Image to World Conversion

Michael Rencheck

The Pinhole Camera Model:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \left(R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + t \right)$$

A is the intrinsic camera matrix

For the application of ball tracking, the image coordinates and the height of the ball is known, so we need to solve for x_w, y_w, and s.

So solving for the world coordinates:

$$A^{-1}R^T \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} s - R^T t = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}$$

From here we can solve for s directly by first multiplying the matrices:

$$M_1 = A^{-1}R^T \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

$$M_2 = R^T t$$

$$M_1 s - M_2 = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}$$

M1 and M2 are both 3x1 matrices and since z_w is known we can use the third row of each matrix to solve for s:

$$M_1[3,1] * s - M_2[3,1] = z_w$$

$$s = \frac{z_w + M_2[3,1]}{M_1[3,1]}$$

Now since s is computed, it can be substituted back in to directly solve for x_w and y_w

$$M_1[1,1] * s - M_2[1,1] = x_w$$

$$M_1[2,1] * s - M_2[2,1] = y_w$$