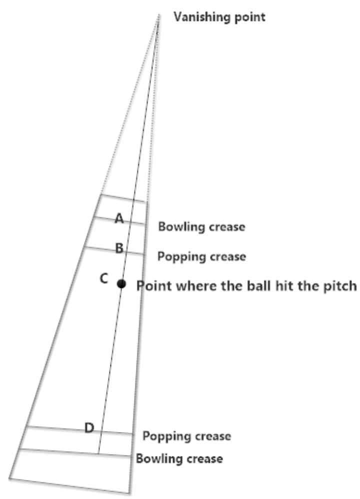
CS 763: Assignment 1

Ayush Baid
Niranjan Thakurdesai
Jainesh Doshi 12D070014

February 2, 2016

Question 1

We'll use vanishing points and cross-ratios to solve this problem.



As all the sides of the pitch are visible in the image, we can extend the lengths of the pitch to find their vanishing point. Draw a line joining the vanishing point and the point where the ball pitched. This line will be parallel to the lengths of the pitch. Let A and B be the intersection points of this line with the bowling crease and the popping crease at the batsman's end and C be the point where the ball pitched. Let D be the intersection point of the line with the popping crease at the bowler's end.
As cross-ratio is invariant under perspective projection, we'll use this property to find the required distance. The lengths $(AC)_i$, $(BD)_i$, $(AD)_i$ and $(BC)_i$ can be calculated from the image (subscript i denotes that the distances are in the pixel coordinate system). The lengths $(AB)_w$ and $(BD)_w$ are known, as the dimensions of the pitch are known (subscript w denotes that the distances

are in the world coordinate system). Using the aforementioned property of cross-ratios,

$$\frac{(AC)_i(BD)_i}{(AD)_i(BC)_i} = \frac{(AC)_w(BD)_w}{(AD)_w(BC)_w}$$

$$\text{Let } (CD)_w \text{be } x.$$

$$\therefore \frac{(AC)_i(BD)_i}{(AD)_i(BC)_i} = \frac{((AD)_w - x)(BD)_w}{(AD)_w((BD)_w - x)}$$

$$\therefore x = \frac{(AD)_w(BD)_w[(AC)_i(BD)_i - (AD)_i(BC)_i]}{(AC)_i(BD)_i(AD)_w - (AD)_i(BC)_i(BD)_w}$$

Now,

$$(AB)_w = 122 \text{ cm}, (BD)_w = 2012 - 2 \cdot 122 = 1768 \text{ cm}, (AD)_w = 2012 - 122 = 1890 \text{ cm}$$

$$\therefore x = \frac{3341520[(AC)_i(BD)_i - (AD)_i(BC)_i]}{1890(AC)_i(BD)_i - 1768(AD)_i(BC)_i} \text{ (in cm)}$$

Let the perpendicular distance from Q to the wickets at the bowler's end be d.

$$\therefore d = x + 122$$

$$= \frac{3341520[(AC)_i(BD)_i - (AD)_i(BC)_i]}{1890(AC)_i(BD)_i - 1768(AD)_i(BC)_i} + 122 \text{ (in cm)}$$

Note that we just need the distances in the pixel and camera coordinate systems to compute the perpendicular distance. We aren't using any extrinsic or intrinsic parameters but just the invariance property of cross-ratio, hence the camera need not be calibrated to make this computation.

Question 2

Let $\mathbf{P_1}$ and $\mathbf{P_2}$ be the camera coordinates of the physical point in the two coordinate systems respectively. Now,

$$\mathbf{P_2} = \mathbf{RP_1} \tag{1}$$

$\mathbf{p_1}$ and $\mathbf{p_2}$ are the pixel coordinates. Let $\mathbf{p_1} = (x_1, y_1)$ and $\mathbf{p_2} = (x_2, y_2)$. We have

$$\begin{pmatrix} \hat{x_1} \\ \hat{y_1} \\ \hat{z_1} \end{pmatrix} = \mathbf{KP_1}, \begin{pmatrix} \hat{x_2} \\ \hat{y_2} \\ \hat{z_2} \end{pmatrix} = \mathbf{KP_2} = \mathbf{KRP_1} \tag{2}$$

where

$$x_1 = \frac{\hat{x_1}}{\hat{z_1}}, y_1 = \frac{\hat{y_1}}{\hat{z_1}}, x_2 = \frac{\hat{x_2}}{\hat{z_2}}, y_2 = \frac{\hat{y_2}}{\hat{z_2}}. \tag{3}$$

Note that

$$\mathbf{K} = \begin{pmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{pmatrix} \tag{4}$$

