Pitch*f/x* Nine Parameter Trajectory Model Fitting

R. Pendleton
Sportvision LLC

November 26, 2011

1 The Pitch*f/x* Measurement System

Each pitch is viewed by two cameras whose video is analyzed by a computer. The computer time stamps each frame of video and processes them to find blobs in the image. The result of each image is then a measurement consisting of time, and screen position. These data are further analyzed to generate a trajectory model for each pitch.

1.1 The Data and Model

The data for a pitch consists of a set of screen pts, s'_i, at time, t_i from a camera with parameters, λ. With the two camera denoted by A and B the data can be notated as:

$$\begin{aligned}\mathbf{d}'_A &= \left\{\left(t_{A_i}, \mathbf{s}'_{A_i}\right), \lambda_A\right\}_{CamA} \\ \mathbf{d}'_B &= \left\{\left(t_{B_j}, \mathbf{s}'_{B_j}\right), \lambda_B\right\}_{CamB}\end{aligned} \tag{1}$$

The first thing we do to the measured blob data eq (1) is convert from *distorted* screen pixels to *undistorted* screen pixels. We use a radial distortion model the details of which are not given here. The result is a set of data consisting of *undistorted* pixel locations and time from two cameras:

$$\begin{aligned}\mathbf{d}_A &= \left\{\left(t_{A_i}, \mathbf{s}_{A_i}\right), \lambda_A\right\}_{CamA} \\ \mathbf{d}_B &= \left\{\left(t_{B_j}, \mathbf{s}_{B_j}\right), \lambda_B\right\}_{CamB}\end{aligned} \tag{2}$$

We want to fit a constant acceleration trajectory model to the data. The model we will use is given by

$$\mathbf{x}_M(t) \quad = \quad \begin{pmatrix} X_o \\ Y_o \\ Z_o \end{pmatrix} + (t - t_o)\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} + \frac{1}{2}(t - t_o)^2 \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{3}$$

Where for now we take t_o to be the minimum of all the values contained in the data eq (2). We can write the above model as a function of t which depends on the nine parameters $\mathbf{P}_{9P} = X_o, Y_o, Z_o, V_x, V_y, V_z, A_x, A_y, A_z$. This model is referred to as the 9P model.

1.2 Mathematical Formulation of Problem

The undistorted screen pixels, s'_i, are images of 3D *world* points - the location of the baseball at specific times along the pitch trajectory. The relationship between 3D points and undistorted screen points is well modelled using the *pin hole* camera model in combination with a *rigid body* transformation. Using *homogeneous* coordinates allows the composition of these two transformations as matrices which results in a 3×4 matrix `M` relating homogeneous 3D world points to homogeneous 2D undistorted screen points [1]:

$$k \begin{pmatrix} s_x \\ s_y \\ 1 \end{pmatrix} = \mathtt{M} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$k \begin{pmatrix} s_x \\ s_y \\ 1 \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \tag{4}$$

where k is an arbitrary overall scale factor.

Now we combine the *world to screen* transform of eq (4) with the 9P model eq (3) and the data eq (2) to generate an equation relating 3D points along a 9P model trajectory to measured screen data which have been undistorted:

$$\begin{pmatrix} s_{x_i} \\ s_{y_i} \end{pmatrix} = \begin{pmatrix} \dfrac{m_{11}X(t_i) + m_{12}Y(t_i) + m_{13}Z(t_i) + m_{14}}{m_{31}X(t_i) + m_{32}Y(t_i) + m_{33}Z(t_i) + m_{34}} \\ \dfrac{m_{21}X(t_i) + m_{22}Y(t_i) + m_{23}Z(t_i) + m_{24}}{m_{31}X(t_i) + m_{32}Y(t_i) + m_{33}Z(t_i) + m_{34}} \end{pmatrix} \tag{5}$$

which we can re-arrange into the matrix equation:

$$\begin{bmatrix} \alpha & \alpha\Delta t_i & \alpha\dfrac{\Delta t_i^2}{2} & \beta & \beta\Delta t_i & \beta\dfrac{\Delta t_i^2}{2} & \gamma & \gamma\Delta t_i & \gamma\dfrac{\Delta t_i^2}{2} \\ \alpha' & \alpha'\Delta t_i & \alpha'\dfrac{\Delta t_i^2}{2} & \beta' & \beta'\Delta t_i & \beta'\dfrac{\Delta t_i^2}{2} & \gamma' & \gamma'\Delta t_i & \gamma'\dfrac{\Delta t_i^2}{2} \end{bmatrix} \begin{pmatrix} X_o \\ V_x \\ A_x \\ Y_o \\ V_y \\ A_y \\ Z_o \\ V_z \\ A_z \end{pmatrix} = \begin{pmatrix} m_{34}s_{x_i} - m_{14} \\ m_{34}s_{y_i} - m_{24} \end{pmatrix} \tag{6}$$

