

55:148 Digital Image Processing

Chapter 11

3D Vision, Geometry

Topics:

- Basics of projective geometry

 - Points and hyperplanes in projective space

 - Homography

 - Estimating homography from point correspondence

- The single perspective camera

 - An overview of single camera calibration

 - Calibration of one camera from the known scene

- Two cameras, stereopsis

 - The geometry of two cameras. The fundamental matrix

 - Relative motion of the camera; the essential matrix

 - Estimation of a fundamental matrix from image point correspondences

 - Applications of the epipolar geometry in vision

- Three and more cameras

 - Stereo correspondence algorithms

3D VISION INTRO

Overall Aim:

Marr 1982: From an image (or a series of images) of a scene, derive an accurate three-dimensional geometric description of the scene and quantitatively determine the properties in the scene.

Examples: <http://www.2d3.com/>

Major challenges

Perspective projection : All points along a line radiating from the optical center are projected onto a single image point → loss of information → need for additional information to solve the inverse task, i.e., mapping each 2D image point into the 3D scene.

3D geometry and intensity shading: A complex relation governed multiple variables.

Mutual occlusion: Further complicates the vision task at conceptual level.

Noise: Additional complexity to many algorithms reducing their sensitivity and accuracy.

3D VISION INTRO

Three major intertwined modules in a computer-based vision system

Feature observability in images: Selection of task-relevant features in the original image data (e.g., points, lines, corners etc.)

Representation: Choice of the models for the observed world (e.g., a triangulated 3D surface representation of the observed scene)

Interpretation: Extraction of high level knowledge from the mathematical model of stored data (e.g., object detection/recognition, correspondence between two partially overlapping scenes etc.)

Three major building blocks/expertise in a computerized 3D vision system

Computational theory: Combines analytic and geometric approaches with device-dependent properties to solve the inverse mapping from a 2D captured image to the 3D scene

primal sketch → 2.5D → full 3D representation

Representation and algorithms: 3D scene data representation and algorithms manipulating them and extracting knowledge (high level information) from 3D scene representation

Implementation: Physical realization of the algorithms (programs + hardware)

Different vision paradigms

Active versus passive vision

Qualitative versus purposive vision

Basics of projective geometry

Single or multiple view geometry deals with mathematics of relation between

- 3D geometric features (points, lines, corners) in the scene
- their camera projections
- relations among multiple camera projections of a 3D scene

Points and hyperplanes in projective space

Scene: $(d + 1)$ -dimensional space excluding the origin, i.e., $\mathbb{R}^{d+1} - \hat{0}$

Why origin is excluded?

