Visión por Computador

Luis Baumela

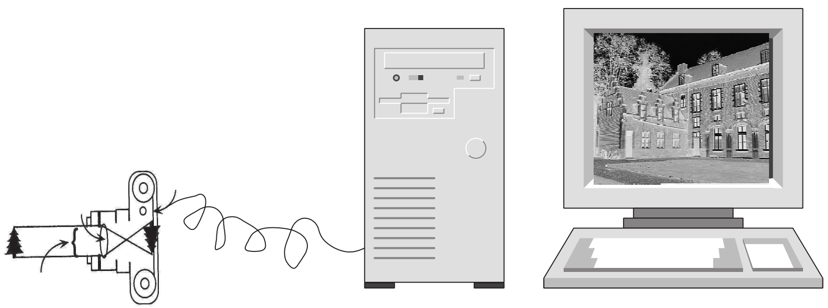
http://www.dia.fi.upm.es/~lbaumela

Departamento de Inteligencia Artificial

Universidad Politécnica de Madrid

Tema 4. Modelado de cámara y calibración

Componentes de un sistema de visión



Formación de imagen

Índice:

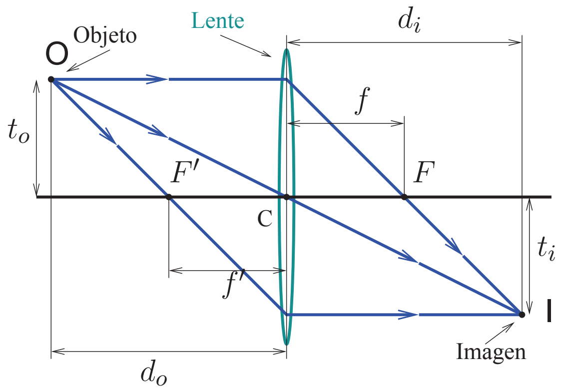
- Modelo de lente fina
- Modelo de proyección perspectiva
- Modelo proyectivo de cámara
- Modelo afín de cámara
- Calibración
- Bibliografía

Modelo de lente fina

- *Hipótesis:* La óptica de la lente puede modelarse como una lente fina.

Modelo de lente fina

- *Hipótesis:* La óptica de la lente puede modelarse como una lente fina.
- *Principio de funcionamiento:* Refracción de la luz.

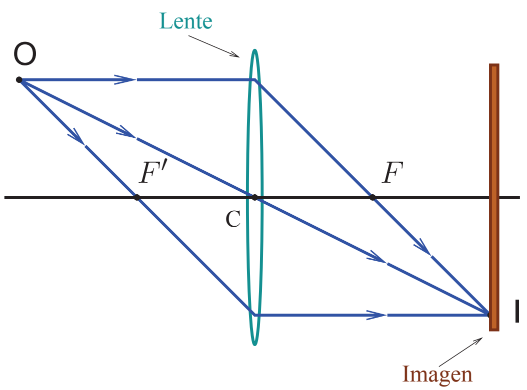


Ec fundamental:

$$\frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_o}$$

Modelo de lente fina

- *Hipótesis:* La óptica de la lente puede modelarse como una lente fina.
- *Principio de funcionamiento:* Refracción de la luz.

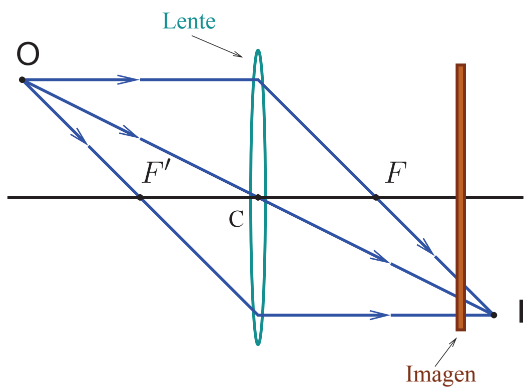


Ec fundamental:

$$\frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_o}$$

Modelo de lente fina

- *Hipótesis:* La óptica de la lente puede modelarse como una lente fina.
- *Principio de funcionamiento:* Refracción de la luz.



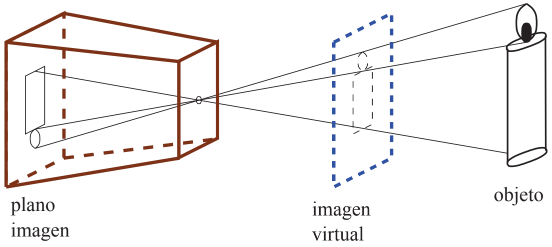
Ec fundamental:

$$\frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_o}$$

$f = d_i$, si $d_o \approx \infty$.

