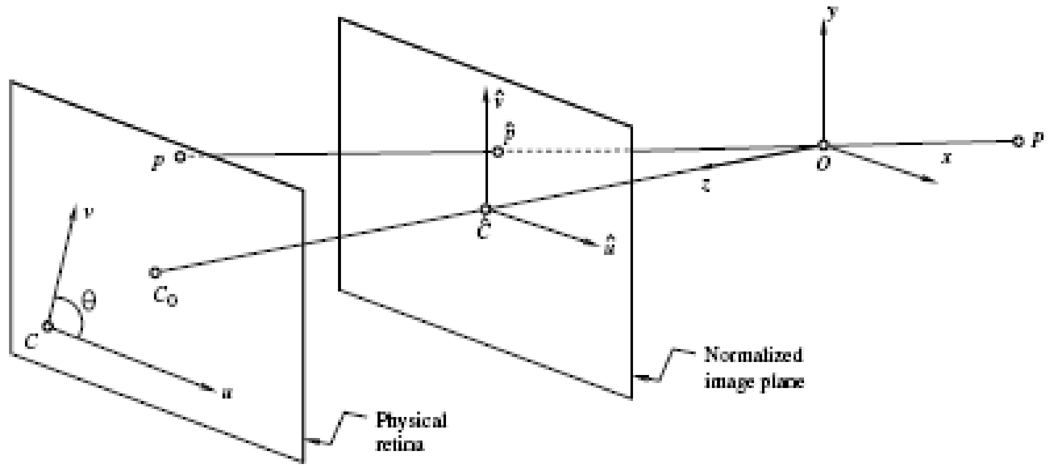
MAPI – Computer Vision

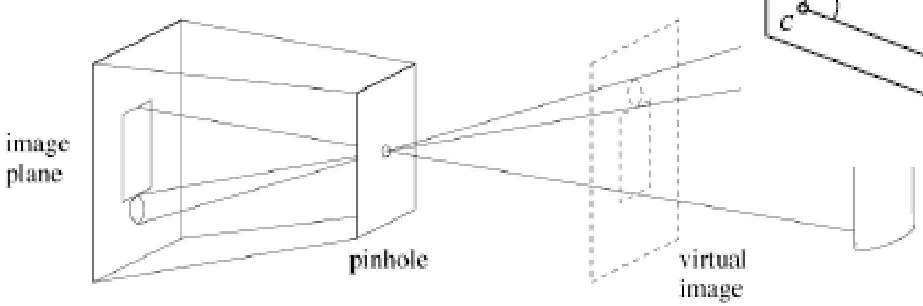
Multiple View Geometry

Intrinsic Camera Parameters

Intrinsic Camera Parameters

- We can associate with a camera two different image planes:
  - the first one is a normalized plane located at a unit distance from the pinhole.
    - We attach to this plane its own coordinate system with an origin located at the point Ĉ where the optical axis pierces it





Intrinsic Camera Parameters

- We can associate with a camera two different image planes:
  - the first one is a normalized plane located at a unit distance from the pinhole.
    - The perspective projection equation can be written in this normalized coordinate system as

$$\begin{cases} \hat{u} = \dfrac{x}{z} \\ \hat{v} = \dfrac{y}{z} \end{cases} \Leftrightarrow \hat{p} = \frac{1}{z}\begin{pmatrix} \mathrm{Id} & 0 \end{pmatrix}\begin{pmatrix} P \\ 1 \end{pmatrix}$$

- where

$$\hat{p}$$

is the vector of homogeneous coordinates of the projection $\hat{p}$ of the point P into the normalized image plane.

Intrinsic Camera Parameters

- The physical retina of the camera is in general different
  - it is located at a distance f from the pinhole
  - measurement in image coordinate system may be in "pixel" units (u,v),
  - pixels may not be rectangular,
    - so the camera has two additional scale parameters k and l
  - origin of image coordinate system may not be at the center of *image* (projection of lens center) (u_0, v_0)
  - axis may be *skewed* (θ)

Intrinsic Camera Parameters

- The physical retina of the camera is in general different
  - f is a distance, expressed in meters for example

