

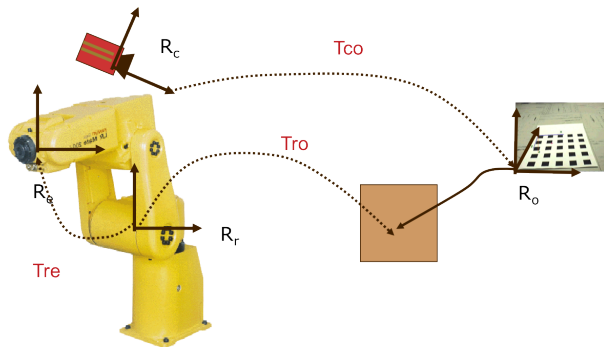
Malik Mallem

Professeur UEVE

27 septembre 2017

Sommaire

Contexte



- R_c : repère de la caméra
- R_r : repère du robot
- R_e : repère effecteur
- R_o : repère environnement

Notations Mathématiques

Notations de base :

Scalars, vectors and points $\in \mathbb{R}^2$: in lowercase

Matrices and points $\in \mathbb{R}^3$: in uppercase

M : Matrice

v : Vecteur

$x \cdot y$: Produit scalaire entre les vecteurs x et y

$x \times y$: Produit vectoriel entre les vecteurs x et y

Notations Mathématiques

Géométrie euclidienne et projective :

\sim : Egalité à un facteur d'échelle près

p : Point 2D ;

\tilde{p} : p en coordonnées homogènes

P : Point 3D

\tilde{P} : P en coordonnées homogènes

d : Droite de l'espace

π : Plan de l'espace

R : Matrice de rotation

t : Vecteur de translation



Modélisation et calibration de capteurs/robot

- Objectif

Selon type de capteurs :

- Caméra → Superposition d'images
- Dispositif 3D → Relevé 3D
- Robot → Relevé 3D, Calibration automatique

- Pré conditions

Indices d'apprentissage :

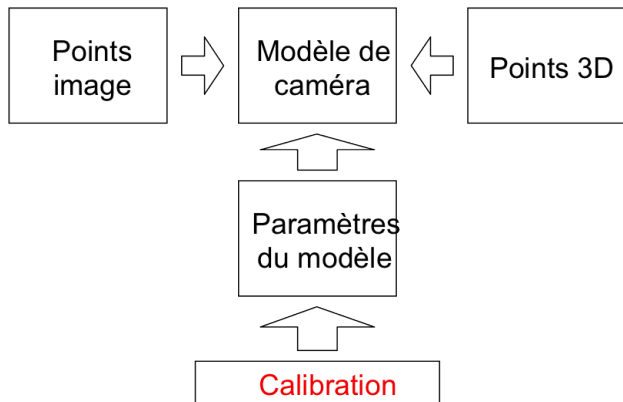
Points, Lignes, ...

- Post conditions

Paramètres du modèle

Physique, Géométrique

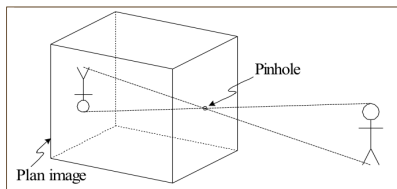
Modélisation et calibration de caméra



Objectif

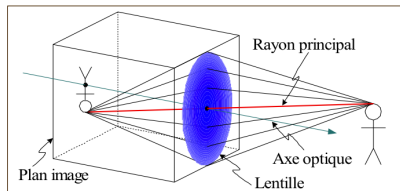
Déterminer la position d'un objet dans l'espace à partir de sa position dans l'image.

Formation des images



Sténopé (modèle idéal)

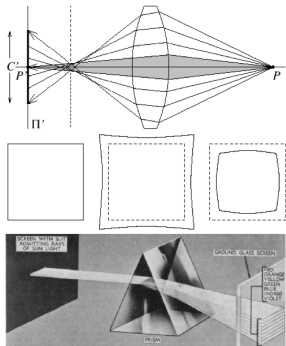
Problème : quantité de lumière
insuffisante



Caméra réelle

Solution : utilisation d'une lentille

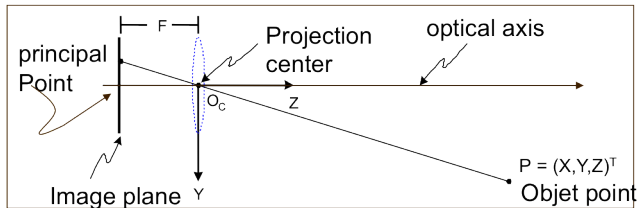
Formation des images



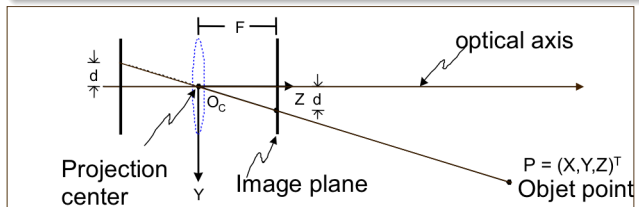
Problèmes liés à l'utilisation d'une lentille :

Aberrations sphériques et chromatiques, Distorsions radiales (Coussinet, Barillet),

Modèle du sténopé



Modèle inverseur



Modèle non inverseur

Modélisation géométrique de caméra

Hypothèses :

- Modèle linéaire à sténopé ("Pinhole"),
- Transformation perspective parfaite de centre O_c (centre optique de la caméra),
- La distance $O_c O'$ est la distance focale f de la caméra,
- L'axe optique $O_c Z_c$ intersecte le plan image en O' (centre de l'image),
- Les axes X_c et Y_c sont parallèles au plan image.

Modélisation d'une caméra, modèle du sténopé

$$\tilde{p} = T_c^e \cdot T_o^c \cdot \tilde{P} \quad (1)$$

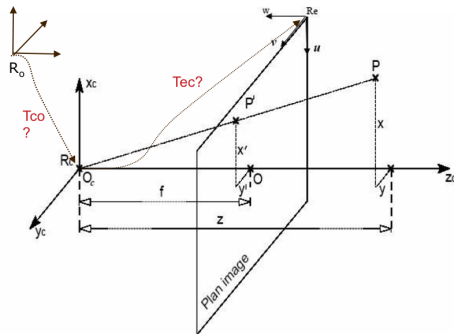
Où :

- (R_e) : repère écran ;
- (R_c) : repère caméra ;
- (R_o) : repère monde ;
- \tilde{p} : point image ;
- T_c^e : modèle interne ;
- T_o^c : modèle externe ;
- \tilde{P} : point 3D ;

avec :

$$\tilde{p} = \begin{pmatrix} u & v & 1 \end{pmatrix}^t;$$

$$\tilde{P} = \begin{pmatrix} X & Y & Z & 1 \end{pmatrix}^t;$$

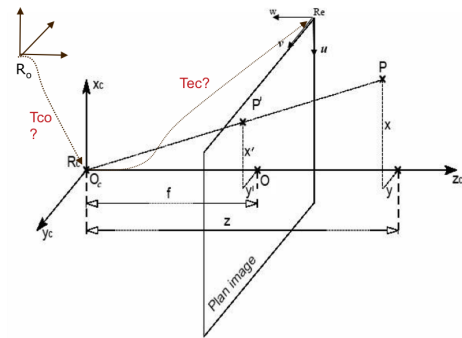


Modèle du sténopé, modèle interne

$$(1) \iff = (T_{hr} \cdot T_p) \cdot T_o^c \cdot \tilde{P}$$

Où :

