ed è necessario aggiungere una quarta coordinata fittizia posta ad 1 per eseguire i calcoli, per uniformità perciò si preferisce espandere la matrice a 4x4 come in figura 3. La quaterna (X, Y, Z, 1) individua un punto in coordinate omogenee. La trasformazione complessiva è individuata in maniera compatta con la notazione:

$$v1 = Tv$$

dove v1 e v sono i vettori colonna delle coordinate e T è la matrice di traslazione.

Scalamento

Allo stesso modo una trasformazione di scala può essere ottenuta moltiplicando le coordinate dei punti per la matrice in figura 4, i fattori Sx, Sy, Sz sono dei numeri reali che rappresentano il fattore di scala di ciascuna coordinata; per esempio volendo ridurre l'immagine della metà basterà porre Sx=0.5 e Sy=0.5 con Sz=1.

Rotazione

Per eseguire una rotazione dell'immagine è naturalmente necessario specificare rispetto a quale asse coordinato si vuole farla; otterremo così tre diverse matrici per ciascun asse, in figura 5 potete vedere le convenzioni adottate per definire il verso positivo di rotazione. In figura 6 è riportata la matrice per la rotazione intorno all'asse Z, in figura 7 quella per l'asse X e in figura 8 la

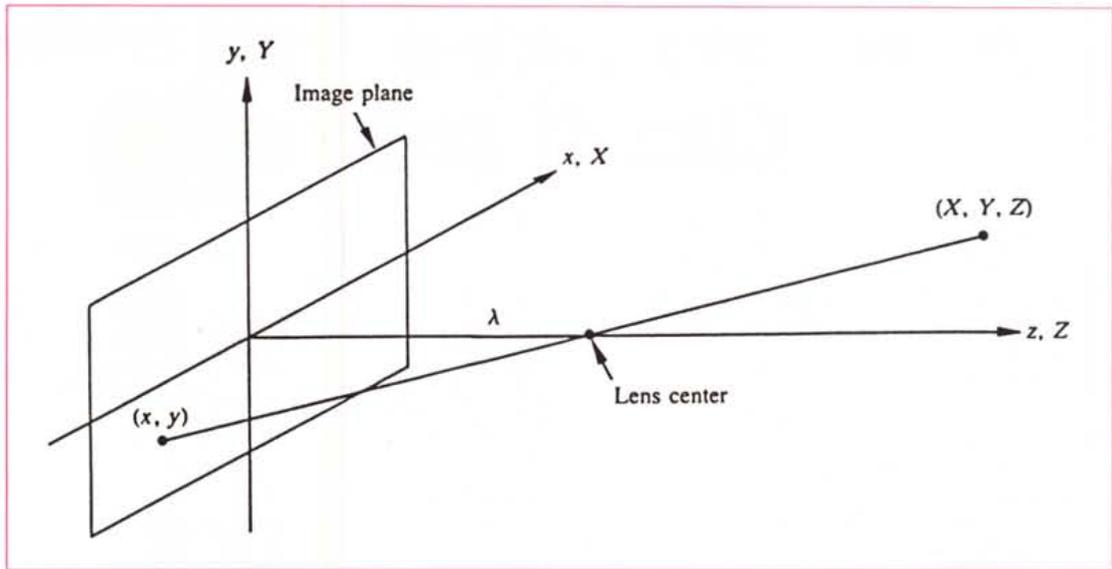
Trasformazioni geometriche

I lettori che hanno confidenza con il CAD sapranno già di cosa stiamo per parlare, tuttavia per i neofiti è necessario introdurre queste semplici trasformazioni. Occorre però una precisazione: tutte le seguenti considerazioni saranno riferite a punti dello spazio tridimensionale cioè ciascun punto sarà individuato dalla tripla (X, Y, Z) mentre i punti dell'immagine dello spazio bidimensionale verranno indicati con (x, y). Dovremo perciò considerare l'immagine bidimensionale in realtà come un piano dove X=x, Y=y e Z=0 vale a dire la nostra immagine è poggiata sul piano Z=0, in questo modo potremo eseguire tutte le operazioni nello spazio tridimensionale. L'origine delle coordinate sarà posto nell'origine dello schermo con l'asse Z che esce dallo schermo verso l'osservatore.

