Disparity Inverse Perspective Image

Paolo Medici
Dipartimento di Ingegneria dell'Informazione di Parma

19 maggio 2008

1 Il problema

Dati due punti in 3 dimensioni definiti come $\tilde{A} = (X, Y, 0)^T$ e $\tilde{B} = (X, Y, Z)^T$ questi verranno proiettati su una immagine bidimensionale attraverso una funzione prospettica nei rispettivi punti a e b. Uno dei punti è alla base e uno a una altezza variabile dal terreno. Tali punti verranno trasferiti in una vista dall'alto, attraverso una funzione di inverse perspective mapping, ignorando la reale altezza dei punti. Tale funzione calcolerà la loro posizione in cui devono finire nell'immagine IPM considerando $z = 0$ costante per entrambi, pertanto riproiettandoli in due punti $\tilde{A}' \cong \tilde{A}$ e $\tilde{B}' \neq \tilde{B}$. Il punto effettivamente sul terreno viene proiettato nel posto corretto, mentre il punto verticale finirà in un altro punto diverso.

È possibile ricavare, dati $\tilde{A}'$ e $\tilde{B}'$, il valore Z?

2 Equazioni base

Da semplici considerazioni geometriche risulta molto evidente che ciò è possibile (proporzione tra triangoli simili). Tuttavia è possibile giungere alla medesima considerazione attraverso complicati conti matematici, che comunque validano l'utilizzo dell'IPM.

L'equazione base per una trasformazione prospettica PM è:

$$F_{pm}(x, y, z) = \left(k_u \frac{r_1 x' + r_2 y' + r_3 z'}{r_7 x' + r_8 y' + r_9 z'}, k_v \frac{r_4 x' + r_5 y' + r_6 z'}{r_7 x' + r_8 y' + r_9 z'} \right)$$

dove $(x', y', z')^T$ sono i punti in coordinate camera, ovvero $(x', y', z')^T = (x, y, z)^T - (x_0, y_0, z_0)^T$, e senza perdita di generalità dichiariamo $(x_0, y_0, z_0)^T = (0, 0, h)^T$, dove h è l'altezza della camera dal suolo.

I parametri k_u e k_v li tralasciamo, perchè si dovrebbero annullare in seguito.

Il punto $B = (X, Y, Z)$ pertanto viene proiettato nel punto (u, v):

$$u = \frac{r_1 X + r_2 Y + r_3 (Z - h)}{r_7 X + r_8 Y + r_9 (Z - h)} \quad (1)$$

$$v = \frac{r_4 X + r_5 Y + r_6 (Z - h)}{r_7 X + r_8 Y + r_9 (Z - h)} \quad (2)$$

L'equazione base per una trasformazione prospettica IPM è invece:

$$X' = \frac{r_1 u + r_4 v + r_7}{r_3 u + r_6 v + r_9}(Z - z_0) + x_0$$

$$Y' = \frac{r_2 u + r_5 v + r_8}{r_3 u + r_6 v + r_9}(Z - z_0) + y_0$$

dove Z è la presunta altezza dal suolo ($Z = 0$) del punto da proiettare, (u, v) è il punto sull'immagine prospettica e (X', Y', Z) il punto nel mondo.

L'equazione di $\tilde{B}' = (X', Y', 0)$ diventa pertanto:

$$X' = -\frac{r_1 u + r_4 v + r_7}{r_3 u + r_6 v + r_9} h \tag{3}$$
$$Y' = -\frac{r_2 u + r_5 v + r_8}{r_3 u + r_6 v + r_9} h \tag{4}$$

3 Soluzione

Sviluppando il sistema

$$X' = -h \frac{r_1(r_1 X + r_2 Y + r_3(Z-h)) + r_4(r_4 X + r_5 Y + r_6(Z-h)) + r_7(r_7 X + r_8 Y + r_9(Z-h))}{r_3(r_1 X + r_2 Y + r_3(Z-h)) + r_6(r_4 X + r_5 Y + r_6(Z-h)) + r_9(r_7 X + r_8 Y + r_9(Z-h))} \tag{5}$$
$$Y' = -h \frac{r_2(r_1 X + r_2 Y + r_3(Z-h)) + r_5(r_4 X + r_5 Y + r_6(Z-h)) + r_8(r_7 X + r_8 Y + r_9(Z-h))}{r_3(r_1 X + r_2 Y + r_3(Z-h)) + r_6(r_4 X + r_5 Y + r_6(Z-h)) + r_9(r_7 X + r_8 Y + r_9(Z-h))} \tag{6}$$

si ottiene la trasformazione omografica:

