RECONSTRUCTION OF A CRICKET BALL USING SINGLE CAMERA

Varun Gupta

June 6, 2005

1 Problem Definition

Our aim is to create an automatic third umpire for giving lbw decisions in cricket matches. Reconstruction of any 3-D object requires atleast 2 non-degenerate views. However if we place more than one cameras to find the location of the ball at any instant during delivery, we need to ensure that the pictures of the ball taken are taken at exactly the same instant. This can be done by synchronising the cameras in hardware. However we wish to develop a purely software based system.

2 Solution I

Let us for the moment assume that the ball follows a trajectory which is quadratic in all three dimensions. That is:

$$x_t = x_0 + u_x t + \frac{1}{2} a_x t^2$$

$$y_t = y_0 + u_y t + \frac{1}{2} a_y t^2$$

$$z_t = z_0 + u_z t + \frac{1}{2} a_z t^2$$

The trajectory can be reconstructed if the initial position, velocity and acceleration values are known. Given (x_t, y_t, z_t, t) for 3 points on the trajectory, we can determine these parameters. However, if we have frames from a single camera then we can only get (x_t, y_t, z_t) as

$$\begin{bmatrix} x_t & y_t & z_t \end{bmatrix}^T = \begin{bmatrix} x_c & y_c & z_c \end{bmatrix}^T + \lambda_t \begin{bmatrix} d_{x_t} & d_{y_t} & d_{z_t} \end{bmatrix}^T$$

where (x_c, y_c, z_c) is the camera center. That is we have an additional parameter λ_t to estimate for every data point. So with n points, we have $9 + n$ parameters to estimate and $3n$ equations. Solving for n,

$$\begin{aligned} 9 + n &< 3n \\ n &> 4.5 \\ n &>= 5 \end{aligned}$$

However we still have a depth scale ambiguity. As shown by the following example.
Example: Consider the setting of Figure 1. The observer is standing at origin O. The only acceleration is an acceleration a_z in the negative Z direction. A ball is thrown from P_1 : $(-v^2/a_z, d, 0)$ with a speed of v at an angle of 45 degrees as shown. The observer (considering he has a single eye), cannot distinguish between this trajectory and the trajectory of a ball thrown from P_2 : $(-2v^2/a_z, 2d, 0)$ with speed $v\sqrt{2}$ (or with speed v but acceleration $a_z/2$)!

To determine the parameters we need to enforce one more constraint. For our case, specifying the moment at which ball touches the pitch ($z = 0$) does the job. (Therefore we need to have the frame in which ball touches the ground). For unambiguity, the observer should not be on $z = 0$ plane, which is usually the case.

3 Solution 2

A very novel way to reconstruct the trajectory of the ball without the assumption of quadratic nature and using only one camera is using a mirror to create a virtual camera 2. Lets say a point P can be be seen by the camera C. Place a mirror, M in the *field of view (FoV)* of C such that the image P' of P in M is also in the FoV of C. This is equivalent to having two distinct cameras (with the same calibration matrix but different rotation and translation matrices) seeing the same point P. The virtual camera C' is the image of C in M. The image of M on the viewing plane of C is the