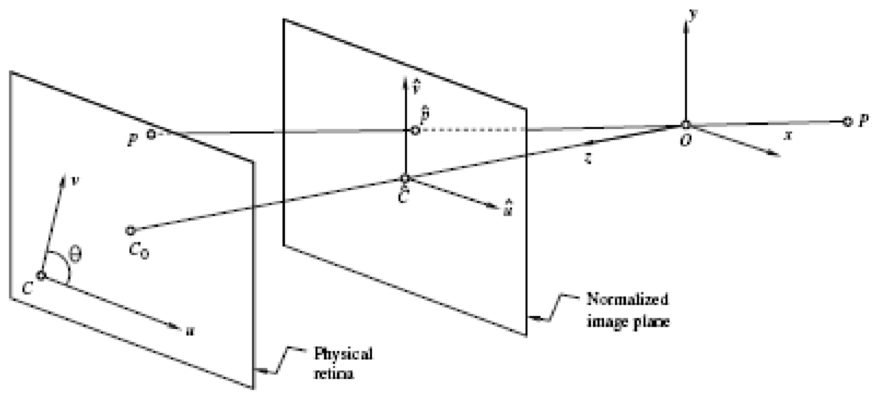
$$\begin{cases} u = kf\,\dfrac{x}{z} \\ v = lf\,\dfrac{y}{z} \end{cases}$$

  - a pixel will have dimensions $1/k$ x $1/l$ where k and l are expressed in pixel×m^{-1}.
  - The parameters k, l and f can be replaced by the
  - magnifications $\alpha = k\,f$ and $\beta = l\,f$ expressed in pixel units.

Intrinsic Camera Parameters

- The physical retina of the camera is in general different
  - The actual origin of the camera coordinate system is at a corner C of the retina and not at its center,
  - The center of the CCD matrix usually does not coincide with the principal point C_0.
  - This adds two parameters u_0 and v_0 that define the position (in pixel units) of C_0 in the retinal coordinate system.

$$\begin{cases} u = \alpha \dfrac{x}{z} + u_0 \\ v = \beta \dfrac{y}{z} + v_0 \end{cases}$$



Intrinsic Camera Parameters

- The physical retina of the camera is in general different
  - the camera coordinate system may also be skewed, due to some manufacturing error,
    - The angle θ between the two image axes is not equal to (but of course not very different from either) 90 degrees.

$$\begin{cases} u = \alpha \dfrac{x}{z} - \alpha \cot\theta \dfrac{y}{z} + u_0 \\ v = \dfrac{\beta}{\sin\theta} \dfrac{y}{z} + v_0 \end{cases}$$

Intrinsic Camera Parameters

- Planar affine transformation
  - between the physical image frame and the normalized:

$$p = \mathrm{K}\hat{p} \qquad p = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad \mathrm{K} = \begin{pmatrix} \alpha & -\alpha \cot\theta & u_0 \\ 0 & \dfrac{\beta}{\sin\theta} & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$p = \frac{1}{z}\mathrm{M}P$$

P denotes the homogeneous coordinate vector of ***P*** in the camera coordinate system:

homogeneous coordinates have allowed us to represent the perspective projection mapping by the 3 ×4 matrix **M**.

$$\mathrm{M} = \begin{pmatrix} \mathrm{K} & 0 \end{pmatrix}$$

Extrinsic Parameters

- Relate camera frame, **C**, to world (object) frame, ***W***
  - General transformation

$$\begin{pmatrix} {}^{B}P \\ 1 \end{pmatrix} = {}^{B}_{A}T \begin{pmatrix} {}^{A}P \\ 1 \end{pmatrix} \qquad {}^{B}_{A}T = \begin{pmatrix} {}^{B}_{A}R & {}^{B}O_{A} \\ 0^{\mathrm{T}} & 1 \end{pmatrix}$$

  - Camera frame

$${}^{C}P = \begin{pmatrix} {}^{C}_{W}R & {}^{C}O_{W} \end{pmatrix} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

Extrinsic Parameters

- Relate camera frame, ***C***, to world (object) frame, ***W***

$$p = \frac{1}{z} \mathrm{M} P \qquad\qquad \mathrm{M} = \mathrm{K}\begin{pmatrix} \mathrm{R} & t \end{pmatrix}$$

- The matrix **M** is also defined with 11 free coefficients.
- Note that there are:
  - 5 intrinsic parameters (α, β, u_0, v_0 and θ) and
  - 6 extrinsic parameters
    - the three angles defining R
    - and the three coordinates of t,
- which matches the number of independent coefficients of M.

Extrinsic Parameters

- Relate camera frame, ***C***, to world (object) frame, ***W***
- The matrix M can be rewritten explicitly as a function of the intrinsic and extrinsic parameters of the camera:

$$\mathrm{M} = \begin{pmatrix} \alpha r_1^T - \alpha \cot\theta r_2^T + u_0 r_3^T & \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \dfrac{\beta}{\sin\theta} r_2^T + v_0 r_3^T & \dfrac{\beta}{\sin\theta} t_y + v_0 t_z \\ r_3^T & t_z \end{pmatrix}$$

- r_1^T, r_2^T and r_3^T denote the three rows of the matrix R
- t_x, t_y and t_z are the coordinates of the vector t in the frame attached to the camera.

