Fundamentals of Photogrammetry

Niclas Börlin, Ph.D.
niclas.borlin@cs.umu.se

Department of Computing Science
Umeå University
Sweden

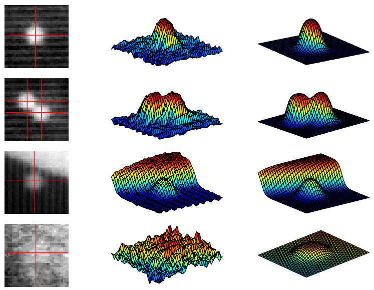
April 1, 2014

Presentation

- Ph.D. in Computing Science (2000).
- Numerical Linear Algebra.
- Non-linear least squares with non-linear equality constraint.
- X-ray photogrammetry — Radiostereometry (RSA).
- Post doc at Harvard Medical School, Boston, MA.

Radiostereometric analysis (RSA)

- Developed by Hallert (1960), Selvik (1974), Kärrholm (1989), Börlin (2002, 2006), Valstar (2005).
- Procedure
  - Dual X-ray setup
  - Calibration cage
  - Marker measurements
  - Reconstruction of projection geometry
  - Motion analysis



- Software *UmRSA Digital Measure* running in Europe, North America, Australia, Asia. Used to produce 150+ scientific papers.

Definition

- *Photogrammetry* — measuring from photographs
  - *photos* — "light"
  - *gramma* — "that which is drawn or written"
  - *metron* — "to measure"
- Definition in *Manual of Photogrammetry*, 1st ed., 1944, American Society for Photogrammetry:

  > *Photogrammetry is the science or art of obtaining reliable measurement by means of photographs*

Overview

- Principles
- History
- Mathematical models
- Processing
- Applications

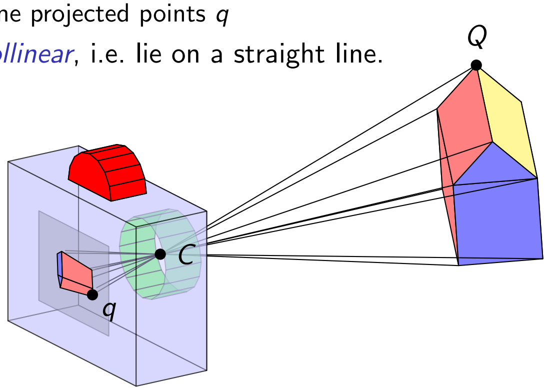
Principles

- Non-contact measurements.
- (Passive sensor.)
- Collinearity.
- Triangulation.

Collinearity

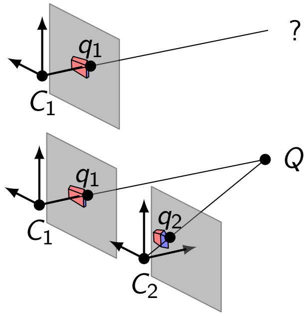
- The collinearity principle is the assumption that
  - the object points Q,
  - the projection center C, and
  - the projected points q

  are *collinear*, i.e. lie on a straight line.



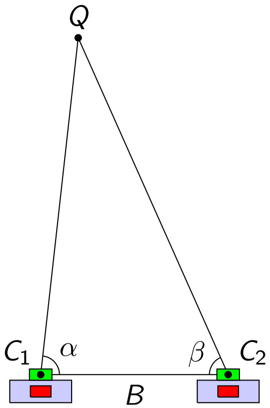
Triangulation

- One image coordinate measurement (x, y) is too little to determine the object point coordinates (X, Y, Z).
- We need at least two measurements of the same point.



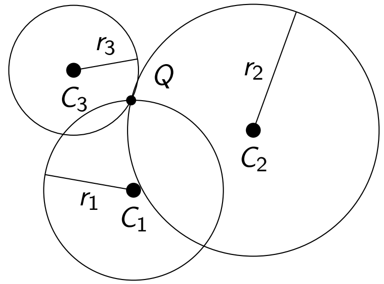
Triangulation (2)

- The position of object points are calculated by *triangulation*, i.e. by *angles*, but without any *range* values.

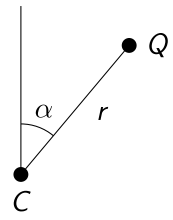


Other techniques

- Trilateration, ranges but no angles (GPS).



- Tachyometry, angles and ranges (surveying, laser scanning)



Overview

- Principles
- History
- Mathematical models
- Processing
- Applications

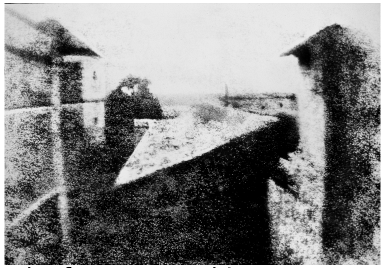
Pre-history

- Geometry, perspective, pinhole camera model — Euclid (300 BC).
- Leonardo da Vinci (1480)

> *Perspective is nothing else than the seeing of an object behind a sheet of glass, smooth and quite transparent, on the surface of which all the things may be marked that are behind this glass. All things transmit their images to the eye by pyramidal lines, and these pyramids are cut by the said glass. The nearer to the eye these are intersected, the smaller the image of their cause will appear.*

First generation — Plane table photogrammetry

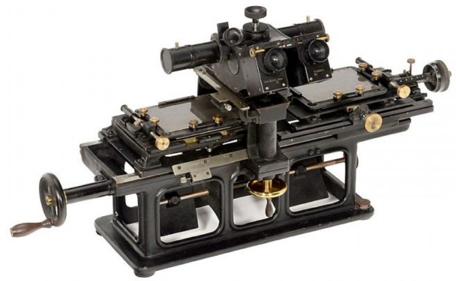
- First photograph — Niépce, 1825. Required 8 hour exposure.
- Glass negative — Hershel, 1839.



- First use of terrestrial photographs for topographic maps — Laussedat, 1849 "Father of photogrammetry". City map of Paris (1851).
- Film — Eastman, 1884.
- Architectural photogrammetry — Meydenbauer, 1893, coined the word "photogrammetry".
- Measurements made on a map on a table. Photographs used to extract angles.

Second generation — Analog photogrammetry

- Stereocomparator (Pulfrich, Fourcade, 1901). Required coplanar photographs. Measurements made by floating mark.



- Aeroplane (Wright 1903). First aerial imagery from aeroplane in 1909.
- Aerial survey camera for overlapping vertical photos (Messter

Second generation — Analog photogrammetry (2)

- Opto-mechanical stereoplotters (von Orel, Thompson, 1908, Zeiss 1921, Wild 1926). Allowed non-coplanar photographs.

Wild A8 Autograph (1950)

- Relative orientation determined by 6 points in overlapping images — von Gruber points (1924)

> *Photogrammetry — the art of avoiding computations*

Third generation — analytical photogrammetry

- Finsterwalder (1899) — equations for analytical photogrammetry, intersection of rays, relative and absolute orientation, least squares theory.
- von Gruber (1924) — projective equations and their differentials,
- Computer (Zuse 1941, Turing, Flowers, 1943, Aiken 1944).
- Schmid, Brown multi-station analytical photogrammetry, bundle block adjustment (1953), adjustment theory.

> *The [Ballistic Research] laboratory had a virtual global monopoly on electronic computing power. This unique circumstance combined with Schmid set the stage for the rapid transistion from classical photogrammetry to the analytic approach (Brown).*

- Ackermann independent models (1966).

