

Modelos Matemáticos en Visión por Ordenador

Modelos matemáticos para representar el funcionamiento de una cámara

Luis Alvarez

Univ. Las Palmas de G.C.

Mayo 2009

- 1 Modelo de proyección (pinhole model)
- 2 Modelo de distorsión de una lente
- 3 Calibración de cámaras
- 4 Bibliografía y problemas seleccionados

El uso de la perspectiva por los artistas del renacimiento



Figure: Ilustración de Albrecht Dürer, pintor alemán del Renacimiento (1527)

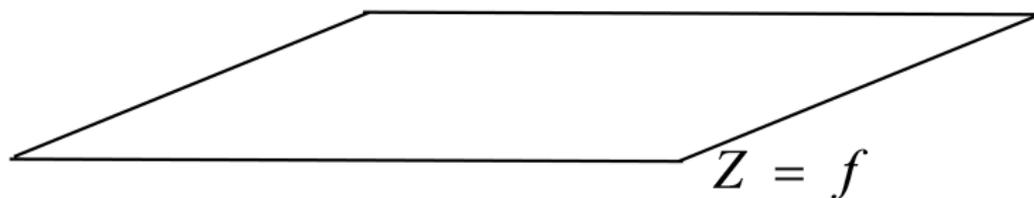


Gérard Desargues (1591-1661), arquitecto, ingeniero y matemático francés, nació en el seno de una familia bien acomodada al servicio de la corona, contemporáneo de Pascal y Descartes. Es el iniciador de la geometría proyectiva, pues fundamentó matemáticamente los métodos de la perspectiva que habían desarrollado los artistas del Renacimiento, y aunque su trabajo fue publicado en 1639, pasó desapercibido durante dos siglos

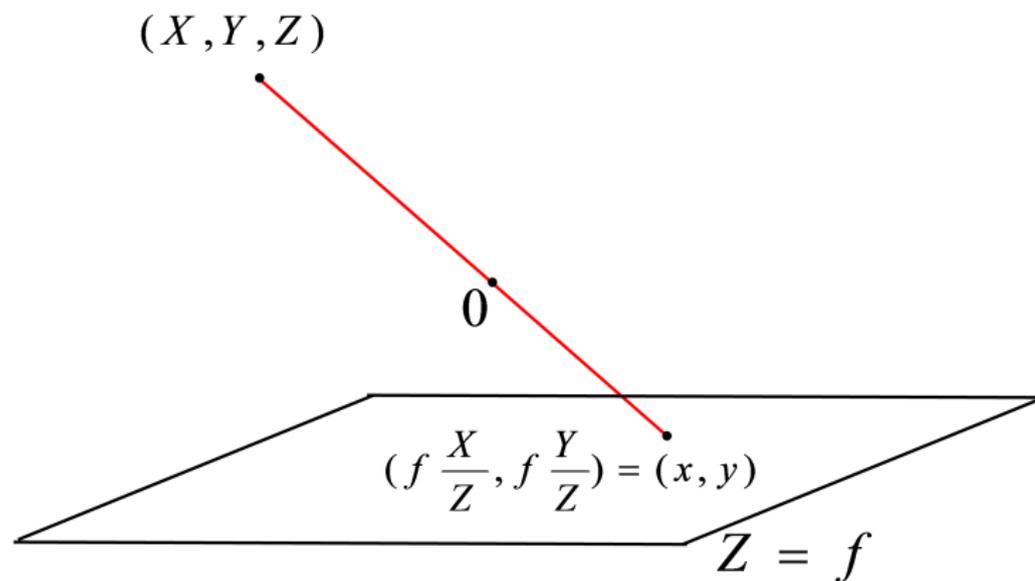
Modelo "pinhole" de proyección de un punto 3D.

