Camera Calibration from Multiple Views

Harvey Rhody  
Chester F. Carlson Center for Imaging Science  
Rochester Institute of Technology  
rhody@cis.rit.edu

March 28, 2006

Abstract
This lecture explores the geometry related to camera calibration derived from multiple views of the same scene. The fundamental matrix is derived and related to the camera matrix (matrices).
Projective Transformation

Let \( x = [x_1, x_2, x_3]^T \) be the homogeneous representation of a point in plane \( \pi \).

Let

\[
\begin{bmatrix}
  x'_1 \\
  x'_2 \\
  x'_3
\end{bmatrix}
= \begin{bmatrix}
  h_{11} & h_{12} & h_{13} \\
  h_{21} & h_{22} & h_{23} \\
  h_{31} & h_{32} & h_{33}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix}
\]

The relationship

\( x' = Hx \)

is a projective transform for any nonsingular matrix \( H \)

Since multiplication of a homogeneous coordinate vector by a nonzero scalar is irrelevant, \( H \) is specific up to a nonzero scale factor.

The projective transform has eight degrees of freedom.
Perspective Projection

The relationship between points in two planes under projective projection is illustrated in the figure below. The points are related by $x'_k = Hx_k$. 

\[ x'_k = Hx_k \]
Planar Projection

Shown below are projective projections of points on a plane to two image planes. The points are related by

\[ x = H_1X \Rightarrow X = H_1^{-1}x \]

\[ x' = H_2X = H_2H_1^{-1}x = Hx \]
Pinhole Camera Model

\[ [X, Y, Z]^T \mapsto [fX/Z, fY/Z]^T \]

Nonhomogeneous coordinates

\[
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\mapsto
\begin{bmatrix}
fX \\
fY \\
Z
\end{bmatrix}
= \begin{bmatrix}
f & 0 \\
f & 0 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]

Homogeneous Coordinates
Camera Plane Coordinates

It is typical to state the position on the camera plane in terms of offset from the principal point. Then

\[
\begin{bmatrix} X, Y, Z \end{bmatrix}^T \mapsto [fX/Z + p_x, fY/Z + p_y]^T
\]

Nonhomogeneous coordinates

\[
\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} f & p_x & 0 \\ f & p_y & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}
\]

Homogeneous Coordinates

The matrix

\[
K = \begin{bmatrix} f & p_x \\ f & p_y \\ 1 & 0 \end{bmatrix}
\]

is called the camera calibration matrix.

\[
x = K [I|0] X
\]
CCD Camera

If the position on the image is measured in pixels then the units can be rescaled by

\[
\begin{bmatrix}
  m_x & m_y \\
  m_y & 1
\end{bmatrix}
\begin{bmatrix}
  f & p_x \\
  f & p_x
\end{bmatrix}
= 
\begin{bmatrix}
  \alpha_x & x_0 \\
  \alpha_y & y_0
\end{bmatrix}
\]

where \( m_x \) and \( m_y \) are in units of pixels per length unit and \( \alpha_x = fm_x \), \( \alpha_y = fm_y \), \( x_0 = m_x p_x \), \( y_0 = m_y p_y \). For added generality we can include a skew parameter, \( s \). Then in pixel units the calibration matrix is

\[
K = 
\begin{bmatrix}
  \alpha_x & s & x_0 \\
  \alpha_y & y_0 & 1
\end{bmatrix}
\]
Imaging 3D to 2D

The Euclidean transformation between world and camera coordinates can be represented as a rotation and translation plus a projection.

\[
x = K \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = K \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} X = PX
\]

where \( \tilde{C} = [C_x, C_y, C_z]^T \) is the camera center in Euclidean coordinates.
Camera Matrix

The matrix $P$ is the *camera matrix*.

$$P = KR[I - \tilde{C}]$$

It is a $3 \times 4$ matrix that maps homogeneous world coordinates to homogeneous image coordinates. The parameters in $K$ form the *interior orientation* and those in $R$ and $C$ form the *exterior orientation*.

Given a camera matrix $P$ we can find the camera orientation and internal parameters. Let

$$P = [M| - MC] = K[R| - RC]$$

Since $M = KR$ and we know that $K$ is upper triangular, we can extract $M$ from $P$ and factor it using RQ decomposition. We can determine $C$ by multiplying the last column of $P$ by $-M^{-1}$.

Note that the camera center is a right null-vector of $P$ since it satisfies

$$PC = P \begin{bmatrix} \tilde{C} \\ 1 \end{bmatrix} = 0$$
Epipolar Geometry and the Fundamental Matrix

The epipolar geometry is the intrinsic geometry between two views of the same scene. It is independent of scene structure, and only depends on the internal camera parameters and relative pose.