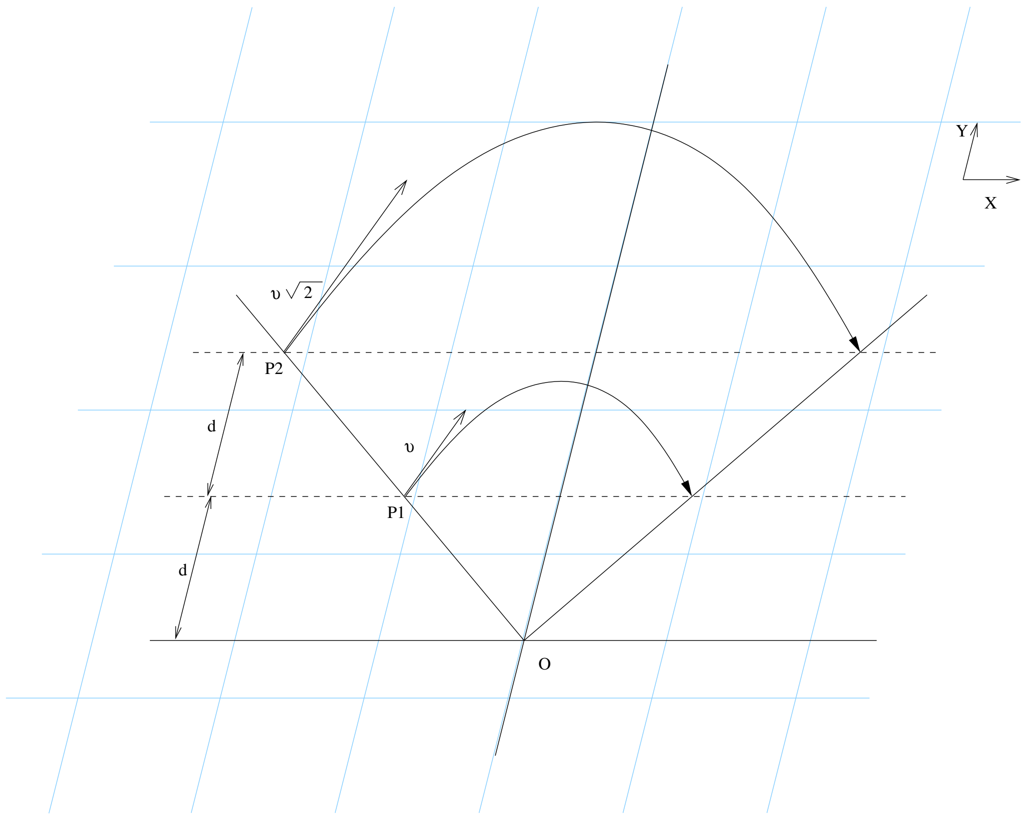

Figure 1: Depth scale ambiguity with one camera

image of the real world on the viewing plane of C'.

Obtaining M: To find the camera matrix of C' we need to find the equation of the plane of M's surface. Let this be expressed as

$$(x - \overline{P}_M).\overrightarrow{n}_M = 0$$

where, $\overline{P}_M$ is a point on the plane and $\overrightarrow{n}_M$ is the direction of normal to the plane. $\overrightarrow{n}_M$ can be obtained by the images on the viewing plane of C of two points $\overline{X}, \overline{Y}$ and their images $\overline{X'}, \overline{Y'}$ in M (we do not need to know the 3D coordinates of $\overline{X}$ and $\overline{Y}$). The plane containing $\overline{C}$, $\overline{X}$ and $\overline{X'}$ is perpendicular to the mirror's plane and so is the plane containing $\overline{C}$, $\overline{Y}$ and $\overline{Y'}$. The equations of these planes can be determined by shooting back rays from $\overline{C}$ through $\overline{X}$,$\overline{X'}$,$\overline{Y}$ and $\overline{Y'}$. Let these rays be

$$\begin{aligned} r_X &= \overline{C} + \lambda(\overrightarrow{d}_X) \\ r_{X'} &= \overline{C} + \lambda(\overrightarrow{d}_{X'}) \\ r_Y &= \overline{C} + \lambda(\overrightarrow{d}_Y) \\ r_{Y'} &= \overline{C} + \lambda(\overrightarrow{d}_{Y'}) \end{aligned}$$

Let $\overrightarrow{n}_X = \overrightarrow{d}_X \times \overrightarrow{d}_{X'}$ and $\overrightarrow{n}_Y = \overrightarrow{d}_Y \times \overrightarrow{d}_{Y'}$ be the normals to these planes. The normal to M is : $\overrightarrow{n}_M = \overrightarrow{n}_X \times \overrightarrow{n}_Y$.

Now to completely determine the mirror, we need the point $\overline{P}_M$. If $\overline{X}$ is a point in the real world and $\overline{X'}$ its image, then the required point is $\overline{P}_M = \frac{\overline{X}+\overline{X'}}{2}$. However now we do need the exact coordinates $\overline{X}$. To get $\overline{X'}$, shoot a ray $r_{X'}$ through $\overline{C}$ passing through $\overline{X'}$. Take the intersection of $r_{X'}$ with

$$x = \overline{X} + \lambda\overrightarrow{n}_M$$

the line passing through $\overline{X}$ and normal to M. This completes the equation of M and therefore we can obtain the rotation and translation matrices for the virtual camera, C'.