(X, Y, Z)
•

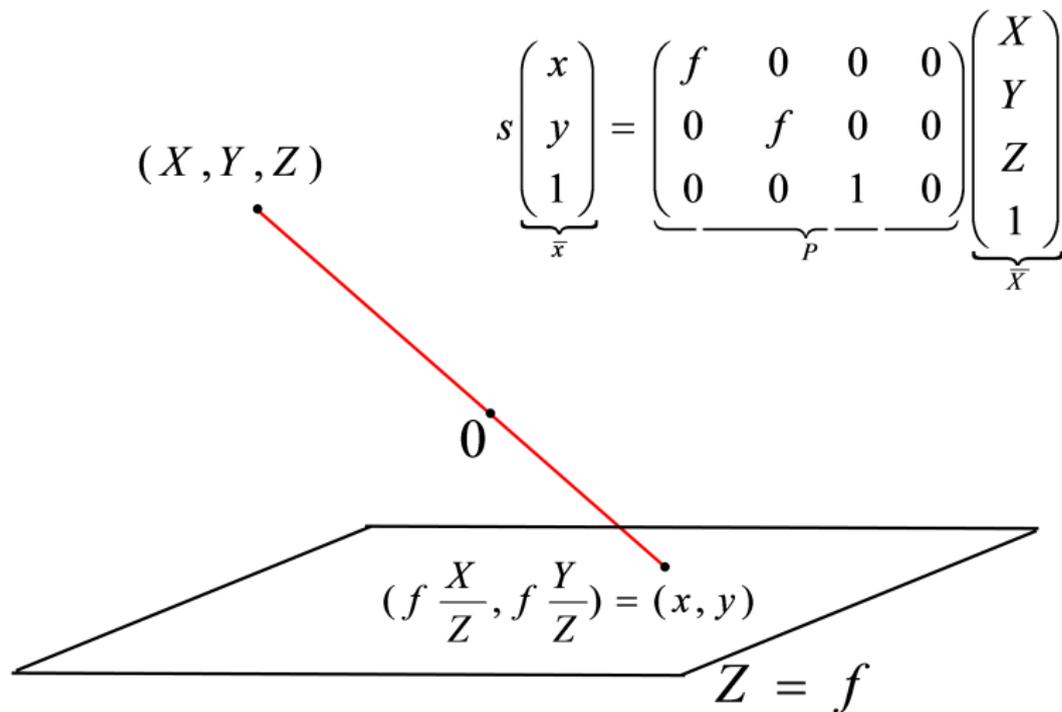
O



Modelo "pinhole" de proyección de un punto 3D.



Modelo "pinhole" de proyección de un punto 3D.



Modelo "pinhole" de proyección de un punto 3D

Ecuación general modelo proyección

$$s \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} f\alpha_x & \gamma & x_0 \\ 0 & f\alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Parámetros Intrínsecos.}} \underbrace{\begin{pmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{pmatrix}}_{\text{Parámetros extrínsecos}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Parámetros Intrínsecos.
Cambio coordenadas
en el Plano $Z = f$

Parámetros extrínsecos
Rotación y traslación en 3D

Modelo "pinhole" de proyección de un punto 3D

Ecuación general modelo proyección

$$s \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} f\alpha_x & \gamma & x_0 \\ 0 & f\alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Parámetros Intrínsecos.}} \underbrace{\begin{pmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{pmatrix}}_{\text{Parámetros extrínsecos}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Parámetros Intrínsecos.
Cambio coordenadas
en el Plano $Z = f$

Parámetros extrínsecos
Rotación y traslación en 3D

matriz de proyección

$$P = \begin{pmatrix} f\alpha_x & 0 & x_0 \\ 0 & f\alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{pmatrix}$$

Condiciones sobre la matriz P

Una matriz P se puede descomponer como

$$P = \begin{pmatrix} f\alpha_x & 0 & x_0 \\ 0 & f\alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{pmatrix}$$

si sólo sí se cumplen las condiciones

$$p_{20}^2 + p_{21}^2 + p_{22}^2 = 1$$

$$((p_{00}, p_{01}, p_{02}) \wedge (p_{20}, p_{21}, p_{22}))^T ((p_{10}, p_{11}, p_{12}) \wedge (p_{20}, p_{21}, p_{22})) = 0$$

Homografías entre planos

Un caso particularmente interesante en la práctica se produce cuando los puntos $3D$ se encuentran en un mismo plano, en ese caso, la matriz de proyección 3×4 P restringida al plano se convierte en una homografía H dada por una matriz 3×3 , de tal manera que la relación

$$s \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{pmatrix}$$

determina la proyección de los puntos del plano $3D$ en la cámara y donde (\tilde{x}, \tilde{y}) representa una parametrización del plano $3D$.