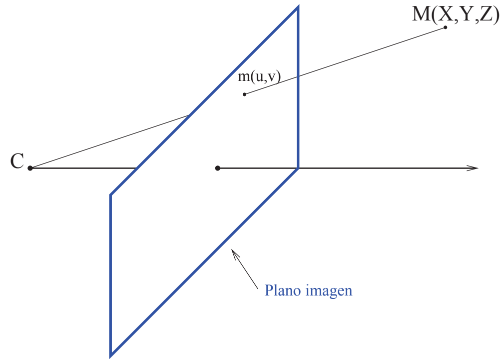
Modelo de proyección perspectiva

- Cámara oscura.



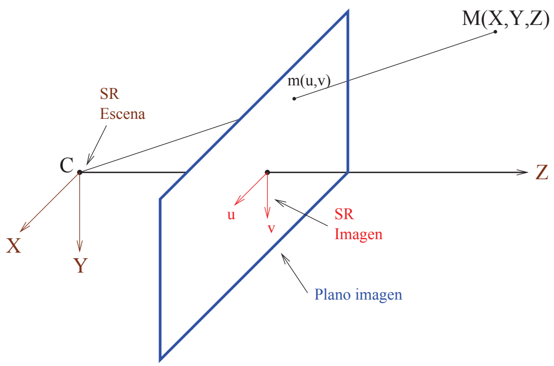
Modelo de proyección perspectiva

- Abstracción geométrica.
  La proyección de un punto viene dada por el corte del plano imagen con la visual que une el punto y el centro óptico de la cámara.



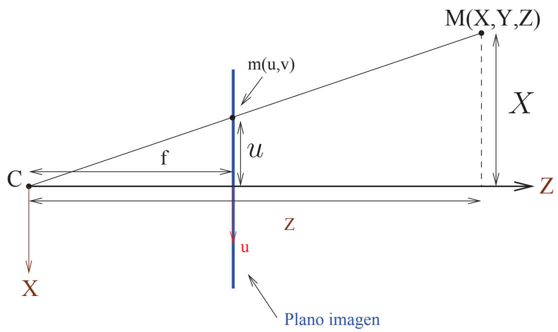
Modelo de proyección perspectiva

- Abstracción geométrica.
  La proyección de un punto viene dada por el corte del plano imagen con la visual que une el punto y el centro óptico de la cámara.



Modelo de proyección perspectiva

- Modelo de proyección.
  Relaciona las posiciones de puntos en la escena con su proyección sobre el plano imagen.



$$u = f\frac{X}{Z}$$

$$v = f\frac{Y}{Z}$$

Modelo proyectivo de cámara

1. El modelo es lineal empleando coordenadas homogéneas.

cartesianas		*homogéneas*
(X,Y,Z)	$\Leftrightarrow$	$(\lambda X, \lambda Y, \lambda Z, \lambda)$
$(\frac{A}{D}, \frac{B}{D}, \frac{C}{D})$	$\Leftrightarrow$	(A,B,C,D)

Modelo proyectivo de cámara

1. El modelo es lineal empleando coordenadas homogéneas.

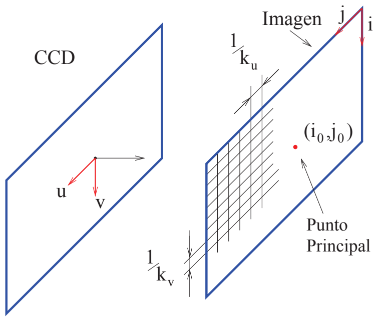
cartesianas		*homogéneas*
(X,Y,Z)	$\Leftrightarrow$	$(\lambda X, \lambda Y, \lambda Z, \lambda)$
$(\frac{A}{D}, \frac{B}{D}, \frac{C}{D})$	$\Leftrightarrow$	(A,B,C,D)

2. El modelo de proyección resultante, en forma matricial:

$$
\begin{array}{l} u = f\dfrac{X}{Z} \\ v = f\dfrac{Y}{Z} \end{array} \quad \equiv \quad \begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}
$$

