Multiple View Geometry

Nie Shicong

November 24, 2019

Contents

1 Introduction - a Tour of Multiple View Geometry

1.1 Introduction - the ubiquitous projective geometry

Preamble:

In projective transformation, straightness is preserved.

In Euclidean geometry, an exception: a pair of lines can be parallel.

Points at infinity are called *ideal points*. By adding ideal points, Euclidean space is transfored into the projective space.

Coordinates:

In projective space, use the honogeneous coordinates. Honogeneous coordinates: (kx, ky, k) represents the same point, for any non-zero value k.

When k is zero, the point becomes ideal point.

In two-dimensional projective space, the set of ideal points forms a line, called *line at infinity*. In three-dimensions, they form the *plane at infinity*.

Honogeneity:

Euclidean geometry treats all points as the same.

Euclidean transformation moves the origin of the space.

Affine transformation: a linear transformation followed by a Euclidean transformation. It moves, rotates and finally stretches linearly the space.

Projective geometry also treats all points as the same.

Projective transformation: $X' = H_{(n+1)\times(n+1)}X$.

In projective transformation, infinity is not preserved. In Euclidean and affine transformation, infinity is preserved.

1.1.1 Affine and Euclidean Geometry

Affine Geometry:

The geometry of the projective plane and a distinguished line is known as affine geometry.

Any projective transformation that maps the distinguished line in one space to the distinguished line of the other space is known as an affine transformation.

Euclidean Geometry:

Affine geometry is seen as spectialization of projective geometry.
Absolute conic.
Circular points.

3D Euclidean geometry:
Perpendicularity of lines.
Two directions are conjugate points w.r.t. the absolute conic.

1.2   Camera projections

Preamble:
Central projection: a ray, center of projection and image plane.
A simple model of a camera: a ray of light, lens and digitial device.
The most general projection is represented by an arbitrary 3×4 matrix of rank 3, acting on the homogeneous coordinates of the point in $\mathbb{P}^3$ mapping it to the imaged point in $\mathbb{P}^2$. This matrix is known as the *camera matrix*.

Cameras as points:
The camera center is the essence: for now, all images acquired with the same camera center are equivalent - they can be mapped onto each other by a projective transformation.

Calibrated cameras:
IAC: Image of the Absolute Conic, denoted ω. If the location of the IAC is known in an image, then we say that the camera is **calibrated**.
Knowing the IAC, one can measure the angle between rays by direct measurements in the image.
Euclidean structure.

Example 1.1. 3D reconstruction from paintings:
Single view reconstruction: using techniques of projective geometry, it's possible in many instances to reconstruct scenes from a single image.
Affine structure.

1.3 Reconstruction from more than one view

Point correspondence: $x_i \leftrightarrow x_i'$.
Reconstruction from two views has unavoidable projective ambiguity.
Reconstruction is possible when there is at least seven point correspondence, and they must not lie in one of various well-defined *critical configuration*.
Basic tool for reconstruction of point sets from two views: *fundamental matrix F*. It encapsulates the complete projective geometry of the pair of cameras.
Given F, a pair of matching points must satisfy the *coplanarity equation*:

$$x_i' F x_i = 0$$

Triangulation.

1.4 Three-view geometry

Basic algebraic entity in three views reconstruction: *trifocal tensor $\mathcal{T}_i^{jk}$*.
Correspondence in three views: $x \leftrightarrow l' \leftrightarrow l''$. It's not only correspondence between points but between points and lines.
A total of seven point correspondences are sufficient to compute the trifocal tensor linearly. And a minimum of six point correspondences using a non-linear method are sufficient.
Advantages to using three-view reconstruction: mixture of line and point correspondences; greater stability.

1.5 Four view geometry and n-view reconstruction

Quadrifocal tensor is seldom used.
Affine camera.
factorization algorithm: using the *Singular Value Decomposition*. This algorithm requires that all the points be visible in all views.
The dominant methodology for the general reconstruction problem is *bundle adjustment*, or *BA*.
BA is an iterative method, which can not be guaranteed to converge to the optimal solution from a arbitray starting point. BA is quite efficient when implemented carefully.

1.6 Transfer

Transfer: given the position of a point in one (or more) image(s), determine where it will appear in all other images of the set.

1.7 Euclidean reconstruction

So far, we have talked about the reconstruction with a set of uncalibrated cameras. If a complete calibration of each of the camera is known then it is possible to remove some of the ambiguity of the reconstructed scene.
Distortions.
The knowledge of the camera calibration is equivalent to being able to determine the Euclidean structure of the scene.

1.8 Auto-calibration

Preamble:
Generally given three cameras known to have the same calibration, it is possible to determine the absolute conic, and hence the calibration of the cameras. However, it remains quite a difficult problem.

Knowing the plane at infinity:
Auto-calibration is relatively simple, once the plane at infinity has been identified. The identification of the plane at infinity itself is substantially more difficult.