Extrayendo información de la matriz P

Si $\bar{C} = (C_x, C_y, C_z)$ es el foco de la cámara debe cumplirse que

$$P \begin{pmatrix} C_x \\ C_y \\ C_z \\ 1 \end{pmatrix} = \bar{0}$$

Extrayendo información de la matriz P

Si $\bar{C} = (C_x, C_y, C_z)$ es el foco de la cámara debe cumplirse que

$$P \begin{pmatrix} C_x \\ C_y \\ C_z \\ 1 \end{pmatrix} = \bar{0} \Rightarrow \bar{C} = - \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix}^{-1} \begin{pmatrix} p_{03} \\ p_{13} \\ p_{23} \end{pmatrix}$$

Extrayendo información de la matriz P

Si $\bar{C} = (C_x, C_y, C_z)$ es el foco de la cámara debe cumplirse que

$$P \begin{pmatrix} C_x \\ C_y \\ C_z \\ 1 \end{pmatrix} = \bar{0} \Rightarrow \bar{C} = - \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix}^{-1} \begin{pmatrix} p_{03} \\ p_{13} \\ p_{23} \end{pmatrix}$$

Dado un punto (x, y) en el plano de proyección, el vector director $\bar{u} = (u_x, u_y, u_z)$ de la recta 3D que une el punto con el foco debe verificar

$$P \begin{pmatrix} u_x \\ u_y \\ u_z \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Extrayendo información de la matriz P

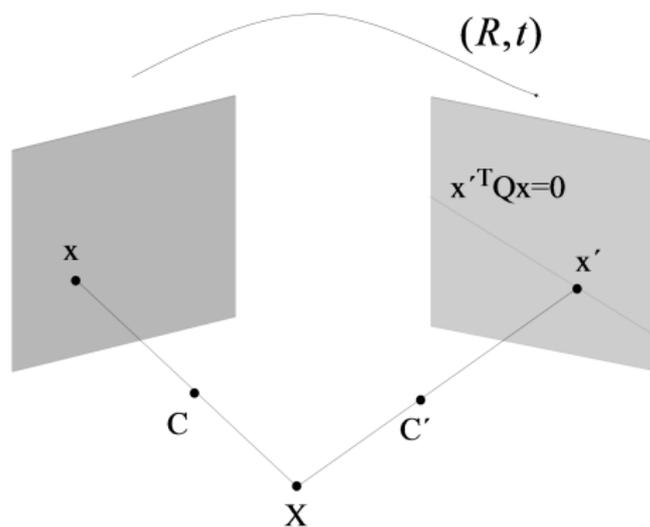
Si $\bar{C} = (C_x, C_y, C_z)$ es el foco de la cámara debe cumplirse que

$$P \begin{pmatrix} C_x \\ C_y \\ C_z \\ 1 \end{pmatrix} = \bar{0} \Rightarrow \bar{C} = - \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix}^{-1} \begin{pmatrix} p_{03} \\ p_{13} \\ p_{23} \end{pmatrix}$$

Dado un punto (x, y) en el plano de proyección, el vector director $\bar{u} = (u_x, u_y, u_z)$ de la recta 3D que une el punto con el foco debe verificar

$$P \begin{pmatrix} u_x \\ u_y \\ u_z \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Visión estéreo. Geometría epipolar



$$Q = \begin{pmatrix} r_{13}t_y - r_{12}t_z & r_{11}t_z - r_{13}t_x & r_{12}t_x - r_{11}t_y \\ r_{23}t_y - r_{22}t_z & r_{21}t_z - r_{23}t_x & r_{22}t_x - r_{21}t_y \\ r_{33}t_y - r_{32}t_z & r_{31}t_z - r_{33}t_x & r_{32}t_x - r_{31}t_y \end{pmatrix}$$

Condiciones que debe cumplir la matriz Q

Una matriz 3×3 se puede descomponer como

$$Q = \begin{pmatrix} r_{02}t_y - r_{01}t_z & r_{00}t_z - r_{02}t_x & r_{01}t_x - r_{00}t_y \\ r_{12}t_y - r_{11}t_z & r_{10}t_z - r_{12}t_x & r_{11}t_x - r_{10}t_y \\ r_{22}t_y - r_{21}t_z & r_{20}t_z - r_{22}t_x & r_{21}t_x - r_{20}t_y \end{pmatrix}$$

si sólo si

$$|Q| = 0$$

$$\frac{1}{2} \left(\text{traza} \left(Q^T Q \right) \right)^2 - \text{traza} \left(\left(Q^T Q \right)^2 \right) = 0$$

Un enfoque geométrico de la visión por ordenador

Olivier Faugeras



Olivier Faugeras (1949), científico francés, alumno de l'Ecole Polytechnique (Cauchy, Poisson, Poincaré), miembro de la Academia de Ciencias Francesa, investigador de gran prestigio en el ámbito de la visión artificial, ha realizado importantes contribuciones en el área, fué uno de los primeros investigadores de prestigio en visión que se interesó por las aplicaciones de las EDP's a la visión.