Bibliografia

R. C. Gonzalez, P. Wintz «Digital Image Processing». Addison-Wesley, 1987.

Figura 9 - Modello base di proiezione. Il sistema di riferimento dell'osservatore è allineato con quello dell'immagine.



corrispondente per Y.

Tenete conto che queste rotazioni sono con centro nell'origine, per eseguire una rotazione rispetto ad un punto arbitrario dobbiamo traslare preventivamente il punto nell'origine, eseguire la rotazione e ritraslare il punto nella posizione originale, in pratica bisogna traslare l'immagine con i parametri ricavati dalla traslazione del punto nell'origine. Per traslare indietro basterà cambiare di segno i parametri della matrice di traslazione.

Trasformazioni bidimensionali

Nel caso in cui volessimo vincolarci a rimanere sul piano Z=0, non è necessario considerare la coordinata Z e le matrici diventano 3x3 semplicemente eliminando la 3° riga e la 3° colonna. Naturalmente sarà possibile eseguire rotazioni soltanto intorno all'asse Z. Le trasformazioni si possono utilmente concatenare moltiplicando a destra le matrici che formano la trasformazione complessa, tuttavia è necessario ricordare che l'ordine è importante e cambiando l'ordine delle trasformazioni il risultato non è il medesimo.

Trasformazione prospettica

Una trasformazione prospettica proietta un punto tridimensionale su di un piano. È questa la trasformazione che, dopo tutte le precedenti trasformazioni in 3d precedenti, ci consente di visualizzare l'immagine sul nostro schermo.

Ci sono molti modelli di trasformazione prospettica, per i nostri scopi non è necessario approfondire eccessivamen-

te quest'argomento più proprio della Computer Graphics e del CAD, illustriamo perciò soltanto un metodo semplice per mappare le nostre immagini trasformate sullo schermo.

In figura 9 trovate un modello per la proiezione dell'immagine, definiamo che il sistema di coordinate dell'osservatore, supponiamo dotato di una macchina fotografica, sia coincidente con quello dell'immagine e in particolare il piano dell'immagine coincida con il piano Z=0 e l'asse Z esca fuori da tale piano coincidente con l'asse focale. Quindi il centro dell'immagine è nell'origine e il centro delle lenti dell'obiettivo è posto nel punto (0,0, l). l è la lunghezza focale dell'obiettivo.

Prendiamo ora un generico punto (X, Y, Z) dello spazio come in figura 9 e supponiamo che Z > l cioè questo punto sia davanti all'obiettivo. Ci interessa trovare la relazione che ci restituisce le coordinate (x, y) della proiezione del punto (X, Y, Z) sul piano dell'immagine (il nostro schermo).

Dalla proprietà dei triangoli simili si ricava:

$$x/l = - (X/(Z-l)) = X/(l-Z)$$

$$y/l = - (Y/(Z-l)) = Y/(l-Z)$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/l & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1-Z/l \end{bmatrix}$$

Figura 10.

quindi

$$x = l*X/(l-Z)$$

$$y = l*Y/(l-Z)$$

Bisogna notare come queste equazioni coinvolgendo la variabile Z non sono lineari e pertanto non sono invertibili, il che significa che da un'immagine bidimensionale non si possono ricostruire le scene tridimensionali rappresentate, come è ovvio.

Per convenienza, visto che questa trasformazione è in cascata con le precedenti, è utile esprimere la proiezione tramite una matrice in modo da poterla inserire nella catena di trasformazione complessiva. In figura 10 trovate la matrice di trasformazione, tenete conto che l'ultimo termine, il quarto del vettore che individua un punto, deve essere pari a 1 perciò bisogna dividere le prime due coordinate per questo numero.

Conclusioni

Questo appuntamento è stato denso di concetti tutti essenziali per la comprensione di determinate operazioni sulle immagini, in particolare il concetto di filtro è d'importanza capitale nell'elaborazione d'immagine. Nei prossimi articoli ce ne serviremo in maniera estensiva per sperimentare vari tipi di operazioni sulle immagini. Rimangono gradite le domande e le proposte dei lettori interessati in modo da poter spiegare il funzionamento di quelle tecniche che ognuno di noi usa utilizzando un qualsiasi pacchetto di grafica.