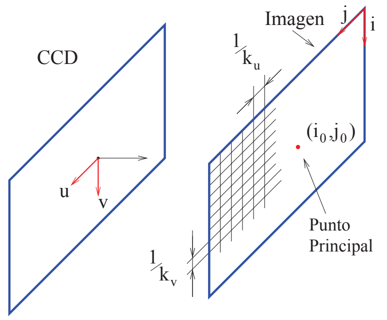
Modelo proyectivo de cámara

3. Modelo intrínseco.
   Digitalización $\equiv$ cambio de unidades $+$ traslación.



Modelo proyectivo de cámara

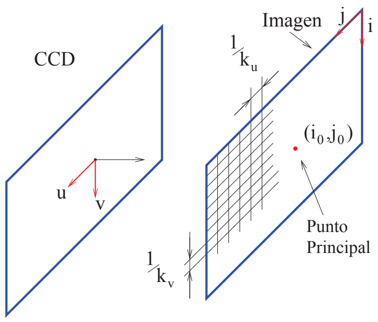
3. Modelo intrínseco.
   Digitalización $\equiv$ cambio de unidades $+$ traslación.



$$
\begin{aligned} j &= k_u u + j_0 \\ i &= k_v v + i_0 \end{aligned}
$$

Modelo proyectivo de cámara

3. Modelo intrínseco.
   Digitalización $\equiv$ cambio de unidades $+$ traslación.



$$\begin{pmatrix} \lambda j \\ \lambda i \\ \lambda \end{pmatrix} = \begin{pmatrix} k_u & 0 & i_0 \\ 0 & k_v & j_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

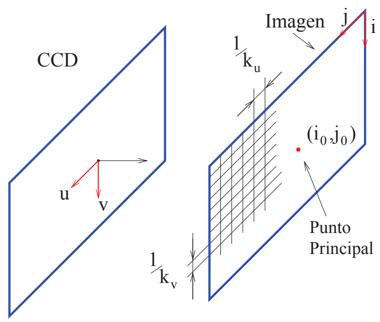
donde

- k_u: Densidad píxeles horizontal
- k_v: Densidad píxeles vertical

$$j = k_u u + j_0$$
$$i = k_v v + i_0$$

Modelo proyectivo de cámara

3. Modelo intrínseco.
   Digitalización $\equiv$ cambio de unidades $+$ traslación.

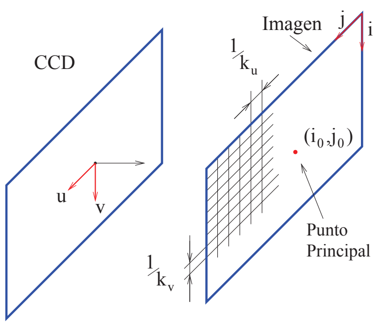


$$\begin{pmatrix} \lambda j \\ \lambda i \\ \lambda \end{pmatrix} = \begin{pmatrix} fk_u & s & i_0 & 0 \\ 0 & fk_v & j_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$j = k_u u + j_0$$
$$i = k_v v + i_0$$

Modelo proyectivo de cámara

3. Modelo intrínseco.
   Digitalización $\equiv$ cambio de unidades $+$ traslación.



$$\begin{pmatrix} \lambda j \\ \lambda i \\ \lambda \end{pmatrix} = \begin{pmatrix} fk_u & s & i_0 & 0 \\ 0 & fk_v & j_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

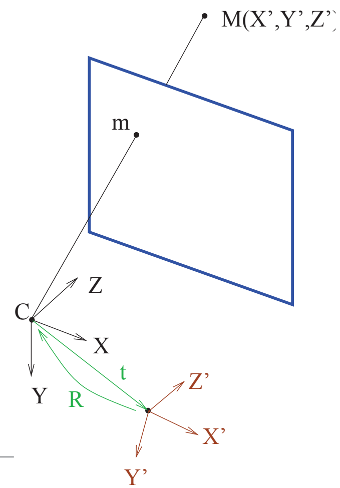
Parámetros intrínsecos:

- Punto principal: (i_0, j_0)
- Focal horizontal: $\alpha_u = fk_u$
- Focal vertical: $\alpha_v = fk_v$
- Sesgo: s

$$j = k_u u + j_0$$
$$i = k_v v + i_0$$

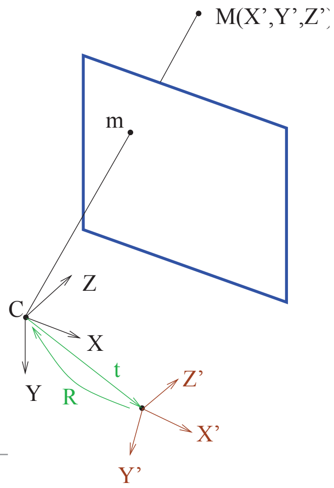
Modelo proyectivo de cámara

4. Modelo extrínseco.
   Movemos el SR de la escena fuera de la cámara



Modelo proyectivo de cámara

4. Modelo extrínseco.
   Movemos el SR de la escena fuera de la cámara



$$\begin{pmatrix} \lambda X \\ \lambda Y \\ \lambda Z \\ \lambda \end{pmatrix} = \left(\begin{array}{c|c} \mathbf{R}_{3\times3} & \bar{t}_{3\times1} \\ \hline \bar{0}_{1\times3} & 1 \end{array} \right) \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}$$

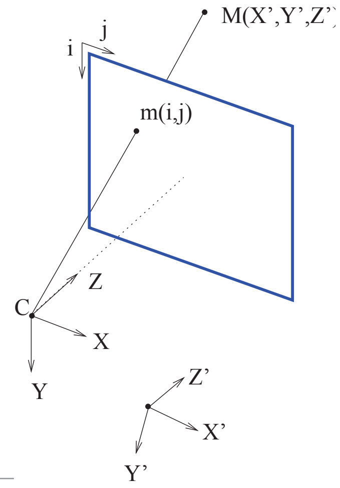
donde

$\mathbf{R}$: Rotación entre SR escena y cámara.

$\bar{t}$: Posición cámara en la escena.

Modelo proyectivo de cámara

5. Modelo de proyección perspectiva



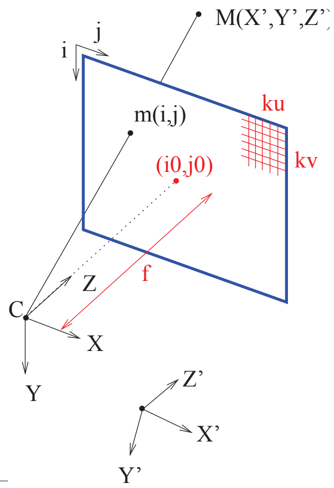
$$\begin{pmatrix} \lambda j \\ \lambda i \\ \lambda \end{pmatrix} = \underbrace{\mathbf{K}\,(\mathbf{R}_{3\times3}\,|\,\bar{t}_{3\times1})}_{\mathbf{P}_{3\times4}} \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}$$

11 parámetros del modelo:

$$\mathbf{P}_{3\times4} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{14} \\ \vdots & \vdots & \ddots & \vdots \\ p_{31} & p_{32} & \cdots & p_{34} \end{pmatrix}$$

Modelo proyectivo de cámara

5. Modelo de proyección perspectiva



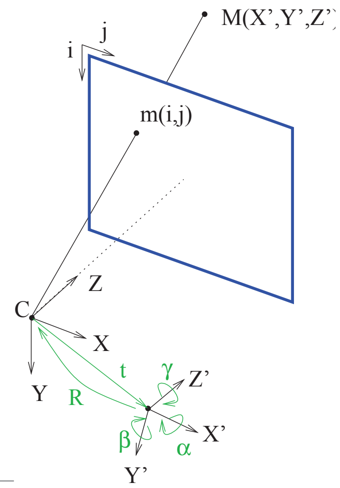
$$\begin{pmatrix} \lambda j \\ \lambda i \\ \lambda \end{pmatrix} = \mathbf{K}\,(\mathbf{R}_{3\times3}\,|\,\bar{t}_{3\times1}) \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}$$

5 parámetros intrínsecos:

$$\mathbf{K} = \begin{pmatrix} \alpha_u & s & j_0 \\ 0 & \alpha_v & i_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Modelo proyectivo de cámara

5. Modelo de proyección perspectiva



$$\begin{pmatrix} \lambda j \\ \lambda i \\ \lambda \end{pmatrix} = \mathbf{K}\,(\mathbf{R}_{3\times3}\,|\,\bar{t}_{3\times1}) \begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix}$$

6 parámetros extrínsecos:

$$\mathbf{R}_{3\times3}(\alpha, \beta, \gamma), \quad \bar{t}_{3\times1} = \begin{pmatrix} t_X \\ t_Y \\ t_Z \end{pmatrix}$$