- 1 Modelo de proyección (pinhole model)
- 2 Modelo de distorsión de una lente
- 3 Calibración de cámaras
- 4 Bibliografía y problemas seleccionados

La difracción de la luz

David Brewster



David Brewster (1781-1868), científico escocés, hijo de un profesor. Aunque realizó estudios teológicos, finalmente se dedicó a la ciencia y realizó importantes contribuciones en el estudio de la difracción de la luz. Propuso y analizó el uso de lentes para desviar los rayos de luz y enfocarlos hacia una zona común. Este es el principio en que se basan las actuales cámaras fotográficas.

Efecto de la distorsión de una lente y su corrección



Modelo matemático de distorsión de una lente

La manera habitual de modelizar la distorsión de una lente consiste en suponer que la magnitud de la distorsión es una función radial, de tal manera que dado un punto (x, y) en la imagen, el correspondiente punto (\tilde{x}, \tilde{y}) con la distorsión corregida viene dado por

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} x_c \\ y_c \end{pmatrix} + L(r^2) \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix}$$

donde (x_c, y_c) es el centro de distorsión, $r^2 = (x - x_c)^2 + (y - y_c)^2$ y $L(r^2) = 1 + d_1 r^2 + d_2 r^4 + \dots + d_{N_d} r^{2N_d}$. Este modelo es sólo una aproximación de la realidad pues no tiene en cuenta que la distorsión cambia también con la profundidad a la que se encuentra el objeto. Estimar el modelo de distorsión consiste en estimar los coeficientes $\{d_i\}$ y el centro de distorsión (x_c, y_c) .

Modelo de proyección de una cámara teniendo en cuenta el modelo de distorsión

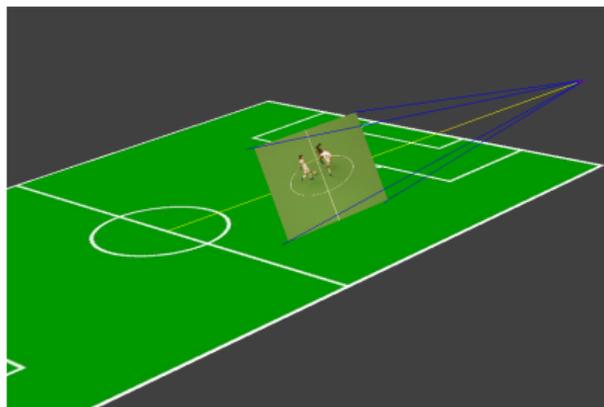
La distorsión introducida por la lente, aleja el modelo de proyección de una cámara del modelo "pinhole" de proyección 3D de un punto. Para restablecer la validez de este modelo es necesario corregir la distorsión de la lente, de tal manera que el modelo de proyección de un punto en la cámara viene dado por :

$$s \left(\begin{pmatrix} x_c \\ y_c \\ 1 \end{pmatrix} + L(r^2) \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix} \right) = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

- 1 Modelo de proyección (pinhole model)
- 2 Modelo de distorsión de una lente
- 3 Calibración de cámaras**
- 4 Bibliografía y problemas seleccionados

Calibración de una cámara.

Donde estoy y hacia donde miro ?



Calibración de cámaras

Algunos escenarios de calibración habituales

- 1 Conocemos la posición 3D de un conjunto de puntos en la escena

Calibración de cámaras

Algunos escenarios de calibración habituales

- 1 Conocemos la posición 3D de un conjunto de puntos en la escena
- 2 Conocemos la posición de un conjunto de puntos situados en un plano en la escena

Calibración de cámaras

Algunos escenarios de calibración habituales

- 1 Conocemos la posición 3D de un conjunto de puntos en la escena
- 2 Conocemos la posición de un conjunto de puntos situados en un plano en la escena
- 3 Conocemos 2 vistas distintas de la misma escena donde son visibles un conjunto de puntos desde las dos cámaras.