The third column of $\mathbf{K}$ cannot be written as a linear combination of the first two columns as the third element of the third column is non-zero while that of both the first and the second column is 0. Hence, $\mathbf{K}$ is invertible.From equation (2), we get

$$\begin{pmatrix} \hat{x_2} \\ \hat{y_2} \\ \hat{z_2} \end{pmatrix} = \mathbf{KRK^{-1}} \begin{pmatrix} \hat{x_1} \\ \hat{y_1} \\ \hat{z_1} \end{pmatrix} \tag{5}$$

From (3), we get

$$\begin{pmatrix} x_2\hat{z_2} \\ y_2\hat{z_2} \\ \hat{z_2} \end{pmatrix} = \mathbf{KRK^{-1}} \begin{pmatrix} x_1\hat{z_1} \\ y_1\hat{z_1} \\ \hat{z_1} \end{pmatrix} \tag{6}$$

Thus,

$$\hat{z_2} \begin{pmatrix} \mathbf{p_2} \\ 1 \end{pmatrix} = \hat{z_1}\mathbf{KRK^{-1}} \begin{pmatrix} \mathbf{p_1} \\ 1 \end{pmatrix} \text{ and } \hat{z_2} = (\mathbf{KRK^{-1}})_3^T \hat{z_1} \tag{7}$$

$$\implies \begin{pmatrix} \mathbf{p_2} \\ 1 \end{pmatrix} = \frac{\mathbf{KRK^{-1}} \begin{pmatrix} \mathbf{p_1} \\ 1 \end{pmatrix}}{(\mathbf{KRK^{-1}})_3^T \begin{pmatrix} \mathbf{p_1} \\ 1 \end{pmatrix}} \tag{8}$$

If $\mathbf{K}$ is known, we can solve for $\mathbf{R}$ using the orthogonal Procrustes method.

Question 3

The projection will be the point where the line joining the pinhole and the 3D point intersects the hemisphere. Let (X, Y, Z) be the 3D point and $(-tX, -tY, -tZ)$ be the projected point. The equation of the hemisphere is

$$x^2 + y^2 + (z + f)^2 = r^2 \tag{9}$$

where r is the radius of the hemisphere.
Substituting $(x, y, z) = (-tX, -tY, -tZ)$ in the above equation and solving the quadratic equation for t, we get

$$t = \frac{fZ \pm \sqrt{f^2Z^2 - (f^2 - R^2)(X^2 + Y^2 + Z^2)}}{X^2 + Y^2 + Z^2} \tag{10}$$

Suppose the 3D point is on the Z axis. Setting X and Y to 0 in the above equation we get

$$t = f \pm R \tag{11}$$

However, in this case, $t = f + R$. Thus,

$$t = \frac{fZ + \sqrt{f^2Z^2 - (f^2 - R^2)(X^2 + Y^2 + Z^2)}}{X^2 + Y^2 + Z^2} \tag{12}$$

and

$$x = -tX, y = -tY, z = -tZ \tag{13}$$

Question 4

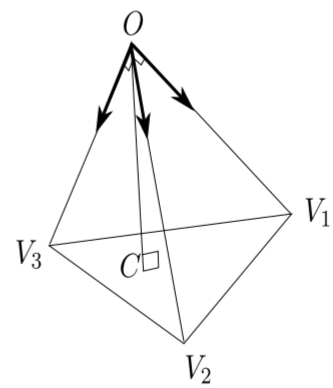


Figure 1: V_1, V_2, V_3 are the vanishing points of a set of Mutually perpendicular directions

In the figure we have V_1, V_2, V_3, C instead of $Q, R, S\&o$. Thus $OQ, OR\&OS$ as mutually perpendicular, while Q, R, S are coplanar on the image plane. $OR - OQ$ will lie in the plane containing $Q, R\&Q$ and OS will the normal to this plane.
$OR - OQ$ lies in the image plane whileOo is optical axis, hence they are orthogonal.
As from above the plane in which OS, So, Oo lie is orthogonal to $OR - OQ$, hence the triangle formed by them too is orthogonal to $OR - OQ = QR$.
The proof rests on the fact that the points Q, R, S are the vanishing points of perpendicular directions, and as we already know that the vanishing points of parallel lines are the same, hence the passing of line through O is not necessary but we will always have a line passing through O that is parallel to the taken line and that will lead us to the vanishing point. Hence the proof remains true in other conditions too.