$$X' = -h \frac{X(r_1 r_1 + r_4 r_4 + r_7 r_7) + Y(r_1 r_2 + r_4 r_5 + r_7 r_8) + (Z-h)(r_1 r_3 + r_4 r_6 + r_7 r_9)}{X(r_3 r_1 + r_6 r_4 + r_9 r_7) + Y(r_3 r_2 + r_6 r_5 + r_9 r_8) + (Z-h)(r_3 r_3 + r_6 r_6 + r_9 r_9)} \tag{7}$$
$$Y' = -h \frac{X(r_2 r_1 + r_5 r_4 + r_8 r_7) + Y(r_2 r_2 + r_5 r_5 + r_8 r_8) + (Z-h)(r_2 r_3 + r_5 r_6 + r_8 r_9)}{X(r_3 r_1 + r_6 r_4 + r_9 r_7) + Y(r_3 r_2 + r_6 r_5 + r_9 r_8) + (Z-h)(r_3 r_3 + r_6 r_6 + r_9 r_9)} \tag{8}$$

da invertire, ma di forma complessa.

Tuttavia l'equazione è molto più semplice del previsto, considerando le proprietà della matrice di rotazione. Ogni riga e colonna della matrice è un vettore ortonormale. Pertanto il modulo di ogni riga e di ogni colonna vale uno e il prodotto scalare di due righe o di due colonne da sempre zero.

Esempio: $r_1 r_1 + r_4 r_4 + r_7 r_7 = 1$, $r_1 r_2 + r_4 r_5 + r_7 r_8 = 0$, $r_1 r_3 + r_4 r_6 + r_7 r_9 = 0$ ...

Il risultato notevole è pertanto:

$$X' = -h \frac{X}{Z-h} \tag{9}$$
$$Y' = -h \frac{Y}{Z-h} \tag{10}$$

come si voleva dimostrare.

Sparendo i parametri che dipendono dalla calibrazione questa relazione vale per qualunque sistema di riferimento, indipendentemente dalla calibrazione e dal modo in cui l'IPM è stata calcolata (calibrazione, LUT, matrice omografa).

La relazione può anche essere scritta nella forma:

$$\frac{X'}{X} = \frac{Y'}{Y} \tag{11}$$

Il coeficiente angolare vale effettivamente

$$m' = \frac{X'}{Y'} = \frac{X}{Y} \tag{12}$$

e come ci aspettavamo tutti gli oggetti verticali convergono verso la proiezione del pin-hole sul piano.

Estendiamo il sistema al caso generico di una camera posta nelle coordinate (x_0, y_0, z_0). L'equazione 9 diventa:

$$X' = -z_0 \frac{X - x_0}{Z - z_0} + x_0 \tag{13}$$
$$Y' = -z_0 \frac{Y - y_0}{Z - z_0} + y_0 \tag{14}$$

Pertanto, tra diverse camere poste in posizioni diverse del mondo, le proiezioni dei punti (X', Y') saranno in posti differenti.

La relazione può perdere una incognita infatti

$$\frac{X' - x_0}{X - x_0} = \frac{Y' - y_0}{Y - y_0} \tag{15}$$

Il coefficiente angolare generale, congiungente il punto (X', Y') con la base (X, Y) (e di conseguenza (x_0, y_0)), vale:

$$m = \frac{X' - X}{Y' - Y} = \frac{X - x_0}{Y - y_0} \tag{16}$$

o la retta (scritta in forma implicita in funzione di X e Y) è:

$$\frac{1}{X - x_0} X' - \frac{1}{Y - y_0} Y' + \frac{y_0}{Y - y_0} - \frac{x_0}{X - x_0} = 0 \tag{17}$$

ricordando che X e Y combaciano tra le due immagini (è il punto sul terreno).

È da ricordare che nel testo sono state usate X, Y, Z e sopratutto X', Y' supponendo che siano espresse nella stessa unità di misura. È ovvio che quando si lavora con immagini, X', Y', sono rimappate da metri a pixel, e perciò bisogna tenere conto di un qualche fattore moltiplicativo e un offset.

4 Estensione a più camere

Prendiamo il caso di due camere allineate separate da una baseline b ($x_1 = x_2 = 0$, $y_1 = \frac{b}{2}$, $y_2 = -\frac{b}{2}$, e $z_1 = z_2 = h$). Le equazioni per i due punti sono:

$$X'_1 = X'_2 = -h \frac{X}{Z - h} = X' \tag{18}$$

$$Y'_1 = -h \frac{Y - \frac{b}{2}}{Z - h} + \frac{b}{2} \tag{19}$$

$$Y'_2 = -h \frac{Y + \frac{b}{2}}{Z - h} - \frac{b}{2} \tag{20}$$

da cui è possibile ottenere la posizione del punto acquisite da una camera su quella dell'altra. L'equazione è infatti:

$$Y'_2 = Y'_1 - b \left(\frac{h}{Z - h} + 1 \right) = Y'_1 + b \frac{Z}{h - Z} \tag{21}$$

Nel caso più classico di $Z < h$ il secondo pezzo dell'equazione è positivo e $Y'_2 > Y'_1$. In questo modo è possibile, dato un potenziale punto su una immagine IPM, verificare per quale data altezza tale punto è soddisfatto in una seconda immagine.