Reconstructing a point from its two images: The process of reconstructing a point from its 2 images on the viewing plane is illustrated in Figure 2. Shoot a ray R through the image of P on the viewing plane of C. This ray will pass through P. Now shoot another ray R' through the image of P' on the viewing plane of C. This ray will intersect the mirror M at a point where we construct the normal n and reflect R' to get the ray R''. The ray R'' will also pass through P. By taking the intersection of rays R and R' (or the point closest to both the rays, since due to numerical errors the rays may not intersect), we can reconstruct the 3D coordinates of P.

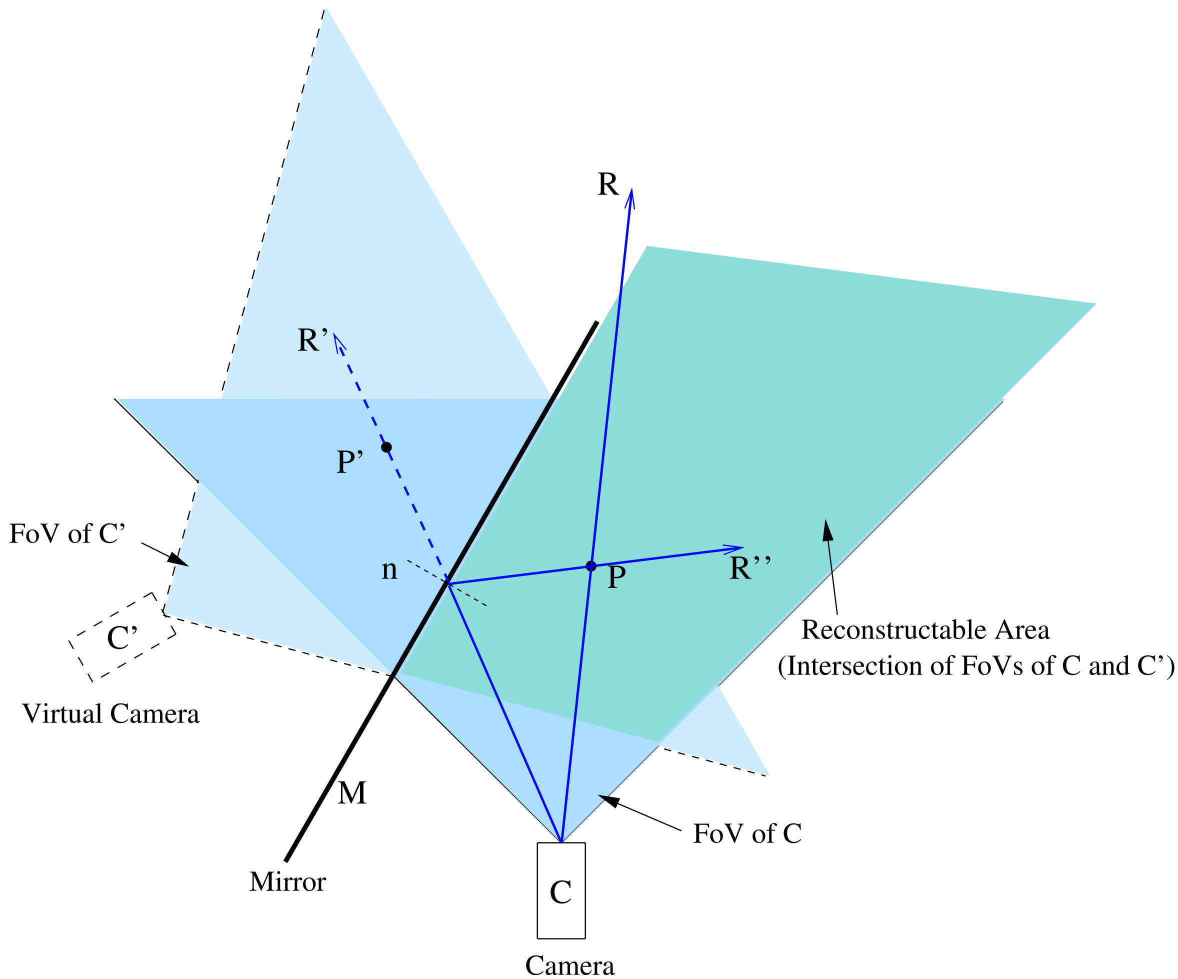

Figure 2: Basic Idea