Third generation — analytical photogrammetry (2)

- Analytical plotter (Helava 1957) - image-map coordinate transformation by electronic computation, servocontrol.

Zeiss Planicomp P3

- Camera calibration (Brown 1966, 1971).
- Direct Linear Transform (DLT) (Abdel-Azis, Karara, 1971).

Digital photogrammetry

- Charge-Coupled Device (CCD) (Boyle, Smith 1969).
- Landsat (1972)
- Digital camera (Sesson (Eastman Kodak) 1975 — 0.01 Mpixels).
- Flash memory (Masuoka (Toshiba) 1980).
- Matching (Förstner 1986, Gruen 1985, Lowe 1999).
- Projective Geometry (Klein 1939)
- 5-point relative orientation (Nistér 2004)

Overview

- Principles
- History
- Mathematical models
- Processing
- Applications

Matrix multiplication

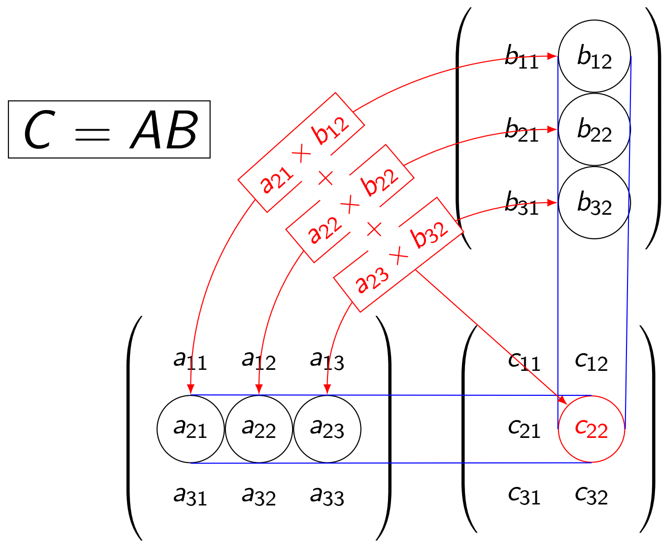
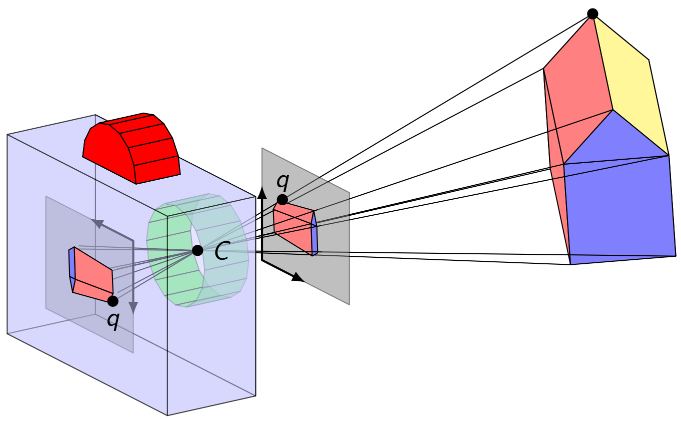


Image plane placement

- The projected coordinates q will be identical
  - if a (negative) sensor is placed *behind* the camera center or
  - if a (positive) sensor is mirrored and placed *in front of* the camera center.

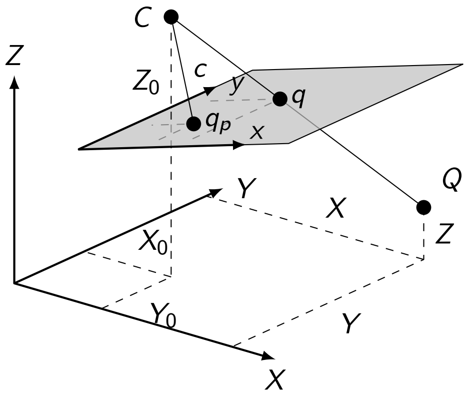


The collinearity equations

- The *collinearity equations*

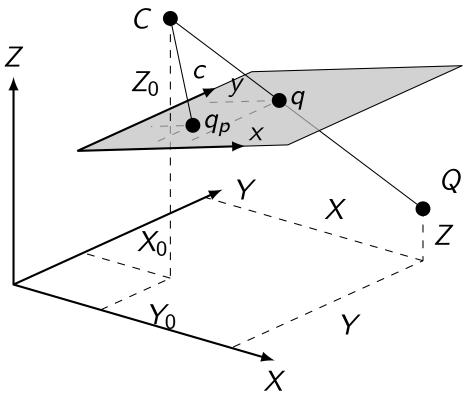
$$\begin{pmatrix} x - x_p \\ y - y_p \\ -c \end{pmatrix} = kR \begin{pmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{pmatrix}$$

describe the relationship between the object point $(X, Y, Z)^T$, the position $C = (X_0, Y_0, Z_0)^T$ of the camera center and the orientation R of the camera.



The collinearity equations (2)

- The distance c is known as the *principal distance* or *camera constant*.
- The point $q_p = (x_p, y_p)^T$ is called the *principal point*.
- The ray passing through the camera center C and the principal point q_p is called the *principal ray*.



The collinearity equations (3)

- From

$$\begin{pmatrix} x - x_p \\ y - y_p \\ -c \end{pmatrix} = kR \begin{pmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{pmatrix}, \text{ and } R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix},$$

we can solve for k and insert:

$$x = x_p - c\frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)},$$

$$y = y_p - c\frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}.$$

Homogenous coordinates

- In projective geometry, points, lines, etc. are represented by *homogenous coordinates*.
- Any cartesian coordinates (x, y) may be transformed to homogenous by adding a unit value as an extra coordinate:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

- All homogenous vector multiplied by a non-zero scalar k belong to the same *equivalence class* and correspond to the same object. Thus,

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \text{ and } k \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} kx \\ ky \\ k \end{pmatrix}, k \neq 0$$

all correspond to the same 2D point $(x, y)^T$.

Homogenous coordinates (2)

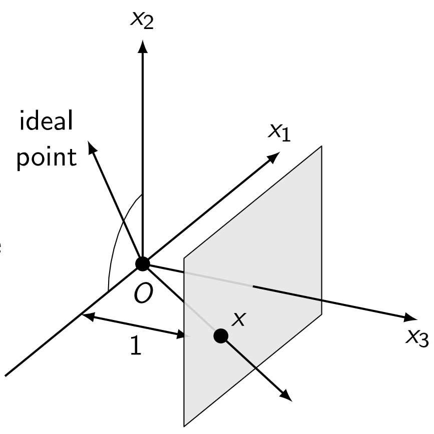
- Any homogenous vector $(x_1, x_2, x_3)^T$, $x_3 \neq 0$ may be transformed to cartesian coordinates by dividing by the last element

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1/x_3 \\ x_2/x_3 \\ x_3/x_3 \end{pmatrix} = \begin{pmatrix} x_1/x_3 \\ x_2/x_3 \\ 1 \end{pmatrix}.$$

- A homogenous vector $(x_1, x_2, x_3)^T$ with $x_3 = 0$ is called an *ideal point* and is "infinitely far away" in the direction of (x_1, x_2).
- The point $(0, 0, 0)^T$ is undefined.
- The space $\Re^3 \setminus (0, 0, 0)^T$ is called the projective plane $\mathcal{P}^2$.
- A homogenous point in $\mathcal{P}^2$ has 2 degrees of freedom.