Calibration Methods

- Geometric camera calibration
  - techniques for estimating the intrinsic and extrinsic parameters of a camera

- Suppose that a camera observes ***n*** geometric features such as points or lines with known positions in some fixed world coordinate system.
  - (1) computing the perspective projection matrix M associated with the camera in this coordinate system
  - (2) computing the intrinsic and extrinsic parameters of the camera from this matrix.
- Once a camera has been calibrated, it is possible to **associate with any image point a well-defined ray** passing through this point and the **camera's optical** center, and to conduct quantitative three-dimensional measurements from digitized pictures

Calibration Methods

A Linear Approach to Camera Calibration

- If the 4-vectors P_i (i = 1, . . ., n) and $m^T{}_j$ (j = 1, 2, 3) denote respectively the homogeneous coordinate vectors of the points P_i and the rows of the matrix M, we can express the position of the image of each point as

$$\begin{cases} u_i = \dfrac{m_1 \cdot P_i}{m_3 \cdot P_i} \\ v_i = \dfrac{m_2 \cdot P_i}{m_3 \cdot P_i} \end{cases} \Leftrightarrow \begin{cases} (m_1 - u_i m_3) \cdot P_i = 0 \\ (m_2 - v_i m_3) \cdot P_i = 0 \end{cases}$$

Calibration Methods

A Linear Approach to Camera Calibration

- Collecting these constraints for all ***n*** points yields a system of ***2n*** homogeneous linear equations in the twelve coefficients of the matrix M
- System of equations:

$$Pm = 0, P = \begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \dots & \dots & \dots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix}, m = \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix}$$

Calibration Methods

A Linear Approach to Camera Calibration

- When $n \geq 6$, the system of equations is in general overconstrained

- The linear least-squares literature provides methods for computing the value of the unit vector $\boldsymbol{m}$ that minimizes $|\boldsymbol{Pm}|^2$.

- Estimating the vector **m** (hence the matrix **M**) reduces to computing the eigenvectors and eigenvalues of the 12×12 matrix $P^T P$.

Linear Least Squares Methods

- consider a system of *n* linear equations in *p* unknowns:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = b_1 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p = b_n \end{cases} \Leftrightarrow \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \ldots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \ldots \\ b_n \end{pmatrix}$$

- Let $\boldsymbol{A}$ denote the $\boldsymbol{n} \times \boldsymbol{p}$ matrix with coefficients a_{ij}, and let $x = (x_1, \ldots, x_p)^T$ and $b = (b_1, \ldots, b_n)^T$
- We know from linear algebra that (in general):
- 1. when $n < p$, there exists an $(p - n)$ dimensional vector space of vectors x that are solutions
- 2. when $n = p$, there is a unique solution;
- 3. when $n > p$, there is no solution.

- This statement is true when the rank of *A* is maximal, i.e., equal to min(n, p)
- When the rank is lower, there exists a higher-dimensional set of solutions.

Linear Least Squares Methods

- Consider the overconstrained case $n > p$.
- Since there is no exact solution in this case,
  - finding the vector x that minimizes the error measure

$$E = \sum_{i=1}^{n} (a_{i1}x_1 + \ldots + a_{ip}x_p - b_i)^2 = |Ax - b|^2$$

Recover the parameters

- Once the projection matrix M has been estimated, it can be used to recover the intrinsic and extrinsic parameters:

$$
\mathrm{M} = \begin{pmatrix} \alpha r_1^T - \alpha \cot\theta r_2^T + u_0 r_3^T & \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \dfrac{\beta}{\sin\theta} r_2^T + v_0 r_3^T & \dfrac{\beta}{\sin\theta} t_y + v_0 t_z \\ r_3^T & t_z \end{pmatrix} = \begin{pmatrix} A & b \end{pmatrix}
$$

Recover the parameters

- Using the fact that the **rows of a rotation matrix have unit length** and **are perpendicular to each** other yields

$$\begin{cases} \rho = \varepsilon / |a_3| \quad \text{, where } \varepsilon = \pm 1 \\ r_3 = \rho a_3 \\ u_0 = \rho^2 (a_1 \cdot a_3) \\ v_0 = \rho^2 (a_2 \cdot a_3) \end{cases}$$

$$\rho \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T \end{pmatrix} = \begin{pmatrix} \alpha r_1^T - \alpha \cot \theta r_2^T + u_0 r_3^T \\ \dfrac{\beta}{\sin \theta} r_2^T + v_0 r_3^T \\ r_3^T \end{pmatrix}$$

$$\begin{cases} \rho^2 (a_1 \times a_3) = -\alpha r_2 - \alpha \cot \theta r_1 \\ \rho^2 (a_2 \times a_3) = \dfrac{\beta}{\sin \theta} r_1 \end{cases}$$

and

$$\begin{cases} \rho^2 |a_1 \times a_3| = \dfrac{|\alpha|}{\sin \theta} \\ \rho^2 |a_2 \times a_3| = \dfrac{|\beta|}{\sin \theta} \end{cases}$$