Modelo proyectivo de cámara

6. Caracterización del modelo de proyección perspectiva.

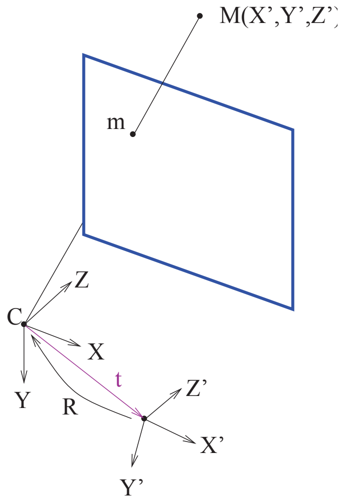
Sea $\mathbf{P} = (\mathbf{A} \,|\, \bar{b})$ una matriz de dimensión 3×4 y sean $(\bar{a}^i)^\top$ $(i = 1 \ldots 3)$ las filas de la matriz $\mathbf{A}$.

- $\mathbf{P}$ es una matriz de proyección perspectiva $\iff \det(\mathbf{A}) \neq 0$.
- $\mathbf{P}$ es una matriz de proyección perspectiva sin sesgo $(s = 0)$ $\iff \det(\mathbf{A}) \neq 0$ y $(\bar{a}^1 \times \bar{a}^3) \cdot (\bar{a}^2 \times \bar{a}^3) = 0$.
- $\mathbf{P}$ es una matriz de proyección perspectiva sin sesgo y con píxeles cuadrados $\left(\frac{\alpha_u}{\alpha_v} = 1\right) \iff \det(\mathbf{A}) \neq 0$ y

$$
\begin{aligned}
(\bar{a}^1 \times \bar{a}^3) \cdot (\bar{a}^2 \times \bar{a}^3) &= 0 \\
(\bar{a}^1 \times \bar{a}^3) \cdot (\bar{a}^1 \times \bar{a}^3) &= (\bar{a}_2 \times \bar{a}^3) \cdot (\bar{a}^2 \times \bar{a}^3)
\end{aligned}
$$

Modelo proyectivo de cámara

7. Propiedades.



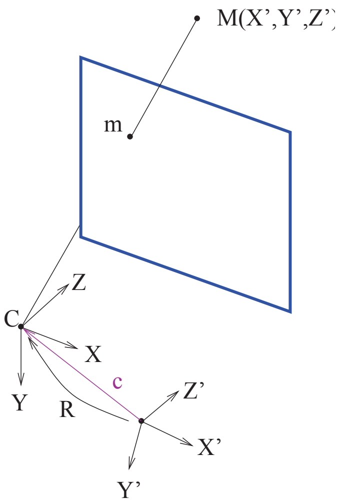
Sea la matriz de proyección $[\mathbf{A} \,|\, \bar{b}]$

Centro Óptico. $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Ya que:
$\mathbf{P}$ también puede expresarse como $\mathbf{P} = \mathbf{K}\,[\mathbf{R} \,|\, -\mathbf{R}\bar{c}]$. Luego, $\mathbf{A} = \mathbf{K}\,\mathbf{R}$ y $\bar{b} = -\mathbf{A}\bar{c}$, con lo que $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Modelo proyectivo de cámara

7. Propiedades.



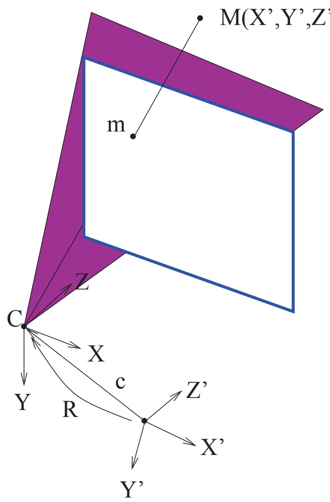
Sea la matriz de proyección $[\mathbf{A} \,|\, \bar{b}]$

Centro Óptico. $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Ya que:
$\mathbf{P}$ también puede expresarse como $\mathbf{P} = \mathbf{K}\,[\mathbf{R} \,|\, -\mathbf{R}\bar{c}]$. Luego, $\mathbf{A} = \mathbf{K}\,\mathbf{R}$ y $\bar{b} = -\mathbf{A}\bar{c}$, con lo que $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Modelo proyectivo de cámara