Interpretation of the projective plane $\mathcal{P}^2$

- A homogenous vector $x \in \mathcal{P}^2$ may be interpreted as a line through the origin in $\Re^3$.
- The intersection with the plane $x_3 = 1$ gives the corresponding cartesian coordinates.



Transformations

- Transformation of homogenous 2D points may be described by multiplication by a 3×3 matrix

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix},$$

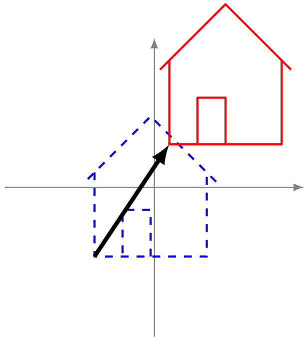
or

$$q = Ap.$$

Basic transformations — Translation

- A *translation* of points in $\Re^2$ may be described using homogenous coordinates as

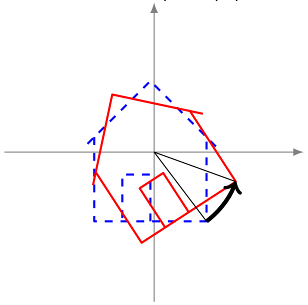
$$q = T(x_0, y_0)p = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + x_0 \\ y + y_0 \\ 1 \end{pmatrix}.$$



Basic transformations — Rotation

- A *rotation* may be described using homogenous coordinates as

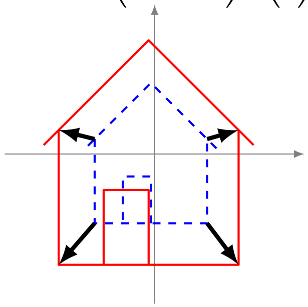
$$R(\varphi)p = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x\cos\varphi - y\sin\varphi \\ x\sin\varphi + y\cos\varphi \\ 1 \end{pmatrix}.$$



Basic transformations — Scaling

- *Scaling* of points in $\Re^2$ along the coordinate axes may be described using homogenous coordinates as

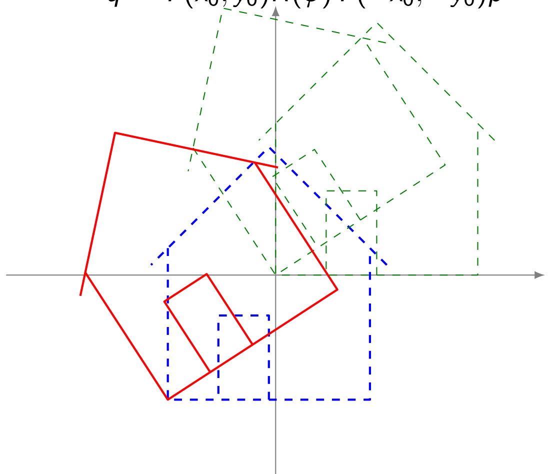
$$q = S(k, l)p = \begin{pmatrix} k & 0 & 0 \\ 0 & l & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} kx \\ ly \\ 1 \end{pmatrix}.$$



Combination of transformations

- Combinations of transformations are constructed by matrix multiplication:

$$q = T(x_0, y_0)R(\varphi)T(-x_0, -y_0)p$$



Transformation classes

- Transformation may be classified based on their properties.
- The most important transformations are
  - Similarity (rigid-body transformation).
  - Affinity.
  - Projectivity (homography).

Similarity

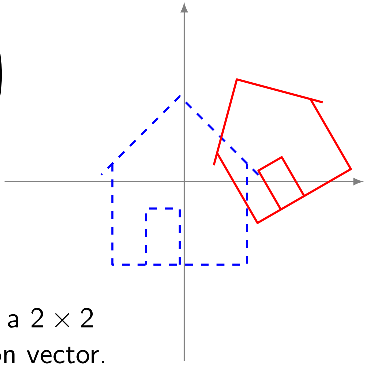
- A similarity transformation consists of a combination of rotations, isotropic scalings, and translations.

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

or

$$\begin{pmatrix} sR & t \\ 0 & 1 \end{pmatrix},$$

where the scalar s is the scaling, R is a 2×2 rotation matrix and t is the translation vector.



- A 2D similarity has 4 degrees of freedom.
- A similarity preserves angles (and "shape").

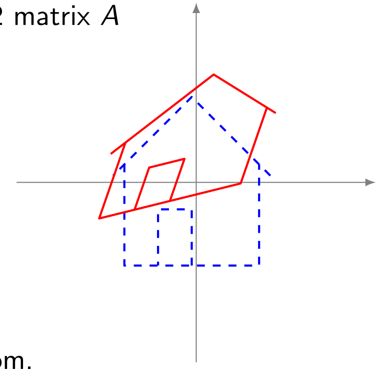
Affinity

- For an affine transformation the rotation and scaling is replaced by any non-singular 2×2 matrix A

$$\begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$$

or

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

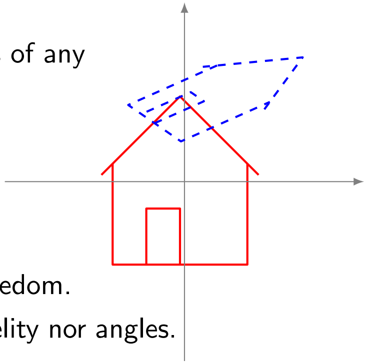


- A 2D affinity has 6 degrees of freedom.
- A similarity preserves parallelity but not angles.

Projectivity (Homography)

- A projectivity or homography consists of any non-singular 3×3 matrix H

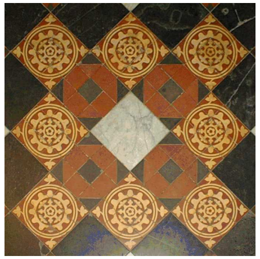
$$\begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}.$$



- A 2D projectivity has 8 degrees of freedom.
- A projectivity preserves neither parallelity nor angles.

The effect of different transformations

Similarity



Affinity

Projectivity

Planar rectification

- If the coordinates for 4 points p_i and their mappings $q_i = Hp_i$ in the image are known, we may calculate the homography H.
- From each point pair $p_i = (x_i, y_i, 1)^T, q_i = (x_i', y_i', 1)^T$ we get the following equations:

$$\begin{pmatrix} x_i' \\ y_i' \\ 1 \end{pmatrix} = \begin{pmatrix} u/w \\ v/w \\ 1 \end{pmatrix}, \text{ where } \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix}$$

or

$$x_i' = u/w = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}},$$
$$y_i' = v/w = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}}.$$

Planar rectification (2)

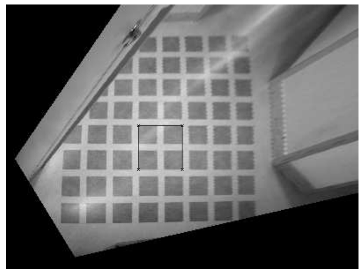
- Rearranging

$$x_i'(h_{31}x_i + h_{32}y_i + h_{33}) = h_{11}x_i + h_{12}y_i + h_{13},$$
$$y_i'(h_{31}x_i + h_{32}y_i + h_{33}) = h_{21}x_i + h_{22}y_i + h_{23}.$$

- This equation is linear in h_{ij}.
- Given 4 points we get 8 equations, enough to uniquely determine H assuming the points are in "standard position", i.e. no 3 points are collinear.

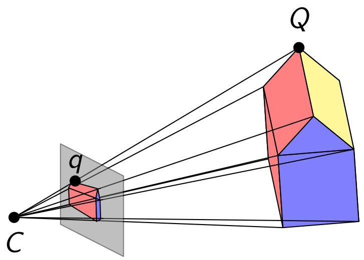
Planar rectification (3)

- Given H we may apply H^{-1} to remove the effect of the homography.