- T_p : Projection perspective du point P en P' exprimé dans (R_c) ;
- T_{hr} : Transformation de (R_c) dans (R_e) (Rotation + homothétie) ;



Modèle du sténopé, modèle interne

1 : T_p , Projection perspective :

- T_p est exprimée dans (R_c) , Elle projette P en P' sur le plan image ;

Les coordonnées de P' sont :

$$\tilde{P}' = \begin{pmatrix} X' & Y' & Z' & 1 \end{pmatrix}^t$$

- On a : $X/Z = X'/f$; $X' = f.X/Z$;
- de même en Y : $Y/Z = Y'/f$; $Y' = f.Y/Z$;
- avec $Z' = f$;
 f : focale de la caméra.

Modele interne, projection perspective

1 : T_p , Projection perspective, représentation matricielle

En utilisant les coordonnées homogènes, on représente la transformation sous forme matricielle :

$$\begin{pmatrix} s \cdot X' \\ s \cdot Y' \\ s \cdot Z' \\ s \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2)$$

Modele interne, projection perspective

1 : T_p , Projection perspective, application

D'où les coordonnées de P' :

$$(2) \iff \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} f \cdot X/Z \\ f \cdot Y/Z \\ f \end{pmatrix} \text{ avec } (s = Z/f)$$

s : facteur d'échelle.

Modele interne, rotation et mise à l'échelle

2 : T_{hr} , Transformation de Rotation et de Mise à l'échelle du plan image
(m \rightarrow pixel)

$$p = T_{hr} \cdot P' + t \iff \begin{pmatrix} u \\ v \\ w \end{pmatrix} = T_{hr} \cdot \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} \iff$$

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & 0 \\ 0 & k_v & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} \quad (3)$$

Modele interne, , transformations $T_{hr}; T_p$

$$(3) \iff \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -k_u & 0 & u_0 & 0 \\ 0 & k_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} \quad (4)$$

avec :

Les paramètres k_u, k_v, f, u_0, v_0 : paramètres intrinsèques de la caméra.

Où k_u et k_v représentent les distances focales exprimées en pixels suivant l'axe u et l'axe v respectivement ;

α_u et α_v : taille focale en pixels selon abscisse et ordonnée ;

$1/k_u$ et $1/k_v$: dimensions en abscisse et en ordonnée du pixel (m/pixel) ;

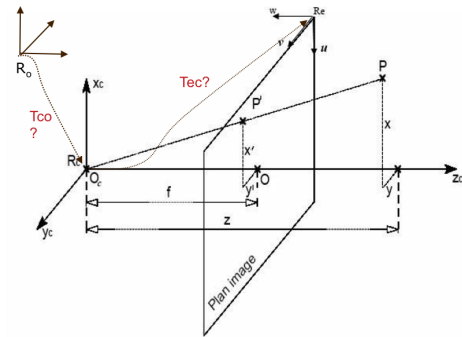
$(u_0, v_0)^t$, coordonnées pixels de la projection de O_c dans le plan image.

Modèle du sténopé, modèle externe

$$(1) \iff \tilde{p} = T_C^e \cdot T_C^c \cdot \tilde{P} \\ = T_C^e \cdot (R \cdot T) \cdot \tilde{P}$$

Où :

- T_O^e : Modèle externe permettant de passer de (R_O) à (R_C) ;
- R : matrice de Rotation ;
- T : matrice de translation ;



Modèle du sténopé, modèle externe

$$\tilde{p} = T_c^e \cdot (R \cdot T) \cdot \tilde{P}$$

$$\Leftrightarrow \begin{pmatrix} s \cdot u \\ s \cdot v \\ s \end{pmatrix} = T_c^e \cdot \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (5)$$

Modèle du sténopé, modèle global

Le modèle global est alors :

$$(5) \iff \begin{pmatrix} s \cdot u \\ s \cdot v \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (6)$$

avec :

$$\alpha_u = -k_u \cdot f \text{ et } \alpha_v = k_v \cdot f$$

Modèle du sténopé, modèle global

Dans les faits, $\gamma \approx 0$ (cf. sujet TD1) :

$$\Leftrightarrow \begin{pmatrix} s \cdot u \\ s \cdot v \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

avec :

$$\alpha_u = -k_u \cdot f \text{ et } \alpha_v = k_v \cdot f$$

Modèle du sténopé, application du modèle externe

Soit un point P de coordonnées $(X, Y, Z)^t$ exprimées dans le repère monde. Ses coordonnées dans le repère caméra $(X_c, Y_c, Z_c)^t$ sont obtenues à partir de la relation suivante :

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = R \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + t = [R|t] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Modèle du sténopé, application du modèle externe

Où $(X, Y, Z, 1)^t$ sont les coordonnées homogènes du point P dans le repère monde. R et t représentent le déplacement rigide entre les deux repères, R étant la matrice de rotation et t le vecteur de translation. Ces paramètres définissent la position et l'orientation de la caméra par rapport au repère monde. Ils représentent les paramètres extrinsèques de la caméra appelés **pose de la caméra**.

Modèle du sténopé, application du modèle externe

Un point P se projette en un point p sur le plan \mathcal{P} selon une projection perspective du centre O_c , origine du repère caméra. Autrement dit le point p correspond à l'intersection de la droite (O_cP) avec le plan image \mathcal{P} . La projection monde/image se décompose donc en :

- Une projection du point P dans le repère associé à la caméra selon les paramètres extrinsèques de la caméra.
- Le point résultant de la transformation monde-caméra se projette sur le plan image en fonction des paramètres intrinsèques de la caméra.

Modèle du sténopé, application du modèle externe

Soit R la matrice de rotation permettant de transformer les axes de (R_o) en ceux de (R_c) . Soit t la translation permettant de déplacer l'origine de (R_o) en celle de (R_c) , le centre optique de la caméra. On suppose que R est appliquée préalablement à t . Notons T_o^c la matrice réalisant les deux transformations :

$$T_o^c = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$

La matrice T_o^c ainsi définie est la matrice de passage de la base (R_o) à la base (R_c) . La transformation que subit X^o lors de son expression dans (R_c) est donc

$$X^o = (T_o^c)^{-1} \cdot X^c = \begin{pmatrix} R^t & -R^t \cdot t \\ 0 & 1 \end{pmatrix} \cdot X^c$$

Modèle du sténopé, application du modèle externe

Dans la suite, pour des raisons de commodité, nous représenterons la pose d'une caméra par la matrice T_o^c permettant de transformer X^o en X^c , soit T^{-1} telle qu'explicitée dans 24 :

$$T_o^c = \begin{pmatrix} R_o^c & t_o^c \\ 0 & 1 \end{pmatrix}$$

On notera généralement (R_c) pour désigner T_o^c .