Calibración de cámaras

Algunos escenarios de calibración habituales

- 1 Conocemos la posición 3D de un conjunto de puntos en la escena
- 2 Conocemos la posición de un conjunto de puntos situados en un plano en la escena
- 3 Conocemos 2 vistas distintas de la misma escena donde son visibles un conjunto de puntos desde las dos cámaras.

Calibración de cámaras

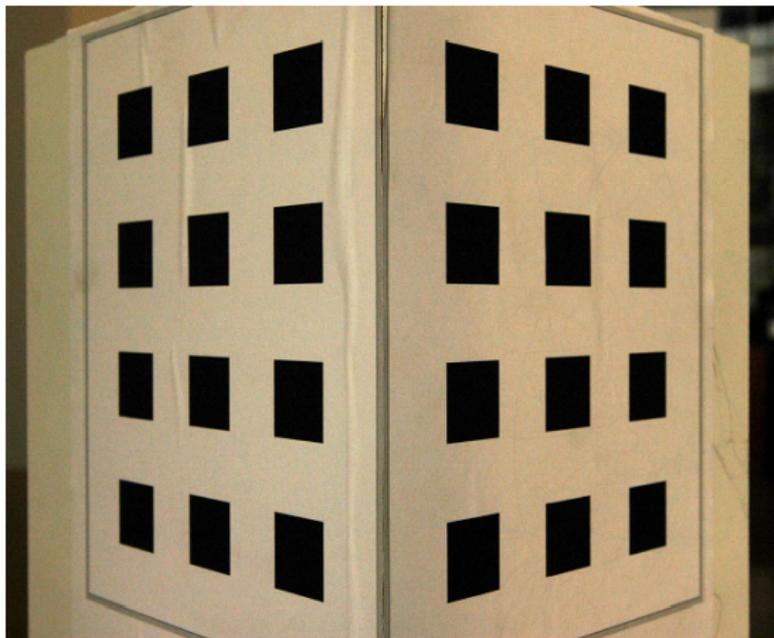
Algunos escenarios de calibración habituales

- 1 Conocemos la posición 3D de un conjunto de puntos en la escena
- 2 Conocemos la posición de un conjunto de puntos situados en un plano en la escena
- 3 Conocemos 2 vistas distintas de la misma escena donde son visibles un conjunto de puntos desde las dos cámaras.

Cálculo modelo distorsión: En el caso en que sean visibles en la escena rectas distorsionadas por la lente, es posible calcular el modelo de distorsión de la lente a partir de la rectificación de estas rectas distorsionadas. Supondremos en lo que sigue que la distorsión de la lente ha sido corregida.

Calibración de cámaras

Conocemos la posición 3D de un conjunto de puntos en la escena. Uso de un calibrador 3D



Calibración de cámaras

Conocemos la posición 3D de un conjunto de puntos en la escena.

por cada punto 3D \bar{X}_i del cual conocemos su proyección en la imagen \bar{x}_i , se cumple $s_i \bar{x}_i = P \bar{X}_i$, si desarrollamos esta igualdad obtenemos

$$s_i x_i = p_{03} + p_{00} X_i + p_{01} Y_i + p_{02} Z_i$$

$$s_i y_i = p_{13} + p_{10} X_i + p_{11} Y_i + p_{12} Z_i$$

$$s_i = p_{23} + p_{20} X_i + p_{21} Y_i + p_{22} Z_i$$

si eliminamos s_i del sistema llegamos a las siguientes ecuaciones lineales en los coeficientes de P

$$p_{03} + p_{00} X_i + p_{01} Y_i + p_{02} Z_i - (p_{23} + p_{20} X_i + p_{21} Y_i + p_{22} Z_i) x_i = 0$$

$$p_{13} + p_{10} X_i + p_{11} Y_i + p_{12} Z_i - (p_{23} + p_{20} X_i + p_{21} Y_i + p_{22} Z_i) y_i = 0$$

Calibración de cámaras

Conocemos la posición 3D de un conjunto de puntos en la escena.

Coleccionando la información de todos los puntos obtenemos un sistema de la forma

$$A\bar{p} = 0$$

donde \bar{p} es el vector formado por todos los coeficientes de la matriz P . La manera habitual de calcular \bar{p} es minimizando

$$E(\bar{p}) = \|A\bar{p}\|^2$$

Calibración de cámaras

Conocemos la posición 3D de un conjunto de puntos en la escena.

