Trasformazione prospettica

Stefano Ceroni, Sara Toia

Luglio 2011

1 La camera e la matrice di proiezione

La camera è definita tramite una terna di vettori:

- $\mathbf{eye} = (eye_x, eye_y, eye_z)$ che rappresenta la posizione della camera nel mondo;
- $\mathbf{at} = (at_x, at_y, at_z)$ che rappresenta il punto in cui sta guardando la camera;
- $\mathbf{up} = (up_x, up_y, up_z)$ che definisce l'inclinazione laterale, detto rollio.

Tramite questi tre vettori è possibile ottenere la terna che rappresenta il sistema di riferimento della camera:

$$\mathbf{F} = \frac{\mathbf{at} - \mathbf{eye}}{\|\mathbf{at} - \mathbf{eye}\|}$$

$$\mathbf{S} = \mathbf{F} \times \mathbf{up}$$

$$\mathbf{U} = \mathbf{S} \times \mathbf{F}$$

Bisogna ora costruire una matrice che proietti i punti in coordinate mondo sul piano immagine; la matrice di proiezione è composta da due matrici:

- la matrice di trasformazione;
- la matrice prospettica.

La matrice di trasformazione porta i punti dal sistema di riferimento mondo a quello della camera tramite una traslazione:

$$T = \begin{bmatrix} 1 & 0 & 0 & -eye_x \\ 0 & 1 & 0 & -eye_y \\ 0 & 0 & 1 & -eye_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

e una rotazione nel sistema di riferimento della camera, definito della base ortonormale:

$$R = \begin{bmatrix} S_x & S_y & S_z & 0 \\ U_x & U_y & U_z & 0 \\ F_x & F_y & F_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

La matrice prospettica provoca la deformazione causata dal frustum di visione: in una proiezione prospettica, le rette parallele diventano incidenti all'infinito e quindi si ha una deformazione dell'immagine che permette, ad esempio, di rappresentare gli oggetti più vicini di dimensioni maggiori rispetto agli oggetti più lontani:

$$P = \begin{bmatrix} \frac{f}{aspect} & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & \frac{far+near}{near-far} & \frac{2(far \cdot near)}{near-far} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{3}$$

dove

$$f = \arctan\left(\frac{\phi}{2}\right)$$

ϕ è l'angolo verticale di visione, *aspect* è il rapporto tra la larghezza e l'altezza del piano immagine, *near* e far rappresentano rispettivamente le distanze tra il punto *eye* e il piano più vicino del frustum di visione e la distanza tra *eye* e il piano più lontano del frustum di visione[1]. La matrice di proiezione finale è la composizione delle trasformazioni e quindi il prodotto delle matrici, ricordando che il prodotto dev'essere eseguito in ordine inverso rispetto a quello di applicazione:

$$\begin{aligned} Proj &= P \cdot (R \cdot T) \\ &= P \cdot R \cdot T \end{aligned} \tag{4}$$

1.1 Proiezione sul piano immagine

Un vettore in coordinate mondo (in tre dimensioni) è composto da tre componenti più una quarta se è in coordinate omogenee. L'insieme dei punti in coordinate omogenee è composto da tutti i punti in coordinate non omogenee più i punti $[x,\, y,\, z,\, 0]$ che rappresentano i punti all'infinito. Per passare dalle coordinate non omogenee (cartesiane) a quelle omogenee è sufficiente aggiungere come quarta coordinata il valore 1:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

[1] Il frustum di visione è il tronco di piramide che rappresenta il volume entro il quale gli oggetti della scena sono inquadrati dalla camera. Tale solido ha quindi una base minore che è il piano *near* e una base maggiore che è il piano far.

Moltiplicando un punto in coordinate omogenee con la matrice di proiezione $Proj$, otteniamo un punto proiettato in coordinate normalizzate:

$$Proj \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \\ w' \end{bmatrix}$$

Riportiamo il punto alla forma canonica dividendo le prime tre componenti per la quarta:

$$\begin{bmatrix} x'/w' \\ y'/w' \\ z'/w' \\ w'/w' \end{bmatrix} = \begin{bmatrix} x'' \\ y'' \\ z'' \\ 1 \end{bmatrix}$$

ottenendo le prime tre componenti in coordinate cartesiane normalizzate:

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix}$$

A questo punto ci basta proiettare il punto sul piano immagine, eliminando la terza componente:

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix}$$

L'ultimo passo consiste nel rimappare le coordinate normalizzate nelle coordinate finestra. Supponiamo che le dimensioni della finestra siano *width* per la larghezza e *height* per l'altezza:

$$\begin{bmatrix} x''' \\ y''' \end{bmatrix} = \begin{bmatrix} \frac{\left(x''+1\right) width}{2} \\ \frac{\left(y''+1\right) height}{2} \end{bmatrix}$$

e si ottengono le coordinate schermo.