Modèle du sténopé, application du modèle global

Le processus complet de la formation d'image est :

$$(1) \iff \tilde{p} = T_e^c \cdot T_o^c \cdot \tilde{P} \\ = T_e^c \cdot (R \cdot T) \cdot \tilde{P}$$

$$s \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \cdot T_o^c \cdot P^o$$

Modèle du sténopé, application du modèle interne

La projection du point P , de coordonnées $(X_c, Y_c, Z_c)^T$ définies dans le repère caméra, est le point P' de coordonnées $(X'_c, Y'_c, Z'_c)^t$ définies dans le repère caméra, telle que :

$$\begin{pmatrix} X'_c \\ Y'_c \\ Z'_c \end{pmatrix} = f \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (7)$$

Modèle du sténopé, application du modèle interne

Dans le repère image $2D$, le point p est exprimé en coordonnées pixel $(u, v)^t$ obtenues grâce à la relation suivante où $\gamma \approx 0$ (cf. TD1) :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} k_u & -k_u \cos \theta & u_0 \\ 0 & k_v \sin \theta & v_0 \end{pmatrix} \begin{pmatrix} X'_c \\ Y'_c \\ 1 \end{pmatrix} \quad (8)$$

Où k_u et k_v représentent les distances focales exprimées en pixels suivant l'axe u et l'axe v respectivement ; $(u_0, v_0)^t$ sont les coordonnées pixel du point centre c ; θ est l'angle entre les deux axes (u et v) dit aussi angle de distorsion. Les paramètres $k_u, k_v, f, u_0, v_0, \theta$ sont les paramètres intrinsèques de la caméra.

Modèle du sténopé, application du modèle interne

De l'équation (7) et de l'équation (8), nous obtenons la relation suivante :

$$s \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (9)$$

Modèle du sténopé, application du modèle interne

Avec $\alpha_u = -f \cdot K_u$ et $\alpha_v = f \cdot k_v$ et θ généralement de l'ordre de $\frac{\pi}{2}$ (cf. TD1). Posons K la matrice des paramètres intrinsèques. Au final, la relation qui relie le point P de coordonnées homogènes $(X, Y, Z, 1)^T$ dans le repère monde et le point p de coordonnées pixel homogènes $(u, v, 1)^t$ est donnée par :

$$s \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K[R|t] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (10)$$

La matrice $K[R|t]$ est appelée matrice de projection perspective. De nombreuses méthodes ont été développées par la communauté de la vision par ordinateur pour le calcul de ces paramètres (méthodes de calibration de caméra). Les paramètres intrinsèques sont généralement supposés connus dans le problème d'estimation de pose.

Modèle du sténopé, application du modèle global

Soit P un point 3D dont les coordonnées homogènes $(X, Y, Z, 1)^t$ sont définies dans le repère monde. Ce point se projette sur le plan image en le point p de coordonnées homogènes $(u, v, 1)^t$. Soit M la matrice de projection, telle que :

(10)

$$\Leftrightarrow \begin{pmatrix} s \cdot u \\ s \cdot v \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (11)$$

Modèle du sténopé, application du modèle global

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} =$$

$$\begin{pmatrix} \alpha_u \cdot r_{11} + u_0 \cdot r_{31} & \alpha_u \cdot r_{12} + u_0 \cdot r_{32} & \alpha_u \cdot r_{13} + u_0 \cdot r_{33} & \alpha_u \cdot t_x + u_0 \cdot t_z \\ \alpha_v \cdot r_{21} + v_0 \cdot r_{31} & \alpha_v \cdot r_{22} + v_0 \cdot r_{32} & \alpha_v \cdot r_{23} + v_0 \cdot r_{33} & \alpha_v \cdot t_y + v_0 \cdot t_z \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix}$$

Modèle du sténopé, application du modèle global

$$M = \begin{pmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{pmatrix} \stackrel{\iff}{=} \begin{pmatrix} \alpha_u \cdot r_1 + u_0 \cdot r_3 & \alpha_u \cdot t_x + u_0 \cdot t_z \\ \alpha_v \cdot r_2 + v_0 \cdot r_3 & \alpha_v \cdot t_y + v_0 \cdot t_z \\ r_3 & t_z \end{pmatrix} \quad (12)$$

En posant :

$$m_i = [m_{i1} m_{i2} m_{i3}]$$

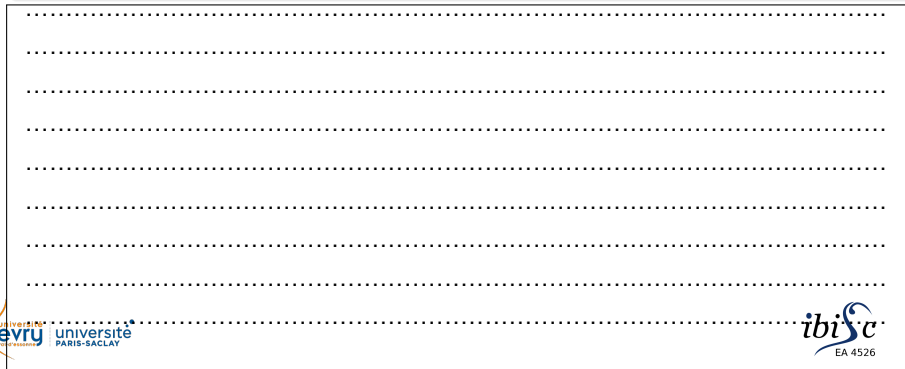
et :

$$r_i = [r_{i1} r_{i2} r_{i3}]$$

Modèle du sténopé, calcul des paramètres intrinsèques et extrinsèques

Exercice :

Le calcul des paramètres intrinsèques et extrinsèques s'obtient en identifiant le système (12). Les m_{ij} sont supposés connus.



(11) équivaut à :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} u = (P^t \cdot m_1 + m_{14}) / (P^t \cdot m_3 + m_{34}) \\ v = (P^t \cdot m_2 + m_{24}) / (P^t \cdot m_3 + m_{34}) \end{cases} \Leftrightarrow$$

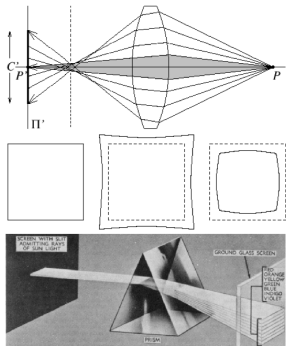
$$\begin{cases} P^t \cdot n_u + a_u = 0 \\ P^t \cdot n_v + a_v = 0 \end{cases}$$

avec :

$$n_u = m_1 - u \cdot m_3; n_v = m_2 - v \cdot m_3; a_u = m_{14} - u \cdot m_{34}; a_v = m_{24} - v \cdot m_{34};$$

Il s'agit de l'équation du rayon visuel porteur de P et de son image p et passant par le centre de la caméra.