The fundamental matrix $F$ is a $3 \times 3$ matrix that captures the epipolar geometry. If $x$ is the image of $X$ in the first view and $x'$ is the image of $X$ in the second view then $x$ and $x'$ are related by

$$x'^T F x = 0$$

The camera matrix $P$ for each view can be obtained, up to a homography $H$, from the fundamental matrix $F$. From this we can determine the camera interior and exterior orientation.
The two cameras with centers $C$ and $C'$ have the image planes shown. A point $X$ in space is imaged to $x$ and $x'$. The point $X$, the camera centers, and the image points $x$ and $x'$ all lie on a common plane $\pi$, called the *epipolar plane*. 
Epipolar Geometry

An image point $x$ pack-projects to a ray in three-space defined by camera center $C$ and $x$. This ray is imaged as a line $l'$ in the second view. The point $X$ must lie on the ray and must be image to a point on $l'$ in the second view. To each point $x$ in one image there must correspond a line $l'$ in the other.
Epipolar Geometry

The camera baseline intersects each image plane at the epipoles \( e \) and \( e' \). Any plane \( \pi \) that contains the baseline is an epipolar plane, and intersects the image planes in corresponding epipolar lines \( l \) and \( l' \).
Epipolar Geometry

As the position of the 3D point $\mathbf{X}$ varies, the epipolar planes “rotate” about the baseline. This family of planes is known as an epipolar pencil. All epipolar lines intersect at the epipole.
Derivation of Fundamental Matrix

Let $\pi$ be a plane in space that does not pass through either camera center. A point $x$ in the first image meets $\pi$ at a point $X$, which then project to a point $x'$ in the second image.

The set of all points $x_k$ in the first image project to points $X_k$ on $\pi$ and then to corresponding points $x'_k$ in the second image. Because the projection is by a plane the points $x_k$ and $x'_k$ must be related by a homography.

$$x' = H_\pi x$$
Derivation of Fundamental Matrix

Given the point $x'$, the epipolar line $l'$ passing through $x'$ and the epipole $e'$ can be written as

$$l' = e' \times x' = e' \times (H_{\pi}x) = [e']_x H_{\pi}x = Fx$$

Since $x'$ must lie on the epipolar line $l'$ we must have

$$x' \times l' = 0$$

This is equivalent to

$$x'Fx = 0$$

The fundamental matrix expresses the relationship between corresponding points in the two images via $x'Fx = 0$.

Although we used a plane $\pi$ in the derivation, the existence of $F$ does not depend upon the choice of the plane.
Relation of F to P and P'

Let \( P \) and \( P' \) be the camera matrices for the two images. Let \( x = PX \) be the image of a point \( X \) in the first image. Then \( X \) must lie on a ray that joins \( C \) and \( x \). This ray can be written as

\[
X_\lambda = P^\dagger x + \lambda C
\]

Note that for any point on the ray, \( PX_\lambda = PP^\dagger x + \lambda PC = Ix + 0 = x \), so all points on the ray project to the point \( x \) in the first image.

The two points \( P^\dagger x \) and \( C \) image to the points \( P'P^\dagger x \) and \( P'C \), respectively, in the second image. The epipolar line is the line joining these two points.

\[
l' = (P'C) \times (P'P^\dagger x) = e' \times (P'P^\dagger x) = [e']_x P'P^\dagger x
\]

Hence

\[
F = [e']_x P'P^\dagger
\]

where \( e' = P'C \) is the epipole in image 2.
Relation of $F$ to $P$ and $P'$

The choice of the world coordinate frame affects $P$ and $P'$ but not the fundamental matrix $F$.

If $H$ is a $4 \times 4$ matrix representing a projective transformation of 3-space then the fundamental matrices corresponding to $(P, P')$ and $(PH, P'H)$ are the same.

**Proof** If $x = PX$ and $x' = P'X$ are matching points under camera matrices $(P, P')$ corresponding to a point $X$ then they are also matching points with respect to the pair of cameras $(PH, P'H)$ corresponding to the point $\hat{X} = H^{-1}X$.

Although a pair of camera matrices determine $F$, the converse is not true. The fundamental matrix determines the camera matrices up to a right multiplication by a 3D projective transform.

This is the case because $F$ does not depend on the coordinate system but $(P, P')$ does.
Relation of $F$ to $P$ and $P'$

Theorem Let $(P, P')$ and $(\tilde{P}, \tilde{P}')$ be two pairs of camera matrices that have the same fundamental matrix $F$. Then there exists a non-singular $4 \times 4$ matrix $H$ such that $\tilde{P} = PH$ and $\tilde{P}' = P'H$.

The proof is straightforward, but will be left to the text.

Theorem A non-zero matrix $F$ is the fundamental matrix corresponding to a pair of camera matrices $(P, P')$ if and only if $P'\!\!T FP$ is skew-symmetric.

Proof The condition that $P'\!\!T FP$ is skew-symmetric is equivalent to $X^T P'\!\!T FPX = 0$ for all $X$. Setting $x' = P'X$ and $x = PX$, this is equivalent to $x'Fx = 0$, which is the defining equation of the fundamental matrix.
Canonical Form of Camera Matrices

It is common to set things up so that one camera of the pair is aligned with the world coordinate system and scaled so that \( P = [I|0] \) where \( I \) is the \( 3 \times 3 \) identity matrix.

Given a fundamental matrix \( F \) we can choose \( P' = [SF|e'] \) where \( S \) is any skew-symmetric matrix and \( e' \) is the epipole such that \( e' \cdot F = 0 \). Then \( F \) is the fundamental matrix corresponding to the pair \((P, P')\).

One common choice for \( S \) is \( S = [e'] \times \).
Computing the Fundamental Matrix

Given a set of $n$ correspondences $x_k \leftrightarrow x'_k$ in two images, the fundamental matrix $F$ satisfies $x'_k^T F x_k = 0$ for all $k = 0, 1, 2, \ldots, n - 1$.

Each correspondence generates one equation that is linear in the components of $F$.

There are eight linearly independent entries in the $3 \times 3$ matrix $F$. Hence, eight or more correspondences are sufficient to solve linearly for the entries of $F$.

If more that eight correspondences are available then the entries can be solved by least-squares techniques.
References