The pinhole camera model

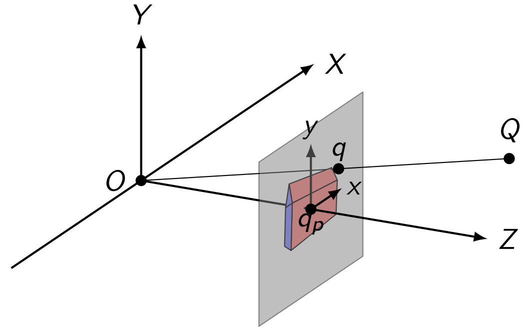
- The most commonly used camera model is called the *pinhole camera*.
- In the pinhole camera model:
  - All object points Q are projected via a *central projection* through the same point C, called the *camera center*.
  - The object point Q, the camera center C, and the projected point q are *collinear*.
  - A pinhole camera is *straight line-preserving*.



The central projection

- If the camera center is at the origin and the image plane is the plane $Z = c$, the world coordinate $(X, Y, Z)^T$ is mapped to the point $(cX/Z, cY/Z, c)^T$ in space or $(cX/Z, cY/Z)$ in the image plane, i.e.

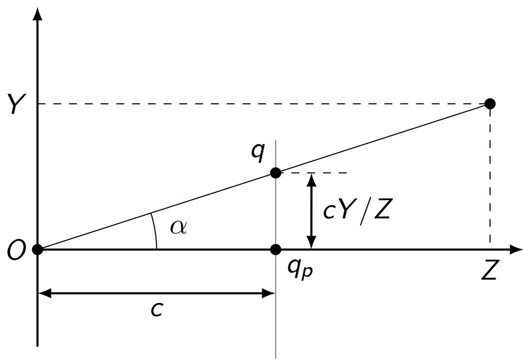
$$(X, Y, Z)^T \mapsto (cX/Z, cY/Z)^T$$



The central projection (2)

- If the camera center is at the origin and the image plane is the plane $Z = c$, the world coordinate $(X, Y, Z)^T$ is mapped to the point $(cX/Z, cY/Z, c)^T$ in space or $(cX/Z, cY/Z)$ in the image plane, i.e.

$$(X, Y, Z)^T \mapsto (cX/Z, cY/Z)^T$$



The central projection (3)

- The corresponding expression in homogenous coordinates may be written as

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \underset{q}{\begin{pmatrix} cX \\ cY \\ Z \end{pmatrix}} = \underset{P}{\begin{pmatrix} c & & & 0 \\ & c & & 0 \\ & & 1 & 0 \end{pmatrix}} \underset{Q}{\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}}.$$

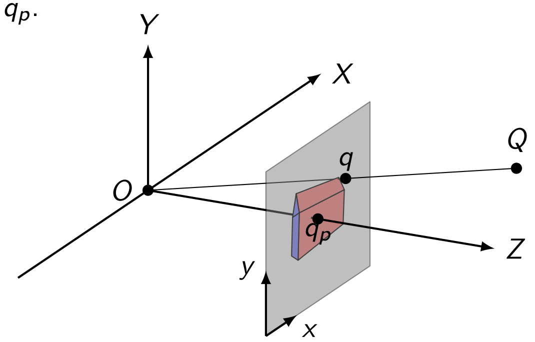
- The matrix P is called the *camera matrix* and maps the world point Q onto the image point q.
- In more compact form P may be written as

$$P = \operatorname{diag}(c, c, 1) \left(I \mid 0 \right),$$

where $\operatorname{diag}(c, c, 1)$ is a diagonal matrix and I is the 3×3 identity matrix.

The principal point

- If the principal point is not at the origin of the image coordinate system, the mapping becomes
$$(X, Y, Z)^T \mapsto (cX/Z + p_x, cY/Z + p_y)^T,$$
where $(p_x, p_y)^T$ are the image coordinates for the principal point q_p.



The principal point (2)

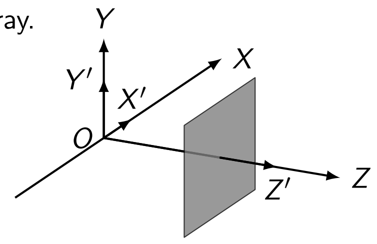
- In homogenous coordinates
$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mapsto \begin{pmatrix} cX/Z + p_x \\ cY/Z + p_y \end{pmatrix}$$
becomes
$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} cX + Zp_x \\ cY + Zp_y \\ Z \end{pmatrix} = \begin{pmatrix} c & & p_x & 0 \\ & c & p_y & 0 \\ & & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

The camera calibration matrix

- If we write
$$K = \begin{pmatrix} c & & p_x \\ & c & p_y \\ & & 1 \end{pmatrix},$$
the projection may be written as
$$q = K \left(I \mid 0 \right) Q.$$
- The matrix K is known as the *camera calibration matrix*.

The camera position and orientation

- Introduce
$$Q' = \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} \text{ and } q = K \left(I \mid 0 \right) Q'.$$
to describe coordinates in the *camera coordinate system*.
- The camera and world coordinate systems are identical if the camera center is at the origin, the X and Y axes coincide with the sensor coordinate system and the Z axes coincide with the principal ray.

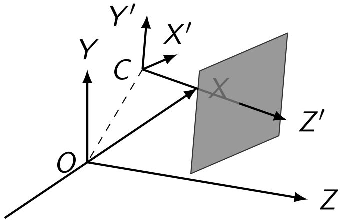


The camera position and orientation (2)

- In the general case, the transformation between the coordinate systems is usually described as

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = R\left(\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix}\right),$$

where $C = (X_0, Y_0, Z_0)^T$ is the camera center in world coordinates and the rotation matrix R describes the rotation from world coordinates to camera coordinates.



The camera position and orientation (3)

- In homogenous coordinates, this transformation becomes

$$Q' = \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I & -C \\ 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} R & -RC \\ 0 & 1 \end{pmatrix} Q.$$

- The full projection is given by

$$q = KR\left(I \mid -C\right) Q.$$

- The equation

$$q = PQ = KR\left(I \mid -C\right) Q,$$

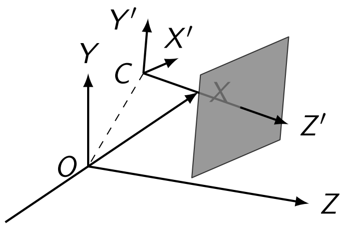
is sometimes referred to as the *camera equation*.
- The 3×4 matrix P is known as the *camera matrix*.

The camera position and orientation (4)

- If the transformation from the world to the camera is written as

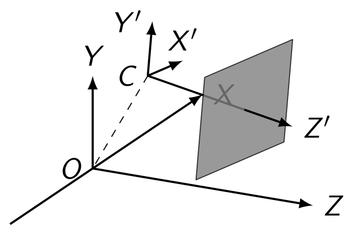
$$Q' = \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = R\left(\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix}\right),$$

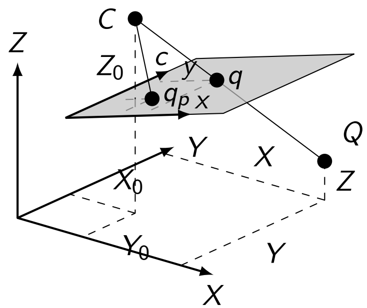
how does the transformation from the camera to the world look like?