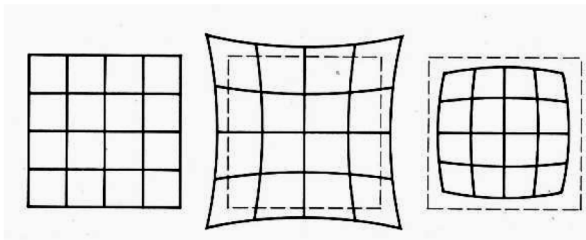
Distorsion de l'image



Problèmes liés à l'utilisation d'une lentille :

Aberrations sphériques et chromatiques, Distorsions radiales (Coussinet, Barillet),

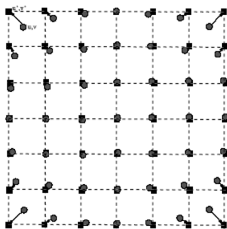
Distorsion radiale



Distorsion radiale de l'image (Coussinet ou Barrillet) :

- Causée par les imperfections de la lentille ;
- Les distorsions sont importantes sur les bords de l'image

Distorsion radiale



L'équation (4) suppose que le modèle de la caméra est linéaire.

Physiquement existent des distorsions optiques (radiales) : (4) devient :

$$\begin{cases} u = \alpha_u \cdot x + u_0 + \delta_u \\ v = \alpha_v \cdot y + v_0 + \delta_v \end{cases}$$

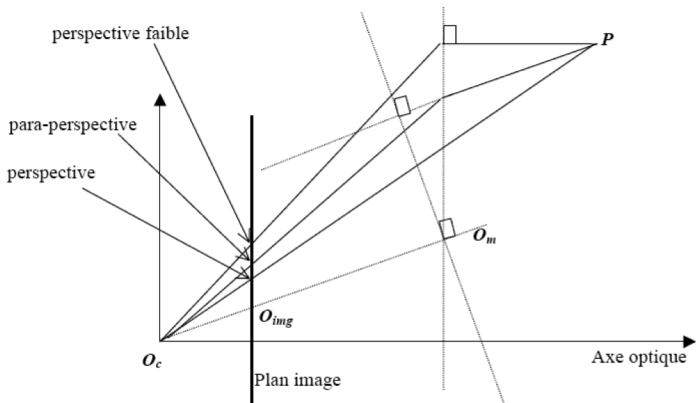
avec :

$$\delta_u = k \cdot ((u - u_0)^2 + (v - v_0)^2)(u - u_0)$$

$$\delta_v = k \cdot ((u - u_0)^2 + (v - v_0)^2)(v - v_0)$$

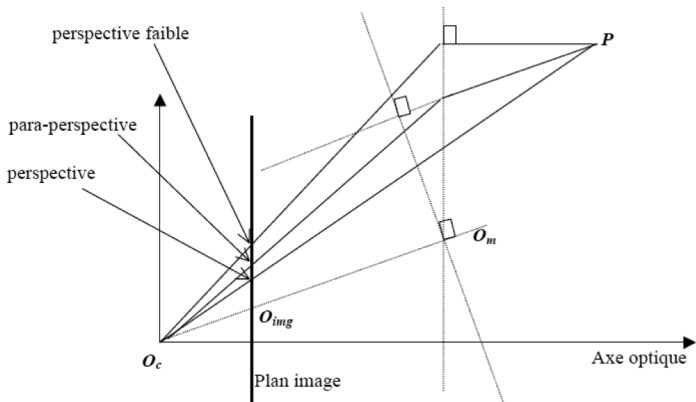
Diverses perspectives

Le modèle de la perspective est général mais complexe (10ddl).
 Considérons les cas particuliers de la perspective faible et de la para perspective



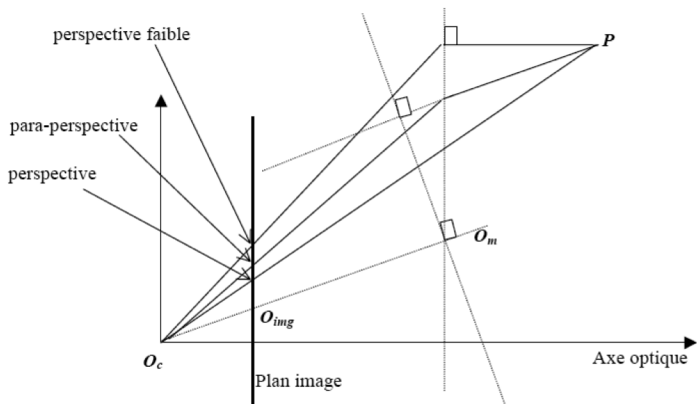
Diverses perspective

Perspective faible : Tout point P est projeté orthogonalement dans un plan // au plan image (Objets éloignés de la caméra et ont même Z). Puis $s=1/tz$ est appliqués aux points obtenus.



Diverses perspectives

Para perspective : Tout point P est projeté dans un plan // au plan image, non de façon orthogonale, mais parallèle à la droite $O_c O_m$. O_m étant l'origine du repère objet.



Modele affine

$$(4) \iff \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} \quad (14)$$

en considérant une caméra de focale $f = 1$:

$\alpha_u = 1, \alpha_v = 1$: taille focale en pixels selon abscisse et ordonnée ;

$(u_o, v_o)^t = (0, 0)^t$, coordonnées pixels de la projection de O_c dans le plan image.

Modèle affine

Le modèle projectif vu précédemment devient :

$$(6) \iff \begin{pmatrix} s \cdot u \\ s \cdot v \\ s \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_1 & t_x \\ r_2 & t_y \\ r_3 & t_z \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Les coordonnées image d'un point P deviennent :

$$\begin{cases} u = (P^t \cdot r_1 + t_x) / (P^t \cdot r_3 + t_z) \\ v = (P^t \cdot r_2 + t_y) / (P^t \cdot r_3 + t_z) \end{cases}$$

Modèle affine : perspective faible

Modèle de projection à perspective faible : En divisant par t_z les 2 équations et en considérant $1/1 + \xi \approx 1$ avec $\xi = P \cdot r_3/t_z$ on obtient le modèle affine de la perspective faible :

$$\begin{cases} u = (P^t \cdot r_1 + t_x)/t_z \\ v = (P^t \cdot r_2 + t_y)/t_z \end{cases}$$

Les points de l'objet 3D sont projetés orthogonalement sur un plan parallèle au plan image, puis un facteur d'échelle $1/t_z$ est appliqué sur les points ainsi obtenus. Ceci revient à considérer que tous les points de la scène ont en commun la même composante z dans le repère caméra ou encore que la différences entres leurs composantes en z sont négligeables par rapport) t_z : ce qui a pour effet d'atténuer les effets de la perspective.

Modèle affine : para perspective

Modèle de projection à para perspective : En divisant par t_z les 2 équations et en considérant $1/1 + \xi \approx 1 - \xi$ avec $\xi = P \cdot r_3/t_z$ on obtient le modèle affine de la perspective faible :

$$\begin{cases} u = (P^t \cdot (r_1 - t_x/t_z \cdot r_3) + t_x)/t_z \\ v = (P^t \cdot (r_2 - t_y/t_z \cdot r_3) + t_y)/t_z \end{cases}$$

Dans ce cas, les points sont projetés sur un plan parallèle au plan image, non de façon orthogonale mais en parallèle de la droite $O_c O_m$ où O_m est l'origine du repère objet. Un facteur d'échelle $1/t_z$ est ensuite appliqué sur les points ainsi obtenus. Là aussi, le modèle est d'autant plus réaliste que les dimensions de l'objet filmé sont négligeables devant leur distance au centre optique.