L'equazione 21 può essere invertita, e permette di calcolare l'altezza del punto del mondo data la disparità $\Delta = Y'_2 - Y'_1$:

$$Z = h \left(1 - \frac{b}{\Delta + b} \right) = h \frac{\Delta}{\Delta + b} \tag{22}$$

che vuol dire che punti alla stessa altezza Z nel mondo, possiedono la stessa disparità Δ nell'immagine IPM.

Quando la disparità $\Delta = 0$ correttamente $Z = 0$. Man mano che la disparità cresce l'equazione tende alla singolarità in cui $Z = h$.

Si può ricavare in maniera analoga la coordinata X usando tutte informazioni comunque conosciute:

$$X = X' \frac{b}{\Delta + b} \tag{23}$$

L'angolo formato dai due segmenti vale:

$$\tan\alpha = -b\frac{X}{X^2+Y^2-\left(\frac{b}{2}\right)^2} \tag{24}$$

anche se poi questo diventa un semplice problema di geometria (angolo in un triangolo).

Esistono curve equi-angolo, in cui α risulta costante. L'angolo non dipende dall'altezza della camera. Non è possibile dato solo l'angolo ricavare la posizione dell'ostacolo nel mondo. È possibile, dato l'angolo generato dall'ostacolo e la sua posizione nel mondo, determinare la baseline del sistema.

5 Estensione al motion stereo

Nel nostro caso le due camere sono allineate ma traslate lungo la coordinata X (nello spazio e nel tempo in generale).

Definiamo perciò $x_1 = 0$, $x_2 = d$ e $y_1 = y_2 = 0$, e $z_1 = z_2 = h$.

Questa volta le equazioni 13 diventano:

$$X'_1 = -h\frac{X}{Z-h} \tag{25}$$

$$X'_2 = -h\frac{X-d}{Z-h} + d \tag{26}$$

$$Y'_1 = Y'_2 = Y' = -h\frac{Y}{Z-h} \tag{27}$$

e si possono fare le medesime considerazioni di sopra, scambiano le due coordinate, tra cui:

$$X'_1 = X'_2 - d\left(\frac{h}{Z-h}+1\right) = X'_2 + d\frac{Z}{h-Z} \tag{28}$$

da cui la precedente equazione sulla disparità vale:

$$Z = h\frac{\Delta}{\Delta+d} \tag{29}$$

6 Problemi del matching

La trasformazione non ha carattere simile, perciò oggetti che in prospettiva erano simili, riproiettati possono risultate differenti.

Un tipico oggetto a X costante e Y variabile è un oggetto con X' costante e Y' che varia lungo una retta lungo l'asse delle Y. L'unica differenza dell'immagine originale è il fattore di scala. Tuttavia lo spostamento lungo Z cambia la forma dell'oggetto.

Tra due immagini che inquadrano la stessa scena il fattore di scala di Y è lo stesso se le camere sono alla stessa altezza (requisito fondamentale del matching, vedi equazione 21).

Ogni punto si sa dove sarà il suo sottostante o sovrastante lungo una retta.

Per ogni punto si può calcolare una probabilità di baseline dalla 21 con Z che va da 0 a un certo valore massimo. Se non si usa un punto ma una finestra va bene lo stesso, ma con il rischio che più la finestra è grande più i piccoli oggetti non vengano rilevati.

Se si rimappa a questo punto l'immagine radialmente (usando la proiezione del pinhole) si dovrebbe riuscire a vedere una specie di profilo (rettificazione della WIPM) su cui vivono gli oggetti.

La WIPM può essere trasformata in una proiezione cilindrica? Chiamiamo ϑ l'angolo tale che

$$X = \rho\cos\vartheta \tag{30}$$

$$Y = \rho\sin\vartheta \tag{31}$$

per cui $X^2 + Y^2 = \rho^2$. Dalla equazione 9, si vede che ϑ dipende solo da X e Y e non da Z. La variazione di Z non fa cambiare ϑ.

$$\tan\vartheta = \frac{Y}{X} \tag{32}$$

e definito

$$\rho_g = \sqrt{X^2 + Y^2} \tag{33}$$

si può dire che

$$\rho = \left|\frac{h}{h - Z}\right| \rho_g \tag{34}$$

e ϑ costante.