Camera coordinates

- What are the (Z) coordinates of points *in front of* the camera?





The collinearity equations (revisited)

- Given

$$K = \begin{pmatrix} -c & & x_q \\ & -c & y_p \\ & & 1 \end{pmatrix},$$

the camera equation

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = q = KR\,(I \mid -C)\,Q = KR\,(I \mid -C)\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

becomes

$$x = x_p - c\frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)},$$

$$y = y_p - c\frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}.$$

Internal and external parameters

- The camera equation

$$q = KR\,(I \mid -C)\,Q$$

that describes the general projection for a pinhole camera has 9 degrees of freedom: 3 in K (the elements c, p_x, p_y), 3 in R (rotation angles) and 3 for C.
- The elements of K describes properties internal to the camera while the parameters of R and C describe the relation between the camera and the world.
- The parameters are therefore called one of

K	R, C
internal parameters	*external parameters*
internal orientation	*external orientation*
intrinsic parameters	*extrinsic parameters*
sensor model	*platform model*

Aspect ratio

- If we have different scale in the x and y directions, i.e. the pixels are not square, we have to include that deformation into the equation.
- Let m_x and m_y be the number of pixels per unit in the x and y direction of the image. Then the camera calibration matrix becomes

$$K = \begin{pmatrix} m_x & & \\ & m_y & \\ & & 1 \end{pmatrix}\begin{pmatrix} c & & p_x \\ & c & p_y \\ & & 1 \end{pmatrix} = \begin{pmatrix} m_x c & & m_x p_x \\ & m_y c & m_y p_y \\ & & 1 \end{pmatrix} = \begin{pmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{pmatrix},$$

where $\alpha_x = fm_x$ and $\alpha_y = fm_y$ is the camera constant in pixels in the x and y directions and $(x_0, y_0)^T = (m_x p_x, m_y p_y)^T$ is the principal point in pixels.
- A camera with unknown aspect ratio has 10 degrees of freedom.

Skew

- For an even more general camera model we can add a *skew* parameter s to describe any non-orthogonality between the image axis. Then the camera calibration matrix becomes

$$K = \begin{pmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{pmatrix}.$$

- The complete 3×4 camera matrix

$$P = KR\,(I \mid -C)$$

has 11 degrees of freedom, the same as a 3×4 homogenous matrix.

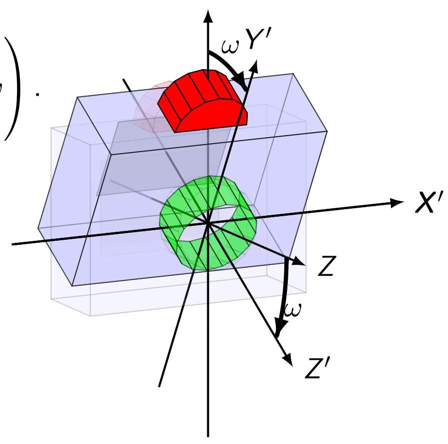
Rotations in $\Re^3$

- A rotation in $\Re^3$ is usually described as a sequence of 3 elementary rotations, by the so called *Euler angles*.
- *Warning: There are many different Euler angles and Euler rotations!*
- Each elementary rotation takes place about a cardinal axis, x, y, or z.
- The sequence of axis determines the actual rotation.
- A common example is the $\omega - \varphi - \kappa$ (omega-phi-kappa or x-y-z) convention that correspond to sequential rotations about the x, y, and z axes, respectively.

Elementary rotations (1)

- The first elementary rotation (ω, omega) is about the x-axis. The rotation matrix is defined as

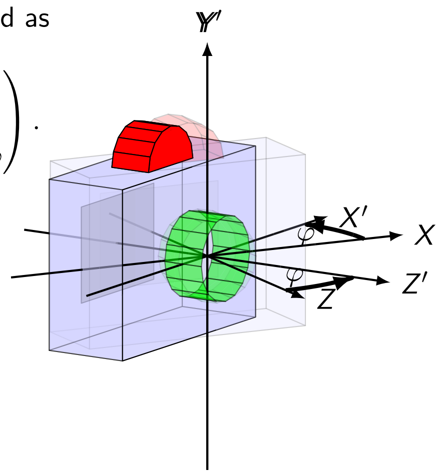
$$R_1(\omega) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{pmatrix}.$$



Elementary rotations (2)

- The second elementary rotation (φ, phi) is about the y-axis. The rotation matrix is defined as

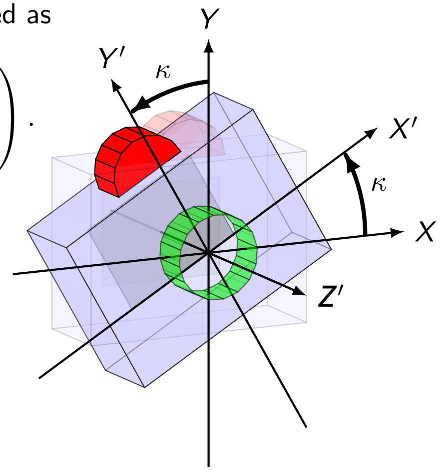
$$R_2(\varphi) = \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix}.$$



Elementary rotations (3)

- The third elementary rotation (κ, kappa) is about the z-axis. The rotation matrix is defined as

$$R_3(\kappa) = \begin{pmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

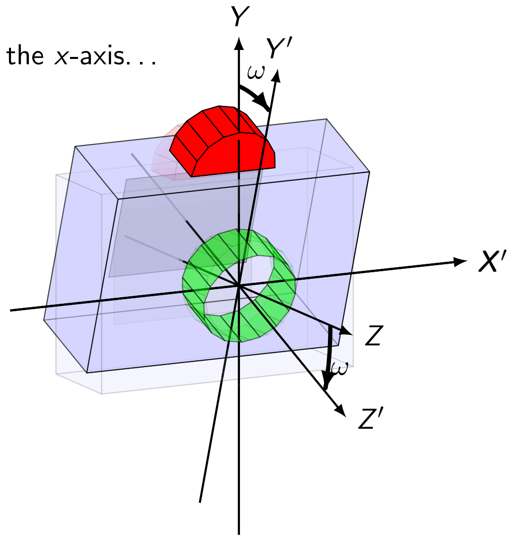


Combined rotations

- The axes follow the rotated object, so the second rotation is about a once-rotated axis, the third about a twice-rotated axis.
- A sequential rotation of 20 degrees about each of the axis is. . .

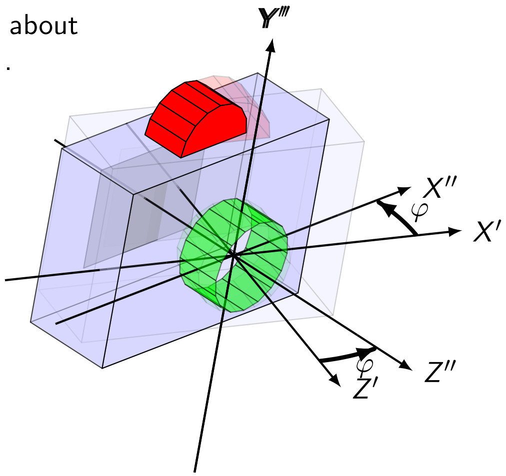
Combined rotations (2)

- . . . first a rotation about the x-axis. . .