Origin \approx pinhole \approx optical center

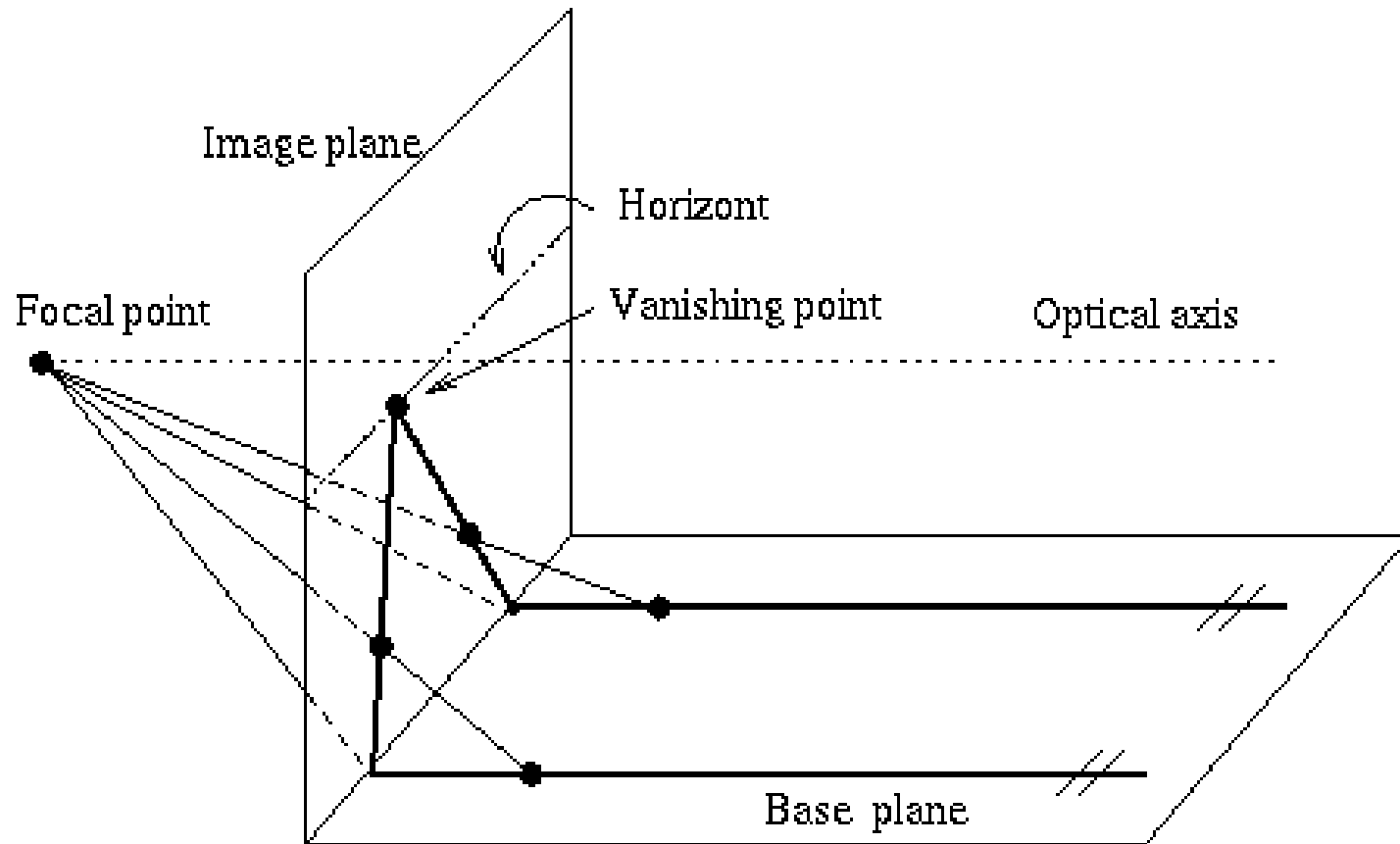
Projective scape: a hyperplane \mathcal{P}^d in the $(d + 1)$ -dimensional scene NOT passing through the origin

An equivalence relation “ \cong ” is defined as follows:

$$[x_1, \dots, x_{d+1}]^T \cong [x'_1, \dots, x'_{d+1}]^T$$

$$\text{iff } \exists \alpha \neq 0 \text{ s. t. } [x_1, \dots, x_{d+1}]^T = \alpha [x'_1, \dots, x'_{d+1}]^T$$

Perspective projection of parallel lines



Homogeneous points

Each equivalent class of the relation " \cong " generates an *open* line from the origin.

Note that the origin is not included in any of these lines and thus the disjoint property of equivalent classes is satisfied

For each line or equivalent class, exactly one point is projected in the acquired image and is the point where the projective hyperplane intersects the line. These points in the projective space are referred to a homogeneous points.

What is the property of homogenous points?

Homogeneous points are coplanar lying on the projection plane.

For simplicity, let us assume that our projection plane is $z = 1$

Homogeneous points

Note that homogeneous points form the image hyperplane.

Thus, to determine the perspective projection of a scene point, we need to determine corresponding homogeneous point

$$[x_1, \dots, x_{d+1}]^T \xrightarrow{P} [x'_1, \dots, x'_{d+1} = 1]^T,$$

where $x_i = \alpha x'_i \mid \alpha$: constant.