7. Propiedades.



Sea la matriz de proyección $[\mathbf{A} \,|\, \bar{b}]$

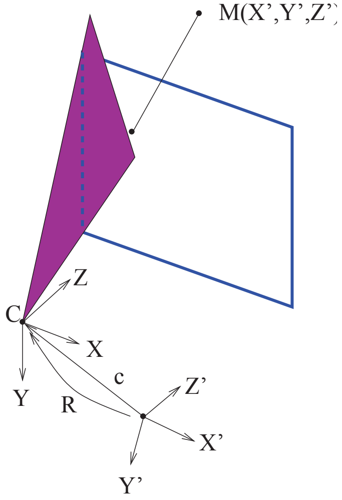
Centro Óptico. $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Filas de $\mathbf{P}$ Planos que pasan por $\bar{c}$ y:

$(\bar{p}^1)^\top$. contiene la línea $j = 0$.

Modelo proyectivo de cámara

7. Propiedades.



Sea la matriz de proyección $[\mathbf{A} \mid \bar{b}]$

Centro Óptico. $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

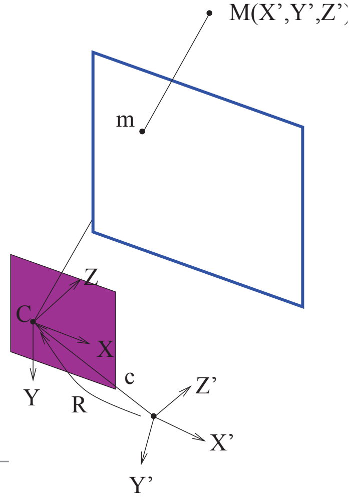
Filas de $\mathbf{P}$ Planos que pasan por $\bar{c}$ y:

$(\bar{p}^1)^\top$. contiene la línea $j = 0$.

$(\bar{p}^2)^\top$. contiene la línea $i = 0$.

Modelo proyectivo de cámara

7. Propiedades.



Sea la matriz de proyección $[\mathbf{A} \mid \bar{b}]$

Centro Óptico. $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Filas de $\mathbf{P}$ Planos que pasan por $\bar{c}$ y:

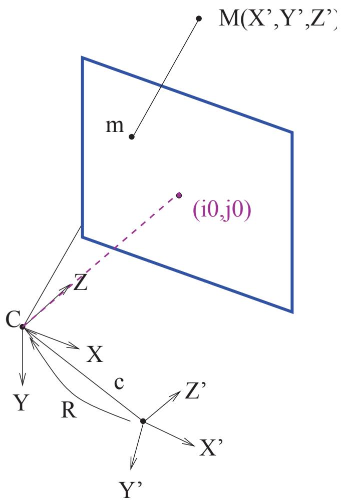
$(\bar{p}^1)^\top$. contiene la línea $j = 0$.

$(\bar{p}^2)^\top$. contiene la línea $i = 0$.

$(\bar{p}^3)^\top$. es paralelo al plano imagen.

Modelo proyectivo de cámara

7. Propiedades.



Sea la matriz de proyección $[\mathbf{A} \mid \bar{b}]$

Centro Óptico. $\bar{c} = -\mathbf{A}^{-1}\bar{b}$.

Filas de $\mathbf{P}$ Planos que pasan por $\bar{c}$ y:

$(\bar{p}^1)^\top$. contiene la línea $j = 0$.

$(\bar{p}^2)^\top$. contiene la línea $i = 0$.

$(\bar{p}^3)^\top$. es paralelo al plano imagen.

Punto principal. $\mathbf{A}\bar{a}^3$. Ya que el eje axial es $\lambda\bar{a}^3 - \mathbf{A}^{-1}\bar{b}$, y su imagen

$$[\mathbf{A} \mid \bar{b}] \begin{bmatrix} \lambda\bar{a}^3 - \mathbf{A}^{-1}\bar{b} \\ 1 \end{bmatrix} = \mathbf{A}\bar{a}^3.$$

Modelo afín de cámara

Si la matriz de proyección $\mathbf{P}_A$ es de la forma

$$\mathbf{P}_A = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} \alpha_u & s & 0 \\ 0 & \alpha_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} \bar{r}^{1\top} & t_1 \\ \bar{r}^{2\top} & t_2 \\ \hline \bar{0} & 1 \end{array}\right]$$

entonces decimos que $\mathbf{P}_A$ representa una cámara afín.

Tiene 8 parámetros:

.

