

# HW7 Part 2 – Uncalibrated Stereo Object Size Estimation

Amir Sabbaghziarani - 002852171

## Part 2 – Uncalibrated Stereo Object Size Estimation (Plain ASCII version for RStudio)

### 1 Setup and Notation

We assume uncalibrated stereo.

We have two unknown cameras (left and right).

A 3D point is written as:

$$X = (X, Y, Z, 1)$$

Its projection in the left image is:

$$x = (u, v, 1)$$

Its projection in the right image is:

$$x_{\text{prime}} = (u_{\text{prime}}, v_{\text{prime}}, 1)$$

The projection equations are:

$$x \propto PX$$

$$x_{\text{prime}} \propto P_{\text{prime}}X$$

P and P\_prime are unknown  $3 \times 4$  projection matrices.

We do NOT know the intrinsic matrix K, focal length f, or the baseline B.

Therefore the formula  $Z = fB/\text{disparity}$  cannot be used.

Instead, we use epipolar geometry.

### 2 Epipolar Geometry and the Fundamental Matrix F

For any corresponding pair  $x \leftrightarrow x_{\text{prime}}$ , the epipolar constraint must hold:

$$x_{\text{prime}}^T F x = 0$$

This equation defines the fundamental matrix F.

#### 2.1 Conceptual Explanation

The two camera centers and any 3D point define an epipolar plane.

This plane intersects each image plane in an epipolar line.

In the right image, the epipolar line is:

$$l_{\text{prime}} = Fx$$

The corresponding point  $x\_prime$  must lie on this line.  
Therefore the constraint:

$$x\_prime^T F x = 0$$

must hold for every correspondence.

## 2.2 Estimating F (8 Point Algorithm)

Given  $N \geq 8$  correspondences  $(u_i, v_i) \leftrightarrow (u\_prime_i, v\_prime_i)$ :

Each correspondence produces a linear equation in the entries of F.  
The equation has the form:

$$u\_prime_i u_i f_{11} + u\_prime_i v_i f_{12} + u\_prime_i f_{13} + v\_prime_i u_i f_{21} + v\_prime_i v_i f_{22} + v\_prime_i f_{23} + u_i f_{31} + v_i f_{32} + f_{33} = 0$$

Stacking all equations produces a system  $Af = 0$ .  
Solve using singular value decomposition.  
Reshape  $f$  into a  $3 \times 3$  matrix to obtain F.  
Force  $\text{rank}(F) = 2$  by zeroing the smallest singular value.

## 3 Constructing Uncalibrated Camera Matrices P and P\_prime

Once F is known, we can construct a canonical pair of projection matrices.

Step 1: Set the left camera to:

$$P = [I \mid 0]$$

Step 2: Compute the right epipole  $e\_prime$  such that:

$$F^T e\_prime = 0$$

Step 3: Form the skew symmetric matrix  $[e\_prime]_x$ .

Step 4: Define the right camera as:

$$P\_prime = ([e\_prime]_x F \mid e\_prime)$$

This pair of cameras is valid for uncalibrated stereo.

**Note:**

The reconstruction is defined only up to a  $4 \times 4$  projective matrix H.

Applying the transformation:

$$P \rightarrow PH, \quad P\_prime \rightarrow P\_prime H, \quad X \rightarrow H^{-1}X$$

does not change the image projections.

## 4 Triangulating 3D Points

For each correspondence  $x \leftrightarrow x\_prime$ :

$$\begin{aligned} x &\propto PX \\ x\_prime &\propto P\_prime X \end{aligned}$$

Use the cross product equations:

$$x \times (PX) = 0$$

$$x\_prime \times (P\_prime X) = 0$$

Together they produce 4 linear equations in the 4 components of X.  
This can be written as:

$$AX = 0$$

Solve using singular value decomposition.  
This gives the homogeneous 3D point X, up to projective scale.  
Repeat this for all correspondences to reconstruct the entire object.

## 5 Converting Projective 3D Into Real Distances

Uncalibrated stereo gives the correct shape but not the correct scale.

To measure real distances, we use a reference length.

Step 1: Choose two reconstructed 3D points  $X_a$  and  $X_b$  whose real distance is known.  
Call this real distance  $d_{real}$ .

Step 2: Compute the reconstructed distance:

$$d_{recon} = \sqrt{(X_a.x - X_b.x)^2 + (X_a.y - X_b.y)^2 + (X_a.z - X_b.z)^2}$$

Step 3: Compute the global scale:

$$s = \frac{d_{real}}{d_{recon}}$$

Step 4: For any pair  $X_i$  and  $X_j$  on the object:

$$\text{object\_distance} = s \sqrt{(X_i.x - X_j.x)^2 + (X_i.y - X_j.y)^2 + (X_i.z - X_j.z)^2}$$

This produces real-world measurements.

## 6 Final Procedure Summary

1. Collect corresponding points between the two images, including one known reference length.
2. Estimate the fundamental matrix F using the 8 point algorithm.
3. Construct canonical cameras P and P\_prime.
4. Triangulate 3D points.
5. Compute global scale using the known reference.
6. Measure object dimensions by computing scaled 3D distances.