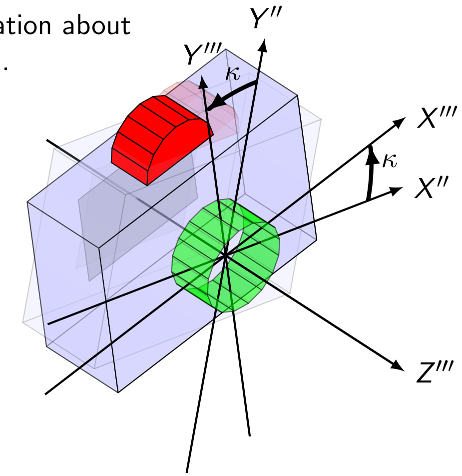
Combined rotations (3)

- . . . followed by a rotation about the once-rotated y-axis. . .



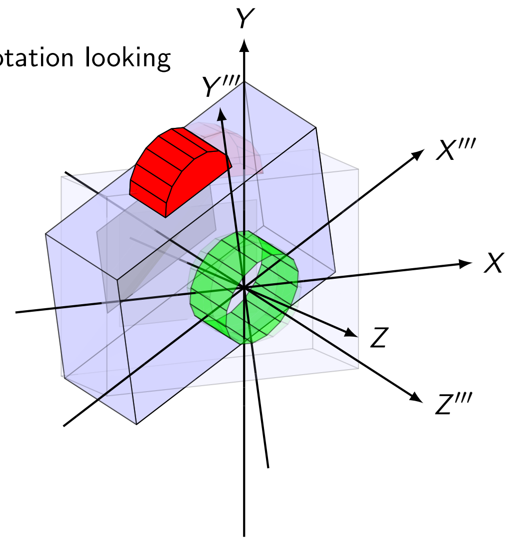
Combined rotations (4)

- . . . followed by a final rotation about the twice-rotated z-axis. . .



Combined rotations (5)

- ... resulting in a total rotation looking like this.

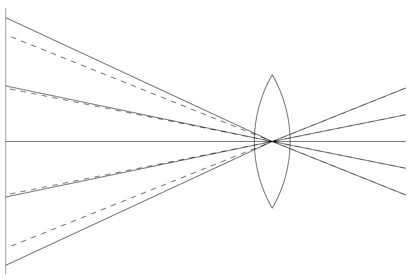


Rotations in $\Re^3$ (2)

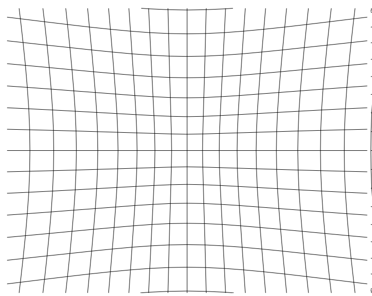
- The inverse rotation is about the same axes in reverse sequence with angles of opposite sign.
- This is sometimes called *roll-pitch-yaw*, where the κ angle is called the *roll* angle.
- Other rotations: azimuth-tilt-swing (z-x-z), axis-and-angle, etc.
- Every 3-parameter-description of a rotation has some rotation without a unique representation.
  - x-y-z if the middle rotation is 90 degrees,
  - z-x-z if the middle rotation is 0 degrees,
  - axis-and-angle when the rotation is zero (axis undefined).
- However, the *rotation* is always well defined.

Lens distortion

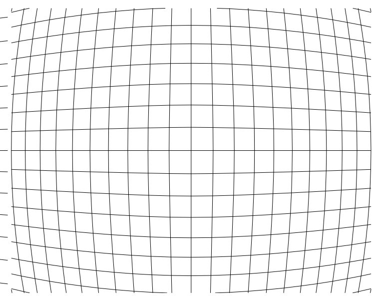
- A lens is designed to bend rays of light to construct a sharp image.
- A side effect is that the collinearity between incoming and outgoing rays is destroyed.



Lens distortion (2)



Positive radial distortion
(*pin-cushion*)



Negative radial distortion
(*barrel*)

Lens distortion (3)

- The effect of lens distortion is that the projected point is moved toward or away from a point of symmetry.
- The most common distortion model is due to Brown (1966, 1971).
- The distortion is separated into a symmetric (*radial*) and asymmetric (*tangential*) about the principal point:

$$\underbrace{\begin{pmatrix} x_c \\ y_c \end{pmatrix}}_{\text{corrected}} = \underbrace{\begin{pmatrix} x_m \\ y_m \end{pmatrix}}_{\text{measured}} + \left(\underbrace{\begin{pmatrix} x_r \\ y_r \end{pmatrix}}_{\text{radial}} + \underbrace{\begin{pmatrix} x_t \\ y_t \end{pmatrix}}_{\text{tangential}} \right).$$

- *Warning: Someone's positive distortion is someone else's negative!*

Lens distortion (4)

- The radial distortion is formulated as

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = (K_1 r^2 + K_2 r^4 + \ldots) \begin{pmatrix} x_m \\ y_m \end{pmatrix},$$

for any number of coefficients (usually 1–2), where r is a function of the distance to the principal point

$$r^2 = \Delta x^2 + \Delta y^2, \text{ and } \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} x_m - x_p \\ y_m - y_p \end{pmatrix}.$$

- The tangential distortion is formulated as follows

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 2P_1 \Delta x \Delta y + P_2(r^2 + 2\Delta x^2) \\ 2P_2 \Delta x \Delta y + P_1(r^2 + 2\Delta y^2) \end{pmatrix},$$

Lens distortion (5)

- The radial distortion follows from that the lens bends rays of light. It is neglectable only for large focal lengths.
- Any tangential distortion is due to de-centering of the optical axis for the various lens components. It is neglectable except for high precision measurements.
- The lens distortion parameters are usually determined at camera calibration.
- *The lens distortion varies with the focal length. To use a calibrated camera, the focal length (and hence any zoom) must be the same as during calibration.*
- *Warning: Some internal parameters are strongly correlated, e.g. the tangential coefficients P_1, P_2 and the principal point. Any calibration including P_1, P_2 must have multiple images at roll angles 0 and 90 degrees.*

Overview

- Principles
- History
- Mathematical models
- Processing
- Applications

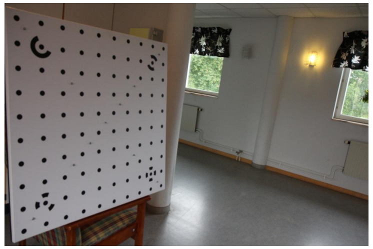
Processing

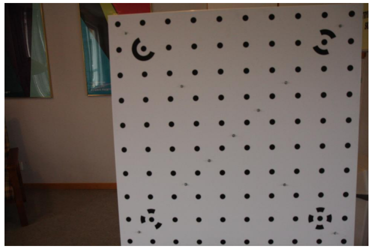
1. Camera calibration
2. Image acquisition
3. Measurements
4. Spatial resection
5. Forward intersection
6. (Bundle adjustment)

1. Camera calibration
2. Image acquisition
3. Measurements
4. Relative orientation
5. Forward intersection
6. (Bundle adjustment)
7. (Absolute orientation)

Camera calibration

- Special cameras may be calibrated by measuring deviation between input/output rays.
- Most of the time, camera calibration is performed by imaging a calibration object or scene.
- A 3D scene is preferable, but may be expensive.
- A 2D object is easier to manufacture and transport.