Recover the parameters

$$\begin{cases} \cos\theta = -\varepsilon_u \varepsilon_v \dfrac{(a_1 \times a_3)\cdot(a_2 \times a_3)}{|a_1 \times a_3||a_2 \times a_3|} \\ \alpha = \varepsilon_u \rho^2 |a_1 \times a_3| \sin\theta \\ \beta = \varepsilon_v \rho^2 |a_2 \times a_3| \sin\theta \end{cases}$$

, where

$$\varepsilon_u = \alpha / |\alpha|$$

$$\varepsilon_v = \beta / |\beta|$$

$$\begin{cases} r_1 = \dfrac{\rho^2 \sin\theta}{\beta}(a_2 \times a_3) = \dfrac{1}{|a_2 \times a_3|}(a_2 \times a_3) \\ r_2 = r_3 \times r_1 \end{cases}$$

Recover the parameters

- the translation parameters are recovered by writing

$$\rho \begin{pmatrix} \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \dfrac{\beta}{\sin\theta} t_y + v_0 t_z \\ t_z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Taking Radial Distortion into Account

- so far that our camera was equipped with a perfect lens.
- real lenses suffer from a number of aberrations.
  - radial distortion
    - a type ofaberration that depends on the distance between the imaged point and the optical axis and can be modelled as

$$p = \begin{pmatrix} 1/\lambda & 0 & 0 \\ 0 & 1/\lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathrm{MP}$$

$$r^2 = u^2 + v^2$$

$$\lambda = 1 + k_1 r^2 + k_2 r^4 + \ldots$$

Taking Radial Distortion into Account

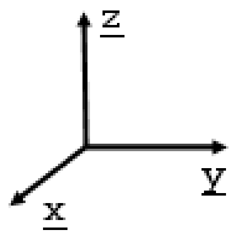
- Geometrically, radial distortion changes the distance between the image center and the image point ***p*** but it does not affect the direction of the vector joining these two points.
- This is called the radial alignment constraint and it can be expressed algebraically by writing

$$\lambda \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \dfrac{m_1.P}{m_3.P} \\ \dfrac{m_2.P}{m_3.P} \end{pmatrix} \Rightarrow v(m_1.P) - u(m_2.P) = 0$$

- This is a linear constraint on the vectors m_1 and m_2.
- Given ***n*** fiducial points we obtain ***n*** equations in the eight coefficients of the vectors $\boldsymbol{m_1}$ and $\boldsymbol{m_2}$

Produto Externo de 2 Vectores

- Denotando por $\underline{x}$, $\underline{y}$ e $\underline{z}$ os vectores unitários dos repectivos eixos, e de acordo com a definição, temos
  - $\mathbf{\underline{x} \times \underline{x} = \underline{y} \times \underline{y} = \underline{z} \times \underline{z} = 0}$ (pois $\underline{x}$ faz um ângulo de 0º com $\underline{x}$
  - $\mathbf{\underline{x} \times \underline{y} = \underline{z}}$; $\mathbf{\underline{x} \times \underline{z} = -\underline{y}}$; $\mathbf{\underline{y} \times \underline{z} = \underline{x}}$ (regra do saca-rolhas) ;
  - $\mathbf{\underline{y} \times \underline{x} = -\underline{z}}$; $\mathbf{\underline{z} \times \underline{x} = \underline{y}}$; $\mathbf{\underline{z} \times \underline{y} = -\underline{x}}$ (regra do saca-rolhas)



- Sendo o produto externo distributivo em relação à soma, dados dois vectores $\mathbf{A = a_x \underline{x} + a_y \underline{y} + a_z \underline{z}}$ e $\mathbf{B = b_x \underline{x} + b_y \underline{y} + b_z \underline{z}}$ o seu produto externo é dado por
- $A \times B = (a_x \underline{x} + a_y \underline{y} + a_z \underline{z}) \times (b_x \underline{x} + b_y \underline{y} + b_z \underline{z}) =$

  $(a_y b_z - a_z b_y) \underline{x} + (a_z b_x - a_x b_z) \underline{y} + (a_x b_y - a_y b_x) \underline{z}$