Modelo afín de cámara

Si la matriz de proyección $\mathbf{P}_A$ es de la forma

$$\mathbf{P}_A = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} \alpha_u & s & 0 \\ 0 & \alpha_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} \bar{r}^{1\top} & t_1 \\ \bar{r}^{2\top} & t_2 \\ \hline \bar{0} & 1 \end{array}\right]$$

entonces decimos que $\mathbf{P}_A$ representa una cámara afín.

Tiene 8 parámetros:

- 3 parámetros intrínsecos . No está definido el punto principal.

$$\mathbf{K}_{2\times 2} = \begin{bmatrix} \alpha_u & s \\ 0 & \alpha_v \end{bmatrix}$$

Modelo afín de cámara

Si la matriz de proyección $\mathbf{P}_A$ es de la forma

$$\mathbf{P}_A = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} \alpha_u & s & 0 \\ 0 & \alpha_v & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} \bar{r}^{1\top} & t_1 \\ \bar{r}^{2\top} & t_2 \\ \hline \bar{0} & 1 \end{array}\right]$$

entonces decimos que $\mathbf{P}_A$ representa una cámara afín.

Tiene 8 parámetros:

- 3 parámetros intrínsecos .
- 5 parámetros extrínsecos.

$$\mathbf{R}_{2\times 3} = \begin{bmatrix} \bar{r}^{1\top} \\ \bar{r}^{2\top} \end{bmatrix}, \quad \bar{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.$$

Modelo afín de cámara

Propiedades:

- Es una buena aproximación a una cámara “real” cuando el relieve de la escena es pequeño en comparación a la distancia media de la escena a la cámara.

Modelo afín de cámara

Propiedades:

- Es una buena aproximación a una cámara “real” cuando el relieve de la escena es pequeño en comparación a la distancia media de la escena a la cámara.
- La proyección del centro de masas de una nube de puntos es el centro de masas de las proyecciones.

  En efecto, la proyección en cartesianas puede expresarse como $\bar{m}_{2\times 1} = \mathbf{A}_{2\times 3}\bar{M}_{3\times 1} + \bar{b}$, entonces basta comprobar que

$$\frac{1}{N}\sum_i (\mathbf{A}\bar{M}_i + \bar{b}) = \mathbf{A}\left(\frac{1}{N}\sum_i \bar{M}_i\right) + \bar{b}$$

Modelo afín de cámara

Propiedades:

- Es una buena aproximación a una cámara “real” cuando el relieve de la escena es pequeño en comparación a la distancia media de la escena a la cámara.
- La proyección del centro de masas de una nube de puntos es el centro de masas de las proyecciones.
- Conserva el paralelismo.

Modelo afín de cámara

Propiedades:

- Es una buena aproximación a una cámara “real” cuando el relieve de la escena es pequeño en comparación a la distancia media de la escena a la cámara.
- La proyección del centro de masas de una nube de puntos es el centro de masas de las proyecciones.
- Conserva el paralelismo.
- La dirección de proyección ortogonal es el vector $\bar{d}$ tal que $\mathbf{A}_{2\times 3}\bar{d} = \bar{0}$.

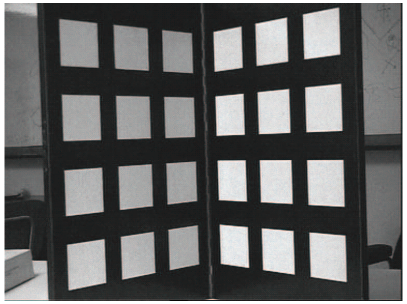
Modelo afín de cámara

Propiedades:

- Es una buena aproximación a una cámara “real” cuando el relieve de la escena es pequeño en comparación a la distancia media de la escena a la cámara.
- La proyección del centro de masas de una nube de puntos es el centro de masas de las proyecciones.
- Conserva el paralelismo.
- La dirección de proyección ortogonal es el vector $\bar{d}$ tal que $\mathbf{A}_{2\times 3}\bar{d} = \bar{0}$.
- *Caracterización.* Una matriz $\mathbf{P}_{2\times 4} = [\mathbf{A}_{2\times 3}\,|\,\bar{b}_{2\times 1}]$ representa una cámara afín $\iff$ el rango de $\mathbf{A}$ es 2.