Questions sur le cours

Q1 : Dans l'équation (4), quelle conséquence sur l'affichage si $k_u < k_v$?

Q2 : Que représente la dernière colonne de la matrice T_{hr} ?

Q3 : Dans l'équation (6), à quoi correspondent les lignes et les Colonnes ?

Q4 : Quelles est la signification physique de m_{34} .

Q5 : Quelle est la condition utilisée pour l'estimation des m_{ij} au sens de la méthode des moindres carrés.

Travail dirigé 1

En pratique les axes abscisse et ordonnée du plan image ne sont pas forcément orthogonaux. Pour représenter cette caractéristique, les paramètres du modèle global du sténopé doivent intégrer l'angle formé par les vecteurs directeurs de ces axes. L'ajout de ce dernier porte à 11 les paramètres à calculer dont 5 internes.

Q : déterminer l'expression des paramètres intrinsèques et extrinsèques en fonction des paramètres globaux m_{ij} , (cf. énoncé détaillé).

Travail dirigé 2

En pratique le modèle de caméra est non sténopé à cause de la structure de la lentille. Pour représenter cette caractéristique, les paramètres intrinsèques doivent intégrer les changements appropriés.

Q : Déterminer les relations entre $(X_o, Y_o, Z_o)^t$, point de la scène, et $(X_c, Y_c, Z_c)^t$, son projeté dans le référentiel caméra, qui donnent le modèle non sténopé, (cf. énoncé détaillé).

Bibliographie

[Abdel-Aziz et Karara, 1971] Y.I Abdel-Aziz et H.M. Karara (1971). Direct linear transformation into object space coordinates in close-range photogrammetry. In Proceedings of the ASP/UI Symposium on Close-Range photogrammetry, pages 1-18, Urbana. University of Illinois at Urbana-Champaign.

Tsai, 1987R. Y. Tsai (1987). A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. IEEE Journal of Robotics and Automation, RA-3(4) :323-344.

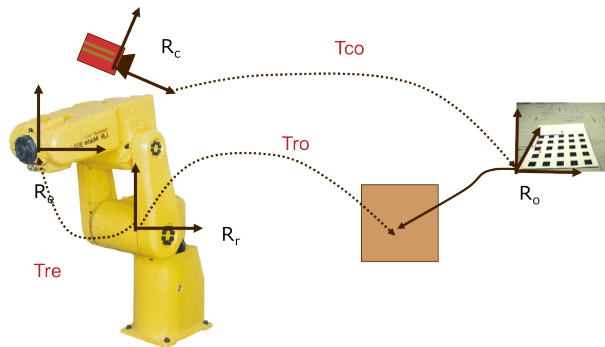
Zhang, 1999Z. Zhang (1999). Flexible camera calibration by viewing a plane from unknown orientations. in International Conference on Computer Vision, volume 1, page 666, Corfu, Greece.

Malik Mallem

Professeur UEVE

21 août 2017

Contexte



- R_c : repère de la caméra
- R_r : repère du robot
- R_e : repère effecteur
- R_o : repère environnement

Caméra RGB

Dans le cadre de la réalité augmentée monoculaire, une seule caméra est utilisée. Dans le cadre de *SLAM* multi-capteurs, par exemple, la caméra restera le capteur principalement utilisé. Nous rappellerons dans un premier temps la modélisation la plus courante du mécanisme de formation d'une image. Puis nous décrivons brièvement une méthode de calibration usuelle. La calibration permet d'estimer les différents paramètres impliqués dans cette modélisation.

Notations Mathématiques

Notations de base :

Scalars, vectors and points $\in \mathbb{R}^2$: in lowercase

Matrices and points $\in \mathbb{R}^3$: in uppercase

M : Matrice

v : Vecteur

$x \cdot y$: Produit scalaire entre les vecteurs x et y

$x \times y$: Produit vectoriel entre les vecteurs x et y

Notations Mathématiques

Géométrie euclidienne et projective :

\sim : Egalité à un facteur d'échelle près

p : Point 2D ;

\tilde{p} : p en coordonnées homogènes

P : Point 3D

\tilde{P} : P en coordonnées homogènes

d : Droite de l'espace

π : Plan de l'espace

R : Matrice de rotation

t : Vecteur de translation



Modélisation et calibration de capteurs/robot

- Objectif

Selon type de capteurs :

- Caméra → Superposition d'images
- Dispositif 3D → Relevé 3D
- Robot → Relevé 3D, Calibration automatique

- Pré conditions

Indices d'apprentissage :

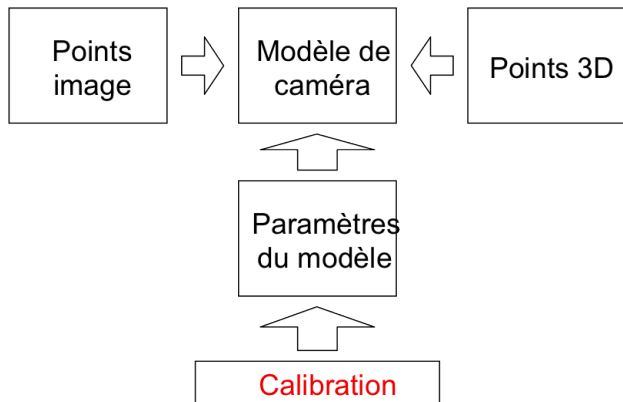
Points, Lignes, ...

- Post conditions

Paramètres du modèle

Physique, Géométrique

Modélisation et calibration de caméra



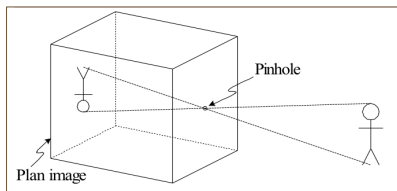
Objectif

Déterminer la position d'un objet dans l'espace à partir de sa position dans l'image.

Modèle du sténopé pour une caméra RGB

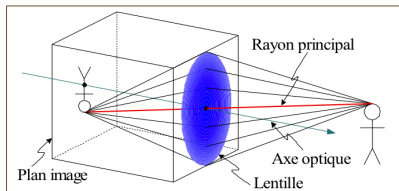
Le modèle du sténopé est une représentation assez simple de la formation d'une image dans une caméra. Dans ce modèle un point P de l'espace, \mathbb{R}^3 , se projette linéairement sur le plan image de la caméra en un point m au travers d'un trou infinitésimal (*pinhole*).