Coleccionando la información de todos los puntos obtenemos un sistema de la forma

$$A\bar{p} = 0$$

donde \bar{p} es el vector formado por todos los coeficientes de la matriz P . La manera habitual de calcular \bar{p} es minimizando

$$E(\bar{p}) = \|A\bar{p}\|^2 = \bar{p}^T A^T A \bar{p}$$

Calibración de cámaras

Conocemos la posición 3D de un conjunto de puntos en la escena.

Coleccionando la información de todos los puntos obtenemos un sistema de la forma

$$A\bar{p} = 0$$

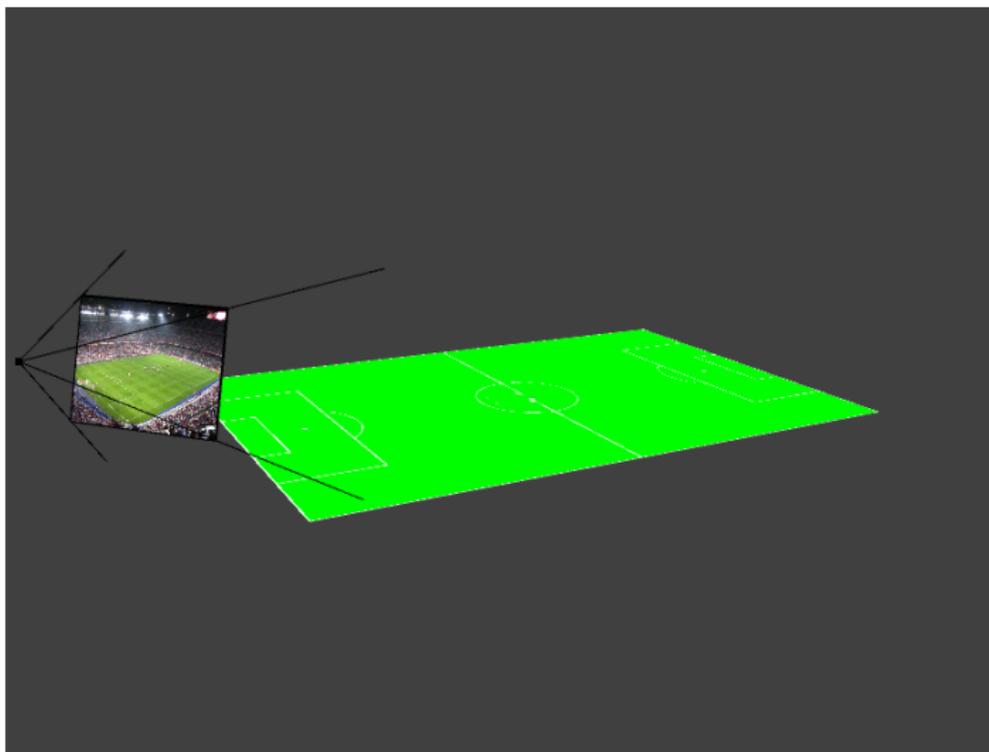
donde \bar{p} es el vector formado por todos los coeficientes de la matriz P . La manera habitual de calcular \bar{p} es minimizando

$$E(\bar{p}) = \|A\bar{p}\|^2 = \bar{p}^T A^T A \bar{p}$$

como \bar{p} está definido módulo la multiplicación por una constante, podemos suponer $\|\bar{p}\| = 1$, en cuyo caso el mínimo de $E(\bar{p})$ es el autovector de autovalor mínimo de la matriz $A^T A$.

Calibración de cámaras

Conocemos información sobre un plano en la escena.



Calibración de cámaras

Conocemos información sobre un plano en la escena.