Question 5

Q-5. Cost functional $E(R,t,\alpha) = \sum_{i=1}^{n} \|p_{2i} - \alpha R p_{1i} - t\|_2^F$

Let $R' = \alpha R \Rightarrow (R')^T R' = \alpha^2$

$$\therefore E(R',t) = \sum_{i=1}^{n} \|p_{2i} - R' p_{1i} - t\|_2^F$$

$$= \sum_{i=1}^{n} tr\left[\left(p_{2i}^T - \cancel{R} p_{1i}^T R'^T - t^T\right)\left(p_{2i} - R' p_{1i} - t\right)\right]$$

Neglecting terms not dependent on R' & t

$$E(R',t) = \sum_{i=1}^{n} tr\left(-p_{2i}^T R' p_{1i} - p_{2i}^T t - p_{1i}^T R'^T p_{2i} + \alpha^2 p_{1i}^T p_{1i} + p_{1i}^T R'^T t - t^T p_{2i} + t^T R' p_{1i} + t^T t\right)\Big]$$

$$\cancel{\frac{\partial E}{\partial t} = \sum_{i=1}^{n}}$$

$$= \sum_{i=1}^{n} tr\left(-2p_{2i}^T R' p_{1i} - 2p_{2i}^T t + \alpha^2 p_{1i}^T p_{1i} + 2p_{1i}^T R'^T t + t^T t\right)$$

$$\frac{\partial E}{\partial t} = \sum_{i=1}^{n}\left(-2p_{2i} + 2R' p_{1i} + 2t\right) = 0$$

$$\Rightarrow \cancel{2} nt = \sum_{i=1}^{n}\left(p_{2i} - R' p_{1i}\right) = 0$$

$$\Rightarrow \boxed{t = \overline{p_2} - R'\overline{p_1}}$$

where $\overline{p_2} = mean(p_{2i})$ and same for $\overline{p_1}$

$$E(R', t) = \sum_{i=1}^{n} \| p_{2i} - R' p_{1i} - t \|_F^2$$

let $p'_{2i} = p_{2i} - \bar{p_2}$ and $p'_{1i} = p_{1i} - \bar{p_1}$

$$E(R', \alpha) = \| P_2' - \alpha R P_1' \|_F^2$$

$$= tr\left[\left(P_2'^T - \alpha P_1'^T R^T \right) \left(P_2' - \alpha R P_1' \right) \right]$$

neglecting terms independent of α & R

$$E(R, \alpha) = -2\alpha \, tr\left(P_2'^T R P_1' \right) + \alpha^2 tr\left(P_1'^T P_1' \right)$$

Question 6

M for dataset 1:

$$\begin{pmatrix} 0.2905 & 0.0532 & -0.1866 & -0.6283 \\ -0.0881 & 0.3264 & -0.0881 & -0.6010 \\ -0.0002 & -0.0002 & -0.0002 & 0.0021 \end{pmatrix}$$

M for dataset 2:

$$\begin{pmatrix} 0.0087 & 0.0011 & -0.0039 & 0.9986 \\ 0.0001 & 0.0092 & 0.0005 & -0.0520 \\ -0.0000 & -0.0000 & -0.0000 & -0.0027 \end{pmatrix}$$

M for dataset 2 with additive gaussian noise:

$$\begin{pmatrix} -0.0001 & 0.0011 & 0.0037 & -0.9455 \\ 0.0017 & -0.0025 & 0.0015 & -0.3254 \\ -0.0000 & -0.0000 & -0.0000 & 0.0022 \end{pmatrix}$$

To validate the results, the given world coordinates f3D are projected to image coordinates and compared with the actual values of f2D. Mean relative error and max relative error are used as metrics (homogeneous coordinates not included).

Dataset 1: mean relative error $= 3.545e-14$; max relative error $= 3.0721e-13$.
Dataset 2: mean relative error $= 0.0035$; max relative error $= 0.1084$.
Dataset 2 with noise: mean relative error $= 3.9744$; max relative error $= 0.2382$.

Question 7

Part (a)

$$\begin{pmatrix} 1.1283 & 0.0385 & -57.1714 \\ 0.0702 & 1.0931 & -40.8860 \\ 0.0005 & 0.0002 & 1.0000 \end{pmatrix}$$

Given Homography Matrix

$$\begin{pmatrix} 1.1714 & 0.0600 & -65.0128 \\ 0.08 & 1.1349 & -46.1425 \\ 0.0005 & 0.0002 & 1.0000 \end{pmatrix}$$

Estimated Homography Matrix

Figure 2: Original, given homography warped and estimated homography warped images

Part (b)

$$\begin{pmatrix} 0.8910 & -0.0359 & 49.5625 \\ 0.0650 & 0.9278 & 34.1957 \\ -0.0002 & -0.0001 & 1.0000 \end{pmatrix}$$

Estimated Homography Matrix

$$\begin{pmatrix} 0.7432 & -0.1477 & 78.8407 \\ -0.0964 & 0.7726 & 50.7226 \\ -0.0003 & 0.0004 & 1.0000 \end{pmatrix}$$

Homography Matrix estimated from Warped image

Note:
Please refer to the attached images for better resolutions of the same.
Bi-linear interpolation was used in image warping.

Figure 3: First, Second and estimated homography warped images