Formation des images



Sténopé (modèle idéal)

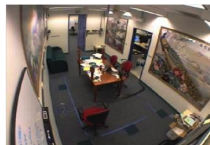
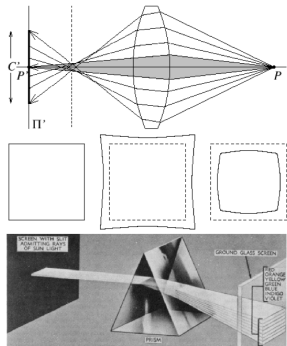
Problème : quantité de lumière insuffisante



Caméra réelle

Solution : utilisation d'une lentille

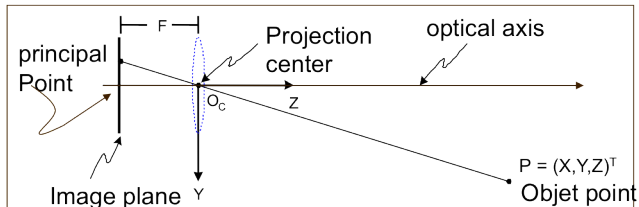
Formation des images



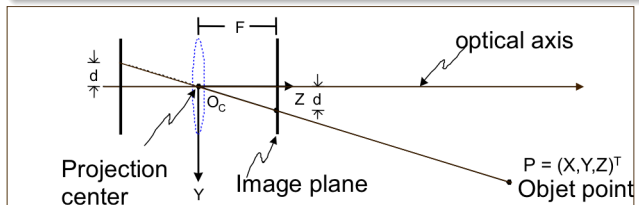
Problèmes liés à l'utilisation d'une lentille :

Aberrations sphériques et chromatiques, Distorsions radiales (Coussinet, Barillet),

Modèle du sténopé



Modèle inverseur



Modèle non inverseur

Modélisation géométrique de caméra

Hypothèses :

- Modèle linéaire à sténopé ("Pinhole"),
- Transformation perspective parfaite de centre O_c (centre optique de la caméra),
- La distance $O_c O'$ est la distance focale f de la caméra,
- L'axe optique $O_c Z_c$ intersecte le plan image en O' (centre de l'image),
- Les axes X_c et Y_c sont parallèles au plan image.

Modélisation d'une caméra, modèle du sténopé

$$\tilde{p} = T_c^e \cdot T_o^c \cdot \tilde{P} \quad (1)$$

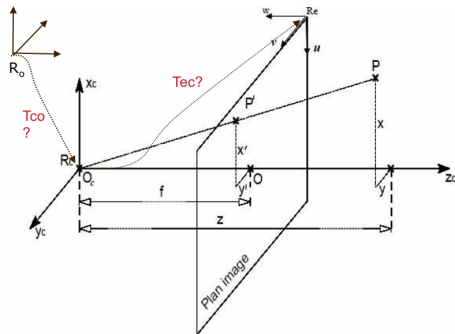
Où :

- (R_e) : repère écran ;
- (R_c) : repère caméra ;
- (R_o) : repère monde ;
- \tilde{p} : point image ;
- T_c^e : modèle interne ;
- T_o^c : modèle externe ;
- \tilde{P} : point 3D ;

avec :

$$\tilde{p} = \begin{pmatrix} u & v & 1 \end{pmatrix}^t;$$

$$\tilde{P} = \begin{pmatrix} X & Y & Z & 1 \end{pmatrix}^t;$$



Géométrie de la caméra RGB- Paramètres extrinsèques

Dans ce schéma de projection, le point de l'espace imagé doit être placé dans le repère de la caméra. Soit un point P^o de \mathbb{R}^3 défini dans un repère global (R_o) que l'on souhaite replacer dans le repère (R_c) de la caméra. On suppose que le déplacement de la caméra dans l'espace suit une transformation rigide. Alors la transformation des axes de (R_o) en ceux de (R_c) est caractérisée par un changement d'orientation et de position. Le changement de base associé au remplacement de P^o dans (R_c) est donc paramétré par une rotation et une translation. On appelle **pose** de la caméra ou encore **paramètres extrinsèques** cette modification d'orientation et cette translation.

Géométrie de la caméra RGB - Paramètres extrinsèques

Soit R la matrice de rotation permettant de transformer les axes de (R_o) en ceux de (R_c) . Soit t la translation permettant de déplacer l'origine de (R_o) en celle de (R_c) , le centre optique de la caméra. On suppose que R est appliquée préalablement à t . Notons T la matrice réalisant les deux transformations :

$$T = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \quad (2)$$

La matrice T ainsi définie est la matrice de passage de la base (R_o) à la base (R_c) . La transformation que subit P^o lors de son expression dans C est donc

$$P^c = T^{-1} \cdot P^o = \begin{pmatrix} R^T & -R^T \cdot t \\ 0 & 1 \end{pmatrix} \cdot P^o \quad (3)$$

Géométrie de la caméra RGB - Paramètres extrinsèques

Dans la suite, pour des raisons de commodité, nous représenterons la pose d'une caméra par la matrice T_o^c permettant de transformer P^o en P^c , soit T^{-1} telle qu'explicitée dans (3) :

$$T_o^c = \begin{pmatrix} R_o^c & t_o^c \\ 0 & 1 \end{pmatrix} \quad (4)$$

On notera généralement (R_c) pour désigner T_o^c .

Géométrie de la caméra RGB - Paramètres intrinsèques

Une fois le point P exprimé dans (R_c) , il est projeté dans le plan image de la caméra. On note :

$$P^c = \begin{pmatrix} x & y & z \end{pmatrix}^T \quad (5)$$

Le plan image se situant à la distance focale f du centre optique le point subit la transformation :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} f \cdot x/z \\ f \cdot y/z \\ f \end{pmatrix} \quad (6)$$

Géométrie de la caméra RGB - Paramètres intrinsèques

Le point s'image donc en $(f \cdot x/z \quad f \cdot y/z)^T$.

En utilisant les coordonnées homogènes, on représente la transformation sous forme matricielle :

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} f \cdot x/z \\ f \cdot y/z \\ 1 \end{pmatrix} \sim \begin{pmatrix} f \cdot x \\ f \cdot y \\ z \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (7)$$

Géométrie de la caméra RGB - Paramètres intrinsèques

Une fois le point projeté dans le plan image, il reste à modéliser la transformation de ses coordonnées dans l'espace des capteurs photo-sensibles de la caméra pour obtenir ses coordonnées pixelliques. Les capteurs sont organisés en grille et nous nous plaçons dans le cas idéal où les axes de celle-ci sont orthogonaux. Généralement l'origine de l'image pixellique se situe à l'extrémité supérieure gauche de la grille et les cases de la grilles sont rectangulaires. Ainsi, entrent en jeu deux transformations : une translation du centre de l'image et une mise à l'échelle.

Géométrie de la caméra RGB - Paramètres intrinsèques

Alors les coordonnées pixelliques sont obtenues par :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} k_1 & 0 & u_0 \\ 0 & k_2 & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \frac{1}{z} \cdot \begin{pmatrix} f \cdot x \\ f \cdot y \\ 1 \end{pmatrix} = \frac{1}{z} \cdot \begin{pmatrix} f \cdot k_1 & 0 & u_0 & 0 \\ 0 & f \cdot k_2 & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (8)$$

Usuellement on note la matrice des paramètres intrinsèques K :

$$K = \begin{pmatrix} f \cdot k_1 & 0 & u_0 & 0 \\ 0 & f \cdot k_2 & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (9)$$

Géométrie de la caméra RGB - Paramètres globaux

Le processus complet de la formation d'image est :

$$s \cdot \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \cdot T_o^c \cdot P^o \quad (10)$$

Géométrie de la caméra RGB - Distorsions

En réalité les ouvertures de caméras ne sont pas des trous infinitésimaux. En effet, de telles ouvertures ne laisseraient passer qu'une quantité infinitésimale de lumière que les capteurs ne pourraient pas percevoir. A la place les ouvertures sont constituées de lentilles pour concentrer les rayons lumineux afin d'en obtenir une quantité suffisante tout en pouvant contrôler leur point de convergence. La contre-partie est que la linéarité du modèle du sténopé n'est plus respectée et des distorsions apparaissent lors de la formation de l'image. Ces distorsions sont de plusieurs types, notamment radiale et tangentielle.