Camera calibration (2)

- Ideally, the calibration situation should mimic the actual scene.
- With a 2D object, multiple images must be taken.
- *Remember: use the same focal setting during calibration and image acquisition!*
- *If possible, include rolled images of the calibration object.*

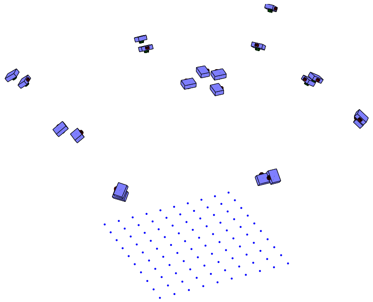
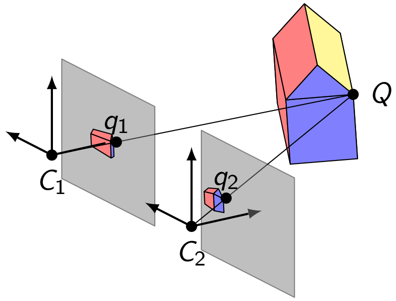


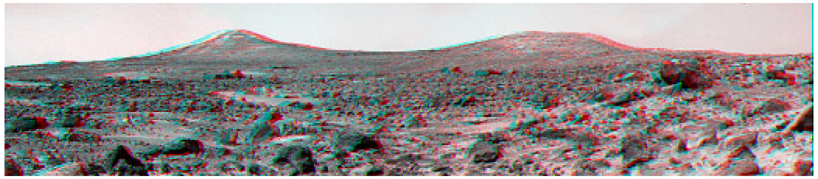
Image acquisition

- Camera networks
  - Parallel (stereo)
  - Convergent
  - Aerial
  - Other

Stereo images

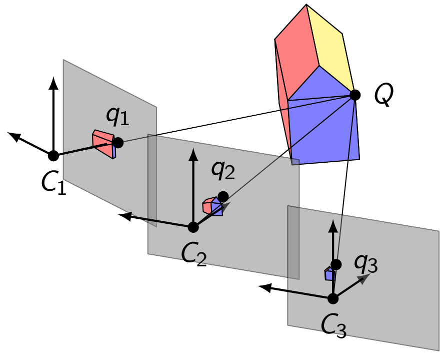
- Simplified measurements.
- Simplified automation.
- May be viewed in “3D”.





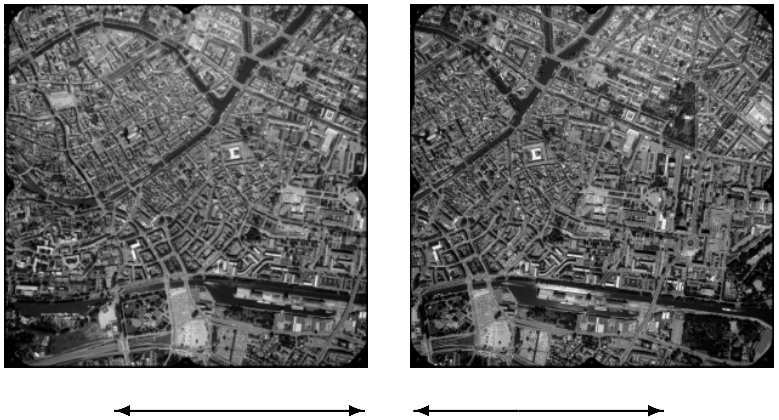
Convergent networks

- Stronger geometry.
- More than 2 measurements per object point.
- Should ideally surround the object.



Aerial networks

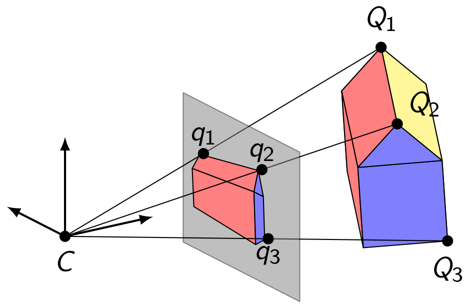
- Highly structurized.
- Typically around 60% overlap (along-track) and 30% sidelap (cross-track).



Measurements

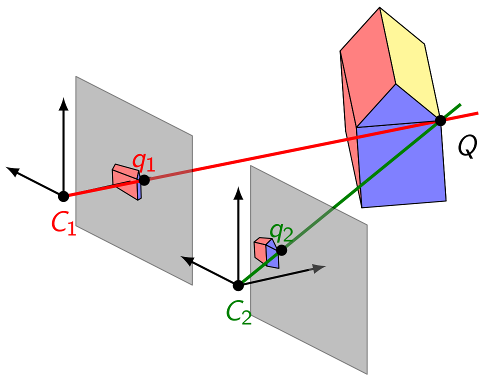
Spatial resection

- Determine the external orientation C, R of the camera from measurements and (ground) control points.
- Direct method from 3 points — solve 4th order polynomial (Grunert 1841, Haralick 1991, 1994). May have multiple solutions.



Forward intersection

- If the camera external orientations are known, an object point may be estimated from measurements in (at least) two images.
- Requires at least two observations.
- Linear estimation, robust.



Forward intersection (2)

- From the left camera we know that

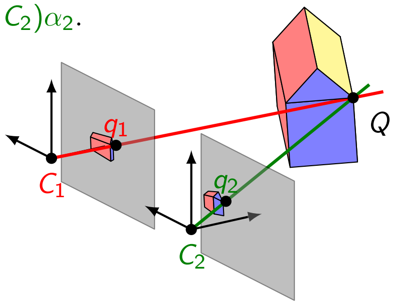
$$Q = C_1 + (v_1 - C_1)\alpha_1,$$

for some value of α_1, where v_1 are the 3D coordinates of q_1.

- Similarly, for the right camera

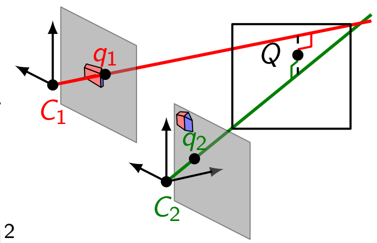
$$Q = C_2 + (v_2 - C_2)\alpha_2.$$

- We have 3+3 equations and 5 unknowns (Q, α_1, α_2).
- In theory, the point Q is at the intersection of the two lines, so we drop 1 equation and solve the remaining 5 to get Q.



Forward intersection (3)

- In reality, the lines may not intersect.
- In that case, we may choose to find the point that is *closest to both lines at the same time*, i.e. that solves the following minimization problem

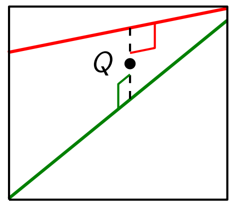


$$\min_{Q,\alpha_1,\alpha_2} \|Q - l_1(\alpha_1)\|^2 + \|Q - l_2(\alpha_2)\|^2,$$

or

$$\min_{Q,\alpha_1,\alpha_2} \left\| \begin{bmatrix} Q - (C_1 + t_1\alpha_1) \\ Q - (C_2 + t_2\alpha_2) \end{bmatrix} \right\|^2,$$

where $t_i = x_i - C_i$.