Calibración

- ¿Qué es?
  Proceso en el que se calculan los parámetros de $\mathbf{P}$ para una cámara concreta.
- ¿Cómo?
  Proyectando un conjunto de puntos de posición conocida

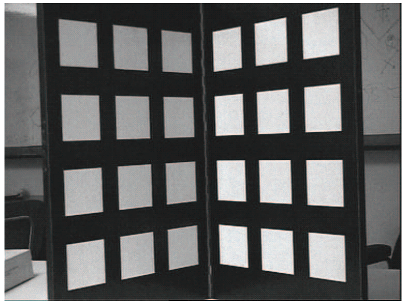


- Ecuaciones. Desarrollando:

$$\begin{bmatrix} \lambda j \\ \lambda i \\ \lambda \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Calibración

- ¿Qué es?
  Proceso en el que se calculan los parámetros de $\mathbf{P}$ para una cámara concreta.
- ¿Cómo?
  Proyectando un conjunto de puntos de posición conocida



- Ecuaciones.

$$p_{11}X + p_{12}Y + p_{13}Z - p_{31}jX - p_{32}jY - p_{33}jZ - p_{34}j + p_{14} = 0$$
$$p_{21}X + p_{22}Y + p_{23}Z - p_{31}iX - p_{32}iY - p_{33}iZ - p_{34}i + p_{24} = 0$$

Calibración

- Algoritmo:
  1. Extrae los bordes de la imagen (p.ej. Canny).
  2. Ajusta líneas rectas a los bordes extraidos.
  3. Calcula las esquinas de los cuadrados intersecando rectas.
  4. Dados un conjunto de $n \geq 6$ correspondencias $\{\bar{m}_i \leftrightarrow \bar{M}_i\}$, resuelve el sistema de ecuaciones:

$$\underbrace{\begin{bmatrix} \bar{M}_i^\top & \bar{0} & -j\bar{M}_i^\top \\ \bar{0} & \bar{M}_i^\top & -i\bar{M}_i^\top \end{bmatrix}}_{\mathbf{A}(\bar{M})_{2n\times 9}} \underbrace{\begin{bmatrix} \bar{p}^1 \\ \bar{p}^2 \\ \bar{p}^3 \end{bmatrix}}_{\bar{x}(\mathbf{P})_{12\times 1}} = \bar{0}_{2n\times 1}$$

Calibración

- Resolución del sistema de ecuaciones:
  1. **Solución lineal.** Calcula $\mathbf{P}$ mediante una solución lineal de $\mathbf{A}\bar{x} = \bar{0}$.
     a) **Normalización.** Normalizamos las coordenadas de los puntos 2D y 3D para mejorar el condicionamiento numérico:
     Sean $\{\tilde{m}_i = \mathbf{T}_{3\times 3}\bar{m}_i, i = 1\ldots n\}$ y $\{\tilde{M}_i = \mathbf{U}_{4\times 4}\bar{M}_i, i = 1\ldots n\}$ sendos conjuntos de datos centrados en el origen y a una distancia media del origen de $\sqrt{2}$ y $\sqrt{3}$ respectivamente.
     b) **Resuelve el sistema de ecuaciones** $\mathbf{A}(\tilde{M})\bar{x}(\tilde{\mathbf{P}}) = \bar{0}$, s.a. $||\bar{x}|| = 1$.
     Sea svd($\mathbf{A}$)=$\mathbf{UDV}^\top$, la solución $\bar{x}$ es la columna de $\mathbf{V}$ asociada al menor valor singular.
  2. **Minimiza el error de reproyección**. Sea $d(\cdot,\cdot)$ una distancia, calculamos

$$\min_{\tilde{\mathbf{P}}} \sum_i d(\tilde{x}_i, \tilde{\mathbf{P}}\tilde{M}_i),$$

     mediante un algoritmo iterativo (ej. Levenberg-Marquardt).
  3. **Desnormalización.** La matriz de proyección para los datos originales $\mathbf{P}$

$$\mathbf{P} = \mathbf{T}^{-1}\tilde{\mathbf{P}}\mathbf{U}.$$

Calibración

- Estimación de $\mathbf{K}$, $\mathbf{R}$ y $\bar{t}$.

  Conocida $\mathbf{P} = [\mathbf{A}\,|\,\bar{b}]$, mediante descomposición RQ de $\mathbf{A}$ (Hartley, 2004) se puede obtener $\mathbf{K}$ y $\mathbf{R}$.

  Conocida $\mathbf{K}$, $\bar{t} = \mathbf{K}^{-1}\bar{b}$.

Bibliografía

1. D. Forsyth, J. Ponce. “Computer Vision. A modern approach.” Prentice Hall. 2003.
2. R. Hartley, A. Zisserman. “Multiple view geometry in computer vision.” Cambridge University Press, 2004.