Auto-calibration give square pixels in the image:
square-pixel constraint: a Euclidean coordinate system is known in each image.
From four images one may determine the calibration.

1.9 The reward I: 3D graphical models

Epipolar geometry: a point in one image defines a line in the other image on which the corresponding point must lie.

Dense 3D model, dense point match.

1.10 The reward II: video augmentation

An application of reconstruction: adding artificial graphics objects in real video sequences.
The only requirement is to compute the correct motion of the camera and hence we are able to generate correct perspective views of the graphics model. It's essential to compute the motion of the camera in a Euclidean frame, because the graphics object and the cameras need to be known in the same coordinate frame.

2 Projective Geometry and Transformations of 2D

2.1 Planar geometry

Geometric approach: coordinate-free.
Algebraic approach: using coordinates and algebraic entities.
The approach in this book is a hybrid approach.
In the algebraic approach, a point or a line is identified as a vector, whilst a conic is represented by a symmetic matrix.
A significant advantage of the algebraic approach is that the results derived in this way may more easily be used to derive algorithms and practical computational methods.

2.2 The 2D projective plane

Row and column vectors. Geometric entities will by default be represented by column vectors.

2.2.1   Points and lines

Homogeneous representation of lines.
A line is represented by the vector $(a, b, c)^\top$.
The vectors $(a, b, c)^\top$ and $k(a, b, c)^\top$ represent the same line for any non-zero k.
In fact, two such vectors related by an overall scaling are considered as being equivalent. An equivalent class of vectors under this equivalence relationship is known as a *homogeneous* vector.
The set of equivalence classes of vectors in $\mathbb{R}^3 - (0, 0, 0)^\top$ forms the *projective space* $\mathbb{P}^2$. The notation $-$ indicates the operand followed is excluded.

Homogeneous representation of points.
$\mathbf{x} = (x_1, x_2, x_3)^\top$ represents the point $(x_1/x_3, x_2/x_3)^\top$ in $\mathbb{R}^2$.
Result 2.1. *the point x lies on the line l iff $x^\top l = 0$.*

Degrees of freedom, dof.
A line is specified by two parameters and so has two degrees of freedom.
These two params could be chosen as the gradient and y intercept of the line.

Intersection of lines.
Result 2.2. *The intersection of two lines l and l' is the point $x = l \times l'$.*
Example 2.3. Determine the intersection of the lines $x = 1$ and $y = 1$.

Line joining points.
Result 2.4. *The line through two points x and x' is $l = x \times x'$.*

2.2.2   Ideal points and the line at infinity

Intersection of parallel lines.
Example 2.5. Determine the intersection point of two lines $x = 1$ and $x = 2$.

Ideal points and the line at infinity.
One may augment $\mathbb{R}^2$ by adding points with last coord $x_3 = 0$. The resulting space is the set of all homogeneous 3-vectors, namely the *projective space* $\mathbb{P}^2$.
Ideal points.
Line at infinity.
First two coords of the ideal point represent the *direction*. Line at infinity

6

can be thought of as the set of directions of lines in the plane.

Projective geometry.
There is no distinction between points at infinity (ideal points) and ordinary points. At times, however, we need to distinguish between them.

A model for the projective plane.
Points and lines of $\mathbb{P}^2$ are represented by rays and planes, respectively, through the origin in $\mathbb{R}^3$.

Duality.
Result 2.6. Duality principle. *To any theorem of 2-dim projective geometry there corresponds a dual theorem, which may be derived by interchaning the roles of points and lines in the original theorem.*

2.2.3   Conics and dual conics.

Hyperbola, ellipse, and parabola.
Conic coefficient matrix C. It's a homogeneous representation of a conic.
A conic has five dofs.

Five points define a conic.
Each point x_i places one constraint on the conic coefficients. Therefore a conic is determined uniquely (up to scale) by five points in general positions.

Tangent lines to conics.
Result 2.7. *The line l tangent to C at a point x on C is given by $l = Cx$.*

Dual conics.
Point conic and *dual* or *line conic.*
It's denoted as C^*. The notation $*$ indicates that C^* is the *adjoint matrix* of C.
In a similar manner, five lines in general position define a dual conic.
Point conic equ: $x^\top C x = 0$, dual (line) conic equ: $l^\top C^* l = 0$.

Degenerate conics.
If the matrix C is not of full rank, then the conic is termed degenerate.
Degenerate point conic and *degenerate line conic.*
Example 2.8. The degenerate point conic: $C = lm^\top + ml^\top$.

2.3 Projective transformations

Geometry is the study of properties invariant under groups of transformations.

2D projective geometry is under a group of transformations known as *projectivities*. A projectivity is also called a *collineation*, a *projective transformation* or a *homography*.

Theorem 2.10. A projectivity is an invertible mapping $h : \mathbb{P}^2 \to \mathbb{P}^2$ represented by a invertible 3×3 matrix H, which is $h(x) = Hx$.
Any invertible linear transformation of homogeneous coords is a projectivity.