Forward intersection (4)

- This problem is linear in the unknowns and may be rewritten

$$\min_x \| \underbrace{\begin{bmatrix} I_3 & -t_1 & 0 \\ I_3 & 0 & -t_2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} Q \\ \alpha_1 \\ \alpha_2 \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}}_{b} \|^2.$$

- The solution is given by the *normal equations*

$$A^T A x = A^T b.$$

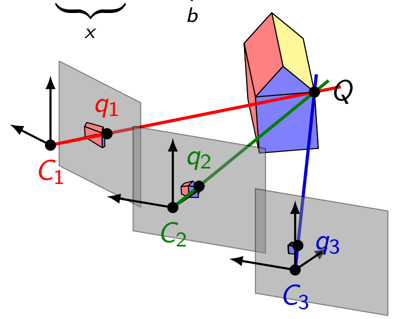
Forward intersection (5)

- Given one more camera, we extend the equation system

$$\min_x \| \underbrace{\begin{bmatrix} I_3 & -t_1 & 0 & 0 \\ I_3 & 0 & -t_2 & 0 \\ I_3 & 0 & 0 & -t_3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} Q \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix}}_{b} \|^2,$$

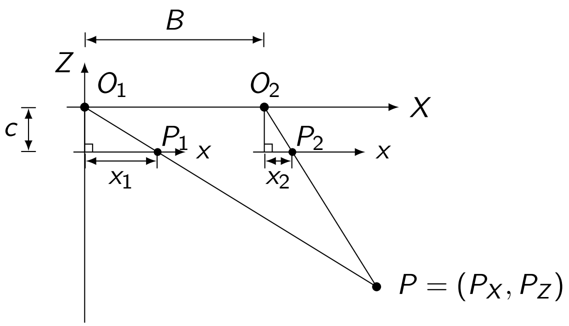
with the solution again given by the normal equations

$$A^T A x = A^T b.$$



Forward intersection (6)

- Stereorestitution ("normal case")



In photo 1:

$$X = Z\frac{x_1}{-c} \qquad Y = Z\frac{y_1}{-c}$$

In photo 2:

$$X = B + Z\frac{x_2}{-c} \qquad Y = Z\frac{y_2}{-c}$$

Forward intersection (7)

- If we have non-zero y parallax, i.e. $p_y = y_1 - y_2 \neq 0$, we must approximate.
- Otherwise,

$$-Z\frac{x_1}{-c} = B - Z\frac{x_2}{-c},$$
$$-Z = \frac{Bc}{x_1 - x_2} = \frac{Bc}{p_x}.$$

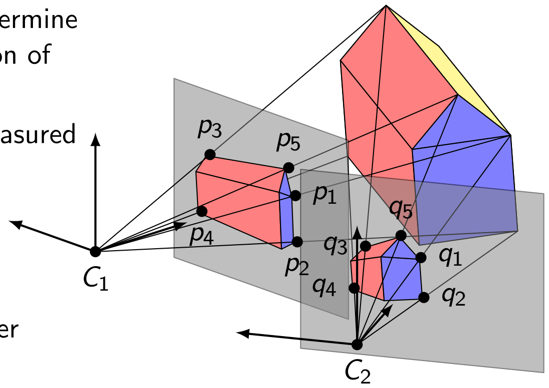
- Error propagation (first order)

$$\sigma_Z = \frac{Bc}{p_x^2}\sigma_x = \frac{Z}{c}\frac{Z}{B}\sigma_x.$$

- The ratio B/Z is the *base/object distance*.
- The ratio Z/c is the *scale factor*.

Relative orientation

- One camera fixed, determine position and orientation of second camera.
- Need 5 point pairs measured in both images.
- No 3D information is necessary.
- Direct method (Nistér 2004). Solve 10th order polynomial. May have multiple solutions.



Absolute orientation

- A 3D similarity transformation.
- 7 degrees of freedom (3 translations, 3 rotations, 1 scale).
- Direct method based on singular value decomposition (Arun 1987) for isotropic errors.

Bundle adjustment

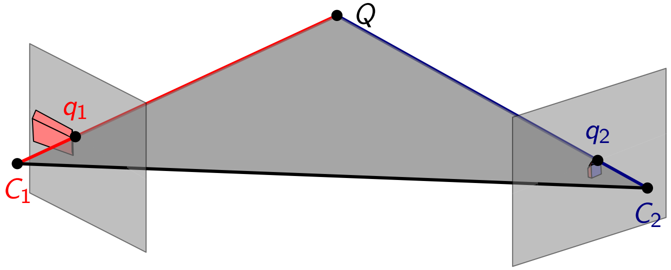
- Simultaneous estimation of camera external orientation and object points.
- Iterative method, needs initial values.
- May diverge.

Applications

- Architecture
- Forensics
- Maps
- Industrial
- Motion analysis
- Movie industry
- Orthopaedics
- Space science
- Microscopy
- GIS

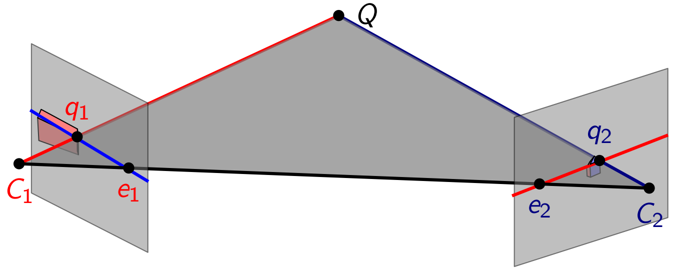
Epipolar geometry

- Let Q be an object point and q_1 and q_2 its projections in two images through the camera centers C_1 and C_2.
- The point Q, the camera centers C_1 and C_2 and the (3D points corresponding to) the projected points q_1 and q_2 will lie in the same plane.
- This plane is called the *epipolar plane* for C_1, C_2 and Q.



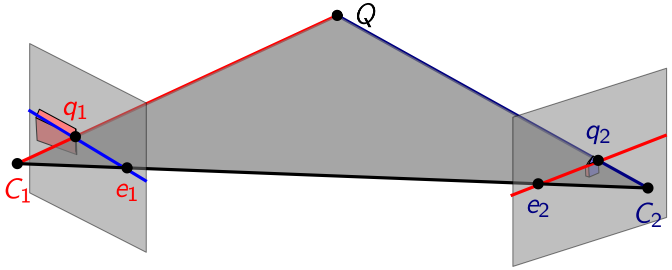
Epipolar lines

- Given a point q_1 in image 1, the epipolar plane is defined by the ray through q_1 and C_1 and the baseline through C_1 and C_2.
- A corresponding point q_2 thus has to lie on the intersecting line l_2 between the epipolar plane and image plane 2.
- The line l_2 is the projection of the ray through q_1 and C_1 in image 2 and is called the *epipolar line* to q_1.

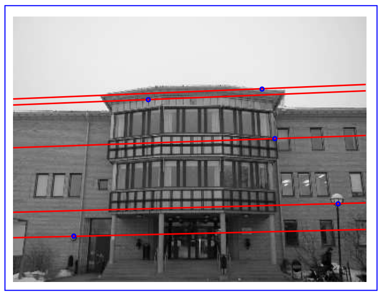


Epipoles

- The intersection points between the base line and the image planes are called *epipoles*.
- The epipole e_2 in image 2 is the mapping of the camera center C_1.
- The epipole e_1 in image 1 is the mapping of the camera center C_2.



Examples



Robust estimation — RANSAC

- The *Random Sample Consensus* (RANSAC) algorithm (Fishler and Bolles, 1981) is an algorithm for handling observations with large errors (outliers).
- Given a model and a data set S containing outliers:
  - Pick randomly s data points from the set S and calculate the model from these points. For a line, pick 2 points.
  - Determine the *consensus set* S_i of s, i.e. the set of points being within t units from the model. The set S_i define the inliers in S.
  - If the number of inliers are larger than a threshold T, recalculate the model based on all points in S_i and terminate.
  - Otherwise repeat with a new random subset.
  - After N tries, choose the largest consensus set S_i, recalculate the model based on all points in S_i and terminate.

Fundamentals of Photogrammetry
└ Epipolar geometry
└ RANSAC