Géométrie de la caméra RGB - Distorsions

La prédiction de la projection d'un point dans l'image est donc erronée si l'on ne prend en compte que le modèle du sténopé. Le redressement de l'image est le nom de la procédure qui consiste à transformer une image présentant des distorsions en celle que le modèle du sténopé produirait. Dans le reste de ce manuscrit, nous supposerons que les images sont redressées, ce qui peut être réalisé efficacement en utilisant la bibliothèque logicielle OpenCV¹.

Géométrie de la caméra RGB - Caractéristiques et limitations

Les caméras sont des capteurs peu chers et apportant beaucoup d'information. Néanmoins, celle-ci n'est qu'une projection de \mathbb{R}^3 dans \mathbb{R}^2 , une dimension entière est perdue : la profondeur est inaccessible. Les fréquences d'acquisition sont élevées puisque les caméras produisent généralement au minimum 30 images par seconde. Pour une telle fréquence une image se forme en $\sim 33\text{ms}$, pour des mouvements suffisamment rapides des effets de flous peuvent se former. On appelle calibration une procédure permettant d'estimer les paramètres d'une caméra.

Calibration de la caméra

On appelle calibration une procédure permettant d'estimer les paramètres intrinsèques d'une caméra.

Plusieurs méthodes de calibration ont été proposées dans la littérature.

L'une d'entre elles a été largement popularisée par son implémentation au travers d'une *toolbox* pour *Matlab* puis dans la bibliothèque logicielle *OpenCV*. Cette méthode est celle de Zhang proposée en 2000. Comme la plupart des méthodes elle s'appuie sur l'utilisation d'une mire planaire représentant un damier dont les dimensions sont connues. La mire est observée depuis plusieurs poses distinctes. Les coins du damier sont détectés à chaque prise de vue et sont les points utilisés dans l'estimation de la transformation d'une image à une autre. Bien que l'objectif soit avant tout de déterminer les paramètres intrinsèques, il est nécessaire de déterminer également les paramètres extrinsèques puisque différents points de vues sont utilisés.

Calibration de la caméra

Le caractère planaire de la mire permet de représenter la transformation des points de la mire réelle vers une des images sous la forme d'une homographie H (voir l'annexe ??). Il permet également en toute généralité de fixer la troisième coordonnée des points de la mire à 0 et ainsi d'exprimer la chaîne de projection présentée dans (7) d'une manière simplifiée et de la mettre directement en relation avec H .

Calibration de la caméra

Cette relation entre les matrices formant la chaîne de projection entre les points de la mire et la matrice d'homographie permet d'exprimer deux relations dépendant seulement de l'homographie et des paramètres intrinsèques de la caméra. A partir de trois prises de vues distinctes il est alors possible de déterminer les paramètres intrinsèques. Une fois ceux-ci déterminés, il est possible de retrouver les paramètres extrinsèques liés à chaque prise de vue. Finalement les résultats sont affinés par un processus d'optimisation mathématique, prenant possiblement en compte le phénomène de distorsion. Les paramètres le définissant étant considérés de faible grandeur, une initialisation à zéro permet à l'optimisation de les déterminer avec précision.

Modèle du sténopé, application du modèle global

Soit P un point 3D dont les coordonnées homogènes $(X, Y, Z, 1)^t$ sont définies dans le repère monde. Ce point se projette sur le plan image en le point p de coordonnées homogènes $(u, v, 1)^t$. Soit M la matrice de projection, telle que :

(??)

$$\Leftrightarrow \begin{pmatrix} s \cdot u \\ s \cdot v \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (11)$$

Modèle du sténopé, application du modèle global

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} =$$

$$\begin{pmatrix} \alpha_u \cdot r_{11} + u_0 \cdot r_{31} & \alpha_u \cdot r_{12} + u_0 \cdot r_{32} & \alpha_u \cdot r_{13} + u_0 \cdot r_{33} & \alpha_u \cdot t_x + u_0 \cdot t_z \\ \alpha_v \cdot r_{21} + v_0 \cdot r_{31} & \alpha_v \cdot r_{22} + v_0 \cdot r_{32} & \alpha_v \cdot r_{23} + v_0 \cdot r_{33} & \alpha_v \cdot t_y + v_0 \cdot t_z \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix}$$

Modèle du sténopé, application du modèle global

$$M = \begin{pmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{pmatrix} \stackrel{\iff}{=} \begin{pmatrix} \alpha_u \cdot r_1 + u_0 \cdot r_3 & \alpha_u \cdot t_x + u_0 \cdot t_z \\ \alpha_v \cdot r_2 + v_0 \cdot r_3 & \alpha_v \cdot t_y + v_0 \cdot t_z \\ r_3 & t_z \end{pmatrix} \quad (12)$$

En posant :

$$m_i = [m_{i1} m_{i2} m_{i3}]$$

et :

$$r_i = [r_{i1} r_{i2} r_{i3}]$$

Modèle du sténopé, calcul des paramètres intrinsèques et extrinsèques

Le calcul des paramètres intrinsèques et extrinsèques s'obtient en identifiant le système (12). Les m_{ij} sont supposés connus.

$$r_3 = m_3$$

$$u_0 = m_1.m_3$$

$$v_0 = m_2.m_3$$

$$\alpha_u = -\|m_1 \times m_3\|$$

$$\alpha_v = \|m_2 \times m_3\|$$

$$r_1 = 1/\alpha_u * (m_1 - u_0 * p_3)$$

Modèle du sténopé, calcul des paramètres intrinsèques et extrinsèques

$$r_2 = 1/\alpha_v * (m_2 - v_0 * m_3)$$

$$t_x = 1/\alpha_u * (m_{14} - u_0 * m_{34})$$

$$t_y = 1/\alpha_v * (m_{24} - v_0 * m_{34})$$

$$t_z = m_{34}$$

(11) équivaut à :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} u = (P^t \cdot m_1 + m_{14}) / (P^t \cdot m_3 + m_{34}) \\ v = (P^t \cdot m_2 + m_{24}) / (P^t \cdot m_3 + m_{34}) \end{cases} \Leftrightarrow$$

$$\begin{cases} P^t \cdot n_u + a_u = 0 \\ P^t \cdot n_v + a_v = 0 \end{cases}$$

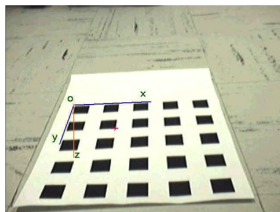
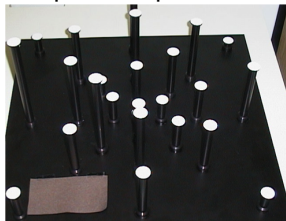
avec :

$$n_u = m_1 - u \cdot m_3; n_v = m_2 - v \cdot m_3; a_u = m_{14} - u \cdot m_{34}; a_v = m_{24} - v \cdot m_{34};$$

Il s'agit de l'équation du rayon visuel porteur de P et de son image p et passant par le centre de la caméra.