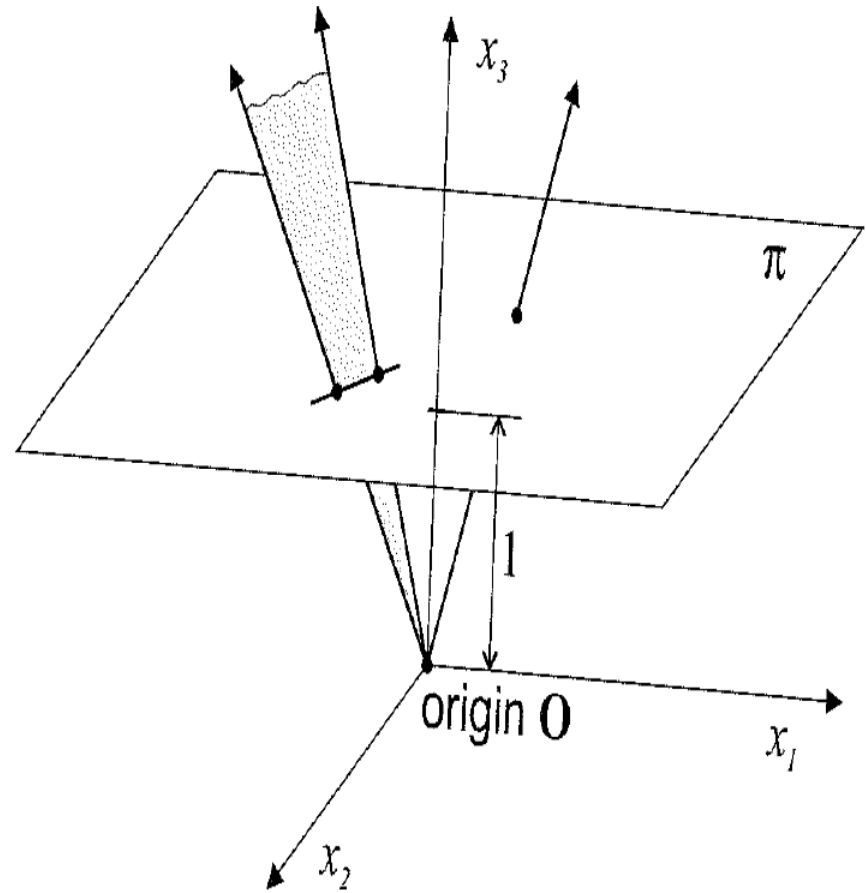
Note that the points $[x_1, \dots, x_d, 0]^T$ do not have an Euclidean counterpart

- Consider the limiting case $[x_1, \dots, x_d, \alpha]^T$ that is projectively equivalent to $[x_1/\alpha, \dots, x_d/\alpha, 1]^T$, and assume that $\alpha \rightarrow 0$.
- This corresponds to a point on the projective hyperplane \mathcal{P}^d going to infinity in the direction of the radius vector $[x_1, \dots, x_d, 0]^T$

Properties of projection

A line in the scene space through (but not including) the origin is mapped onto a point in the projective plane

A plane in the scene space through the origin (but not including) is mapped to a line on the projection plane



Homography

Homography \approx Collineation \approx Projective transformation

is a mapping from one projection plane to another projection plane for the same $(d + 1)$ -dimensional scene and the common origin

$$\mathcal{P}^d \xrightarrow{H} \hat{\mathcal{P}}^d.$$

Also, expressed as

$$\mathbf{u}' \cong H\mathbf{u},$$

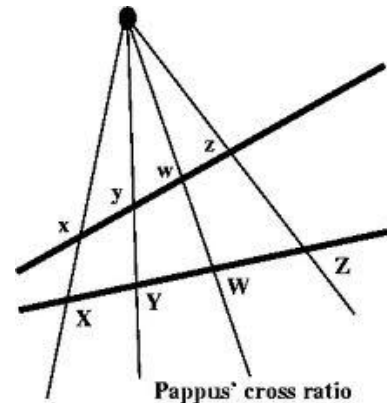
where H is a $(d + 1) \times (d + 1)$ matrix.

Property:

Any three collinear points in \mathcal{P}^d remain collinear in $\hat{\mathcal{P}}^d$

Prove!