With $\Delta t_i = t_i - t_o$ and,

$$\begin{matrix} \alpha = m^C_{11} - m^C_{31}s_x & \beta = m^C_{12} - m^C_{32}s_x & \gamma = m^C_{13} - m^C_{33}s_x \\ \alpha' = m^C_{21} - m^C_{31}s_y & \beta' = m^C_{22} - m^C_{32}s_y & \gamma' = m^C_{23} - m^C_{33}s_y \end{matrix} \tag{7}$$

For each blob measurement two equations as given by eq (6) are generated in the overall system matrix and thus for $N \geq 5$ measurements we generate an over-constrained linear system for the 9 parameters of eq (3) in the form:

$$\mathtt{A}\mathbf{x} = \mathbf{b} \tag{8}$$

which can be solved by the method of Least Squares for the solution vector $\mathbf{x}$.

1.2.1 Modificaiton of 9P - Known Position

We can incorporate a known initial position into the constant acceleration trajectory by a reformulating eq (6) by moving the known quantities involving X_o, Y_o, Z_o to the right hand side vector.

$$\begin{bmatrix} \alpha\Delta t_i & \alpha\frac{\Delta t_i^2}{2} & \beta\Delta t_i & \beta\frac{\Delta t_i^2}{2} & \gamma\Delta t_i & \gamma\frac{\Delta t_i^2}{2} \\ \alpha'\Delta t_i & \alpha'\frac{\Delta t_i^2}{2} & \beta'\Delta t_i & \beta'\frac{\Delta t_i^2}{2} & \gamma'\Delta t_i & \gamma'\frac{\Delta t_i^2}{2} \end{bmatrix} \begin{pmatrix} V_x \\ A_x \\ V_y \\ A_y \\ V_z \\ A_z \end{pmatrix} = \begin{pmatrix} m_{34}s_{x_i} - m_{14} - \alpha X_o - \beta Y_o - \gamma Z_o \\ m_{34}s_{y_i} - m_{24} - \alpha' X_o - \beta' Y_o - \gamma' Z_o \end{pmatrix} \tag{9}$$

2 Solution of Pitch*f/x* Parameters - Linear SVD

The SVD can be applied to the system given by (8), (6) to find the parameters, the residual and estimages of the standard errors in the paramters [2, Ch. 15 - Modeling of Data]. First the SVD of the m×n, $m > n$, matrix $\mathtt{A}$ is computed:

$$\mathtt{A} = \mathtt{U}\Sigma\mathtt{V}^T \tag{10}$$

where,

$$\mathtt{U} = \begin{bmatrix}\mathbf{u}_1 & \dots & \mathbf{u}_n\end{bmatrix} \tag{11}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & \dots & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \dots & 0 & \sigma_p \\ 0 & \dots & \dots & 0 \end{bmatrix} \tag{12}$$

$$\mathtt{V} = \begin{bmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_m\end{bmatrix} \tag{13}$$

The diagonal elements of the m×n matrix Σ contain the *singular values* of $\mathtt{A}$ of which there are at most p which are non-zero. Next the singular are edited, that is thresholded, so that only values larger than the threshold are used - the others are set to 0. Then the 9P parameters, $\mathbf{x}_{9P}$ are calculated via:

$$\begin{aligned} \mathbf{x}_{9P} &= \sum_{\substack{i=1 \\ \sigma_i \neq 0}}^{m} \frac{\mathbf{v}_i \mathbf{u}_i^T}{\sigma_i} \mathbf{b} \\ &= \mathtt{A}^{\dagger} \mathbf{b} \end{aligned} \tag{14}$$

Where, $\mathtt{A}^{\dagger}$ is the *pseudo-inverse* of $\mathtt{A}$. The $2n$ length residual vector is given by:

$$\mathbf{b}_{res} \equiv \mathbf{b} - \mathtt{A}\mathbf{x}_{9P} \tag{15}$$

and the covariance matrix is given by:

$$\mathtt{C}_{jk} = \sum_{\substack{i=1 \\ \sigma_i \neq 0}}^{m} \frac{\mathtt{V}_{ji} \mathtt{V}_{ki}}{\sigma_i^2} \tag{16}$$

References

[1] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second ed., 2004.

[2] W. H. Press, S. A. Teukolsky, W. T. Vettering, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing, Third Edition*. Cambridge University Press, 2007.