Definition 2.11. Projective transformation.
H is a homogeneous matrix. It has eight dofs.

Mappings between planes.
Central projection. It preserves lines.
Perspectivity.

Example 2.12. Removing the projective distortion from a perspective image of a plane.
Shape is distored under perspective imaging. In order to remove the distortion, we need to compute the inverse of the projective transformation.
The method is using four point correspondences wherein each point is represented in inhomogeneous form. When the projective transformation is determined, its inverse can be easily computed. Note that the four points should be in "general position", which means that no three points are colinear.
This way does not require knowledge of **any** of the camera's params or the pose of the plane.
It's not the preferred way.

2.3.1 Transformations of lines and conics.

Transformation of lines.
Under a point transform $x' = Hx$, a line transforms as $l' = H^{-\top}l$, or alternatively write $l'^{\top} = l'H^{-1}$.
One says that points tranform *contravariantly* and lines transform *covariantly*.

Transformation of conics.

Result 2.13. *Under a point transformation $x' = Hx$, a conic transforms to $C' = H^{-\top}CH^{-1}$.*

Result 2.14. *Under a point transformation $x' = Hx$, a dual conic C^* transforms to $C^{*\prime} = HC^*H^\top$.*

2.4   A hierarchy of transformations

Preamble.
Projective linear group, denoted as $PL(n)$.
General linear group, denoted as $GL(n)$.
Affine group is an important *subgroup* of $PL(3)$, and the *Euclidean group* is a subgroup of the affine group.
Isometries.
Not all classes of transformation form a group, e.g. the perspectivities.

Invariants.
An alternative to describing the transformation *algebraically*, i.e. as a matrix acting on coords of a point or curve, is to describe the transformation in terms of those elements or quantities that are preserved or *invariant*.

2.4.1   Class I: Isometries

Isometries are transformations of the plane $\mathbb{R}^2$ that preserve Euclidean distance.
There is *orientation-preserving* isometries and *orientation-reversing* isometries.
Euclidean transformation, composed of a *rotation matrix*, a *translation matrix*, a null vector and a pure number one. A planar Euclidean transformation has three dofs, one for rotation and two for translation.

Invariants.
Length, angle, and area.

Groups and orientation.
Orientation-preserving isometries form a group, whilst orientation-reversing

ones do not.

2.4.2   Class II: Similarity transformations

A similarity tranformation (or more simply a similarity) is an isometry composed with an isotropic scaling s.
It's also known as an *equi-form* transformation, because it preserves shape.
A planer similarity has four dof as the addition of scaling s.

Invariants.
Angles, ratio of two lengths and ratio of areas.

Metric structure.
The description *metric structure* implies that the structure is defined up to a similarity.

2.4.3   Class III: Affine transformations

An affine transformation (or more simply an *affinity*) is a non-singular linear transformation followed by a translation.
It can be seen as the composition of two fundamental transformations, namely rotations and non-isotropic scalings. It can always be decomposed as

$$A = R(\theta)R(-\phi)DR(\phi)$$

This decomposition follows directly from the *SVD*.
The new non-isotropic scaling accounts for the two extra dofs: the angle ϕ specifying the scaling direction, and the ratio of the scaling params $\lambda_1 : \lambda_2$.
The essence of an affinity is this scaling in **orthogonal** directions, oriented at a particular angle.

Invariants.
Parallel lines, ratio of lengths of parallel line segments, and ratio of areas.

2.4.4   Class IV: Projective transformations

A projective transformation is a general non-singular linear transformation of **homogeneous** coords. It's a generalization of affine transformation.

A projective transformation in $\mathbb{P}^2$ has eight dofs.

Invariants.
The *cross ratio* (ratio of ratios) of four collinear points.

2.4.5 Summary and comparison

Angles are not preserved, so that shapes are skewed.
Area scaling varies with position, and the orientation of a transformed line depends on both the orientation and position of the source line.
The key difference between a projective and affine transformation is that the vector v is not null for a projectivity. This is responsible for the non-linear effects of the projectivity.
Affinity maps the ideal points to finite points whilst the ideal points over projectivity remain ideal points.

2.4.6 Decomposition of a projective transformation

A projective transformation can be decomposed into a chain of transformations, where each matrix in the chain represents a transformation higher in the hierarchy than the previous one:

$$H = H_S H_A H_P$$

or in the reverse order:

$$H = H_P H_A H_S$$

Note that the actual values of K, R, t and v will be different.
H_P is an *elation*.

Example 2.15. $H = \begin{bmatrix} 1.707 & 0.586 & 1.0 \\ 2.707 & 8.242 & 2.0 \\ 1.0 & 2.0 & 1.0 \end{bmatrix}$ Decompose this projectivity.

2.4.7 The number of invariants

Number of invariants. It's functionally independent.

Result 2.16. *the number of functionally independent invariants is equal to, or greater than, the number of degrees of freedom of the configuration less the number of degrees of freedom of the transformation*

2.5 The projective geometry of 1D