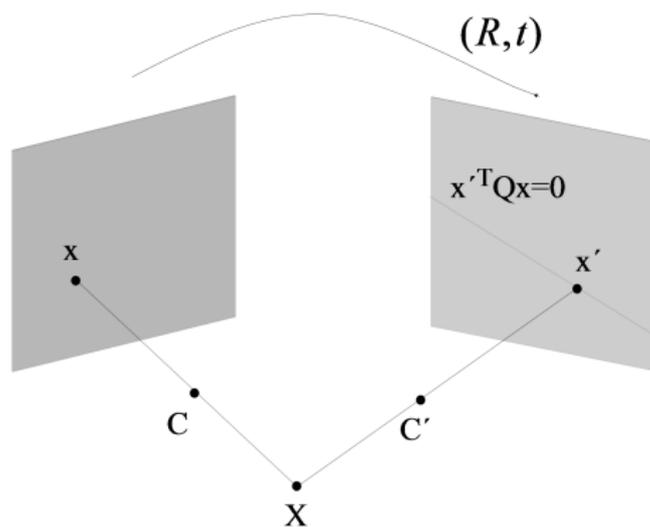
Para cada punto \tilde{x}_i del plano en la escena para el cual conocemos su proyección \bar{x}_i tenemos la relación $s_i \bar{x}_i = H \tilde{x}_i$. A partir de un número suficiente de relaciones de este tipo es posible, como en el cálculo de P , calcular la homografía H . Una vez calculada H , si conocemos los parámetros intrínsecos de la cámara es posible calcular los parámetros extrínsecos (R y t). Para ello se sitúa el sistema de coordenadas de tal forma que el plano en la escena corresponda a $Z = 0$, y explotando la relación algebraica

$$s \begin{pmatrix} f\alpha_x & 0 & x_0 \\ 0 & f\alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} H = \begin{pmatrix} r_{00} & r_{01} & t_x \\ r_{10} & r_{11} & t_y \\ r_{20} & r_{21} & t_z \end{pmatrix}$$

Disponemos de 2 vistas de la escena



Disponemos de 2 vistas de la escena



$$Q = \begin{pmatrix} r_{13}t_y - r_{12}t_z & r_{11}t_z - r_{13}t_x & r_{12}t_x - r_{11}t_y \\ r_{23}t_y - r_{22}t_z & r_{21}t_z - r_{23}t_x & r_{22}t_x - r_{21}t_y \\ r_{33}t_y - r_{32}t_z & r_{31}t_z - r_{33}t_x & r_{32}t_x - r_{31}t_y \end{pmatrix}$$

Disponemos de 2 vistas de la escena

Para cada par de puntos en correspondencia entre las 2 vistas tenemos la relación lineal en los coeficientes de $\bar{x}_i^T Q \bar{x}'_i = 0$. A partir de un número suficiente de relaciones de este tipo es posible, como en el cálculo de P , calcular la matriz Q . Una vez calculada Q , es posible calcular los parámetros extrínsecos (R y t). Para ello se explota la relación algebraica

$$\begin{pmatrix} r_{02}t_y - r_{01}t_z & r_{00}t_z - r_{02}t_x & r_{01}t_x - r_{00}t_y \\ r_{12}t_y - r_{11}t_z & r_{10}t_z - r_{12}t_x & r_{11}t_x - r_{10}t_y \\ r_{22}t_y - r_{21}t_z & r_{20}t_z - r_{22}t_x & r_{21}t_x - r_{20}t_y \end{pmatrix} = sQ$$

Video Reconstrucción 3D

- 1 Modelo de proyección (pinhole model)
- 2 Modelo de distorsión de una lente
- 3 Calibración de cámaras
- 4 Bibliografía y problemas seleccionados**

Bibliografía seleccionada

- O. Faugeras. Three-Dimensional Computer Vision. MIT Press, 1993.
- R. I. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2004.

- 1 Dado un plano parametrizado en $3D$ como $\tilde{X} = \bar{p} + \tilde{x}\bar{u} + \tilde{y}\bar{v}$ (con $\bar{p}, \bar{u}, \bar{v} \in \mathbb{R}^3$, $\tilde{x}, \tilde{y} \in \mathbb{R}$). Calcular, utilizando la matriz de proyección P , la expresión de homografía H que proyecta dicho plano en el plano de proyección.
- 2 Demostrar que el mínimo de $E(\bar{p}) = \|A\bar{p}\|^2$ con la restricción $\|\bar{p}\| = 1$ es el autovector de autovalor mínimo de $A^T A$.
- 3 Dadas dos cámaras de matrices de proyección P_1 y P_2 y las proyecciones de un punto \tilde{X} en ambas cámaras dadas por \bar{x}_1 y \bar{x}_2 . Deducir como podríamos recuperar las coordenadas $3D$ de \tilde{X} a partir de la información suministrada por P_1 , P_2 , \bar{x}_1 y \bar{x}_2 .