Satisfies cross ratio property (see the figure)



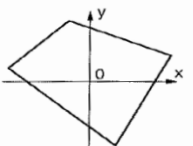
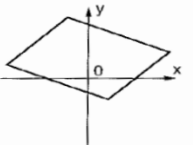
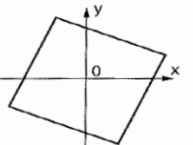
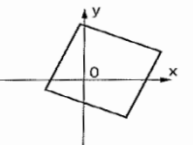
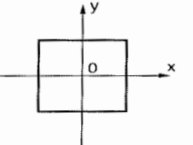
Matrix formulation for Homography

$$\alpha \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

The scale factor $\alpha \neq 0$ and $\det H \neq 0$; otherwise everything is mapped onto a single point.
Eliminating the scale factor α , we get

$$u' = \frac{h_{11}u + h_{12}v + h_{13}}{h_{31}u + h_{32}v + h_{33}} \quad \text{and} \quad v' = \frac{h_{21}u + h_{22}v + h_{23}}{h_{31}u + h_{32}v + h_{33}}$$

Various linear transformations

Name	Constraints on H	2D example	Invariants
projective	$\det H \neq 0$		collinearity tangency cross ratio
affine	$H = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}$ $\det A \neq 0$		projective invariants + parallelism + length ratio on parallels + area ration + linear combinations of vectors centroid
similarity	$H = \begin{bmatrix} sR & -R\mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}$ $R^\top R = I$ $\det R = 1$ $s > 0$		affine invariants + angles + ratio of lengths
metric (Euclidean, isometric)	$H = \begin{bmatrix} R & -R\mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}$ $R^\top R = I$ $\det R = 1$		similarity invariants + length + area (volume)
identity	$H = I$		trivial case everything is invariant

Sub groups of homographys

Any homography can be uniquely decomposed as

$$H = H_P H_A H_S$$

where

$$H_P = \begin{bmatrix} I & \hat{0} \\ \mathbf{a}^T & b \end{bmatrix}, \quad H_A = \begin{bmatrix} K & \hat{0} \\ \hat{0}^T & 1 \end{bmatrix}, \quad H_S = \begin{bmatrix} R & -R\mathbf{t} \\ \hat{0}^T & 1 \end{bmatrix}$$

Estimating homography from point correspondence

Given a set of ordered pairs of points $\{(u_i, u'_i)\}_{i=1}^m$

To solve the homogeneous system of linear equations

$$\alpha_i u'_i = H u_i, \quad i = 1, \dots, m$$

for H and α_i .

Number of equations : $m(d + 1)$

Number of unknowns: $m + (d + 1)^2 - 1$

Degenerative configuration, i.e., H may not be uniquely solved even if $m \geq d + 2$ and caused when d or more points are coplanar

Correspondence of more than sufficient points lead to the notion of optimal fitting reducing the effect of noise

Maximum likelihood estimation

$[\tilde{u}_i, \tilde{v}_i]^T$ and $[\tilde{u}'_i, \tilde{v}'_i]^T \mid i = 1, \dots, m$ are identified corresponding points in two different projection planes

Principle: Find the homography (i.e., the transformation matrix H) that maximizes the likelihood mapping of the points $[\tilde{u}_i, \tilde{v}_i]^T$ on the first plane to $[\tilde{u}'_i, \tilde{v}'_i]^T$ on to the second plane

Model:

Ideal points are in the vicinity of the identified points, i.e., there noise in the process of locating the points $[\tilde{u}_i, \tilde{v}_i]^T$ and $[\tilde{u}'_i, \tilde{v}'_i]^T$

Method to solve the problem

- Determine the ML function using Gaussian model
- It contains several multiplicative terms
- Take log \rightarrow multiplications are converted to addition
- Remove the minus sign (see the Gaussian expression)
- Maximization is converted to a minimization term

Final expression for maximum likelihood estimation

$$\min_{h, u_i, v_i} \sum_{i=1}^m \left[(u_i - \tilde{u}_i)^2 + (v_i - \tilde{v}_i)^2 + \left(\frac{h_{11}u_i + h_{12}v_i + h_{13}}{h_{31}u_i + h_{32}v_i + h_{33}} - \tilde{u}'_i \right)^2 + \left(\frac{h_{21}u_i + h_{22}v_i + h_{23}}{h_{31}u_i + h_{32}v_i + h_{33}} - \tilde{v}'_i \right)^2 \right]$$