# Handouts: Geometry

Ronen Basri

Introduction to Computer and Human Vision, Nov. 17, 2002

## 1 Projection

The process of forming a 2D image from the 3D scene is called projection. Below we define persective and orthographic projections.

### 1.1 Perspective

(Also called central projection, pinhole model.)

The pinhole is called *focal center* or *center of projection* and is denoted by $O = (0, 0, 0)$. The X- and Y-axes are assigned parallel to the image plane, and the Z-axis is set to be perpendicular to the image plane. (The Z-axis is called the *optical axis.*) For convenience, to avoid the inversion of pictures, the image plane is put in front of the center of projection.

The distance between $O$ and the image plane is called the *focal length* (denoted by $f$). Therefore image plane points have the form $(x, y, f)$. The projected position $p = (x, y)$ of a point $P = (X, Y, Z)$ is given by

$$x = \frac{fX}{Z} \qquad\qquad y = \frac{fY}{Z}.$$

(This can be derived easily using similar triangles).

One important aspect is that we see a larger part of the world as the viewing distance grows. Parallel lines appear intersecting in a single point, called the *vanishing point*, whose position is dependent on the orientation of the lines.

#### 1.1.1 Orthographic

Orthographic projection is a useful approximation to perspective projection. When an object is relatively far from the camera we can assume that all the rays are parallel to the optical axis.

$$(X, Y, Z) \Longrightarrow (x, y) \qquad \text{where} \quad x = X, \quad y = Y$$

To account for size differences due to distance we introduce a scale factor (scaled-orthographic or weak-perspective)

$$x = sX, \quad y = sY$$

where $s$ is constant for all points.

$$s = \frac{f}{Z_0}$$

This approximation holds when the objects considered are far relatively to their size.

Note that both projections are merely approximations. They do not account for distortions due to the lens, what it brings to focus and what it does not, abberations, etc..

## 2 Projective Geometry

In Euclidean geometry a point is denoted $(x, y)$ and a line is denoted $ax + by + c = 0$. Two lines can either meet at a point or never meet (if they are parallel). In projective geometry two lines always meet at a point. If the lines are parallel the points would meet at infinity (*ideal point*). Line equation is defined up to scale: $ax + by + c = 0$ and $(ka)x + (kb)y + (kc) = 0$ are the same line $k \neq 0$ (equivalence class). We can represent a line by a triplet $k(a, b, c)$. Lines are rays in $\mathbb{R}^3 - (0, 0, 0)$.

Homogeneous coordinates: We can write the coordinates of a point $(x, y, 1)$, then a line would be $ax + by + cz = 0$ and $(x, y, z)$ and $k(x, y, z)$ represent the same point. The Euclidean coordinates of the point are $(x/z, y/z)$. Points, as homogeneous vectors are elements in $\mathbb{P}^2$.

Question: where in $\mathbb{P}^2$ is $(0, 0)$?

Rewrite perspective projection in homogeneous coordinates $(x, y, f) = \frac{f}{Z}(X, Y, Z)$. In projective geometry $(x, y, f)$ and $(X, Y, Z)$ are the same point. A point lies on a line iff $l^T p = 0$. The intersection of two lines is $p = l \times l'$, since $l^T(l \times l') = l'^T(l \times l') = 0$.

Duality: $l^T p = 0$ implies $p^T l = 0$. Any result about points is also true for lines and vice versa.

The line through two points is $l = p \times p'$.

The intersection of two parallel lines $l = (a, b, c)$ and $l' = (a, b, c')$ is $(c - c')(b, -a, 0) \propto (b, -a, 0)$. This point is ideal; it does not exist in the Euclidean plane $(b/0, -a/0)$.

How many ideals points exist? 1-D family. We refer to this as the *line at infinity*.

Conics: $ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = 0$, or $p^T C p = 0$ with

$$C = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}.$$

$C$ is symmetric and the quadratic form is homogeneous, so $C$ is homogeneous (defined up to a scale factor). Define a unit circle. A conic that is not full rank is called *degenerate*.

Five points define a conic. $p_i^T C p_i = 0$ is linear in $C$ and homogeneous. $C$ has 6 variables determined up to a scale factor, $(x_i^2, x_i y_i, y_i^2, x_i z_i, y_i z_i, z_i^2)c = 0$.

Tangent to a conic $l^T p = p^T C p = 0$, so $l = Cp$.

Dual conic $p = C^{-1}l$ thus $(C^{-1}l)C(C^{-1}l) = l^T C^{-1} l = 0$. The lines $l$ are tangent to the conic (envelop).

Projective transformation (homography, collineation) $h : \mathbb{P}^2 \to \mathbb{P}^2$ is non-invertible transformation such that 3 points on a line $p_1$, $p_2$, $p_3$ iff $h(p_1)$, $h(p_2)$, $h(p_3)$ on a line.

Projective transformation is represented by a $3 \times 3$ non-singular matrix $H$, $q \propto Hp$ (that is, $\alpha q = Hp$). Four points determine homography.

Projective transformation is a product of perspectivities.

$p' \propto Hp$ implies $l' \propto H^{-T} l$, since $l'^T p' \propto l^T H^{-1} H p = l^T p = 0$. Likewise $C' = H^{-T} C H$.

Hierarchy of transformations:

1. Isometries (Euclidean), rotation $R^T R = I$, $det(R) = 1$ and translation, reflection.

   $$\begin{pmatrix} \pm\cos\theta & -\sin\theta & t_x \\ \pm\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix}.$$

   Isometries preserve angles, distances and areas (invariants).

2. Similarity

   $$\begin{pmatrix} sR & t \\ 0^T & 1 \end{pmatrix}.$$

   Preserves angles, distance ratio, area ratio.

3. Affine

   $$\begin{pmatrix} A & t \\ 0^T & 1 \end{pmatrix}.$$

   Preserves parallel lines, distance ration along parallel lines, ratio of areas.

4. Projective

   $$\begin{pmatrix} A & t \\ u^T & v \end{pmatrix}.$$

   Preserves cross ratio

   $$\text{Cross}(a, b, c, d) = \frac{\overline{ab}\,\overline{cd}}{\overline{ac}\,\overline{bd}},$$

   where $\overline{ab}$ is the signed distance between $a$ and $b$ for some choice of orientation.

Example: square grid after projection. Infer line at infinity (from intersections of parallel lines. Then rectify.

# 3 Two views related by homography

1. Camera rotation: $P' = RP$ implies $p' \propto Rp$.

2. Planar scene: $P = R(u, v, 0)^T + t'$, $P' = R'(u, v, 0)^T + t'$. Let $\hat{R} = [r_1, r_2, t]$ and $\hat{R}' = [r'_1, r'_2, t']$. Then $P = \hat{R}(u, v, 1)^T$ and $P' = \hat{R}'(u, v, 1)^T$. Therefore, $P = \hat{R}\hat{R}'^{-1}P' = HP'$.

# 4 Algorithms

## 4.1 RANSAC

Given $m$ and $n$ feature points, choose $k$ from each image and treat them as matches. Recover transformation based on these $k$-touples. Verify the transformation. If correct, you are done. If not, try another $k$-touple. Complexity $m^k n^k \times$ verification.

## 4.2 Hough transform

Prepare a quantized table for possible transformations. Given $m$ and $n$ feature points, choose $k$ from each image and treat them as matches. Recover transformation based on these $k$-touples. Vote to the corresponding bin in the table. After trying all $k$-touples select the bin with highest vote.