Calibration d'une caméra : procédure

1. Placer une cible d'étalonnage (figures ci-dessous) devant la caméra ;
2. Repérer la position de chaque marqueur de la cible dans l'image.



3. Connaissant les coordonnées 3D des marqueurs, leurs projections dans l'image sont prélevées.

On obtient ainsi une liste de correspondances 3D-2D.

Calibration d'une caméra : estimation des paramètres

Procédure

1. Chaque pt nous donne 2 équations de plans et on a 11 inconnues (5 intrinsèques + 6 extrinsèques) ;
2. On construit un système d'équations linéaires qu'on peut résoudre à l'aide des techniques d'algèbre linéaire standards (moindre carrés) -> estimation des m_{ij} ;
3. On calcule, s'il y a lieu, les paramètres explicites (intrinsèques et extrinsèques)

Calibration d'une caméra : estimation des paramètres m_{ij}

Soit P un point 3D dont les coordonnées homogènes $(X, Y, Z, 1)^t$ sont définies dans le repère monde. Ce point se projette sur le plan image en le point p de coordonnées homogènes $(u, v, 1)^t$. Soit M la matrice de projection, telle que :

$$\Leftrightarrow \begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (14)$$

Calibration d'une caméra : estimation des paramètres m_{ij}

Avec s un facteur d'échelle. Cette équation se réécrit :

$$u = \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}}; v = \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \quad (15)$$

Calibrer la caméra revient à estimer les 12 coefficients de M : 6 points minimum, non tous coplanaires, sont alors requis.

Le système précédent peut s'écrire pour chaque point P_i :

$$m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14} - u_i m_{31}X_i - u_i m_{32}Y_i - u_i m_{33}Z_i = u_i m_{34}$$

$$m_{21}X_i + m_{22}Y_i + m_{23}Z_i + m_{24} - v_i m_{31}X_i - v_i m_{32}Y_i - v_i m_{33}Z_i = v_i m_{34} \quad (16)$$

Estimation des m_{ij} : méthode des moindres carrés

Une mise en forme matricielle de (16) permet d'obtenir $2N$ équations :

$$\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_i X_i & -u_i Y_i \\ -u_i Z_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_i X_i & -v_i Y_i \\ -v_i Z_i \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \end{pmatrix} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ -u_i m_{34} \\ -v_i m_{34} \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad (17)$$

Un système linéaire de la forme : $H \cdot m = h$

Une solution au sens des moindres carrés donne : $m = (H^t \cdot H)^{-1} \cdot H^t \cdot h$

Estimation des m_{ij} : méthode des moindres carrés

A partir d'un ensemble de points (u_i, v_i) et un ensemble de correspondants de points 3D (X_i, Y_i, Z_i) , l'approche estime la matrice de projection M (matrice 3x4) en minimisant le critère suivant

$$Q = \|H.m - h\|^2 \quad (18)$$

Où :

$m^t = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{21} & m_{22} & m_{23} & m_{24} & m_{31} & m_{32} \\ m_{33} \end{pmatrix}$ est le

vecteur à identifier ;

h est l'échantillon des N points images mesurés ;

et H : la matrice de mesures ($2N$ lignes, 11 colonnes).

Estimation des m_{ij} : méthode de Faugeras-Toscani

A partir d'un ensemble de points (u_i, v_i) et un ensemble de correspondant de points 3D (X_i, Y_i, Z_i) , l'approche estime la matrice de projection M (matrice 3x4) en minimisant le critère suivant

$$Q = \|BX_9 + CX_3\|^2 \quad (19)$$

Où

$$B = \begin{pmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_i \end{pmatrix} \quad (20)$$

$$C = \begin{pmatrix} -u_i X_i & -u_i Y_i & -u_i Z_i \\ -v_i X_i & -v_i Y_i & -v_i Z_i \end{pmatrix} \quad (21)$$

Estimation des m_{ij} : méthode de Faugeras-Toscani

Et,

$$X_9 = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{21} & m_{22} & m_{23} & m_{24} & m_{34} \end{pmatrix}^t$$

$$= \begin{pmatrix} m_1 & m_{14} & m_2 & m_{24} & m_{34} \end{pmatrix}^t$$

$$X_3 = \begin{pmatrix} m_{31} & m_{32} & m_{33} \end{pmatrix}^t = m_3^t(22)$$

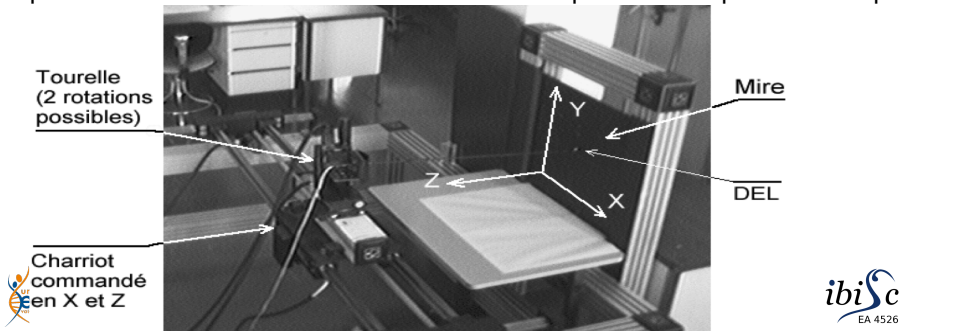
Estimation des m_{ij} : méthode de Faugeras-Toscani

Exercice :

1. Modifier le critère (19) de manière à y intégrer la contrainte $\|m_3\| = 1$;
2. Minimiser ce dernier en imposant que les dérivées partielles par rapport à X_9 et X_3 sont nulles ;
3. Calculer X_9 puis X_3

Expérimentation : utilisation d'un banc de calibration

Cette expérimentation s'est basée sur l'utilisation d'un banc de calibration industriel, consistant en une table XZ précise au 1/10 de mm et une tourelle site et azimut précise au 1/100 de degré. Une tige est solidaire de la tourelle faisant office d'un robot à 4 degrés de liberté permettant de réaliser l'opératoin de saisie des correspondances 3D et 2D préalable à la procédure de calibration de la caméra. Ce dispositif est représenté ci-après :



Expérimentation : équipement et conditions

1. Caméra CCD :
 - focale de 25 mm, définition : 756x581 photoéléments.
2. Caméra à tube :
 - focale de 25 mm
3. Banc de calibration :
 - Distance caméra-plans de calibration : 1-2mètres ;
 - Robot à 4 d.d.l.(PPRR)($\delta\theta = 1/100^\circ$, $\delta X = 1/10mm$)
4. Calculateur :
 - Définition de la carte graphique : 768x576 pixels.

Expérimentation : protocole

1. Calibration du mini robot / à la mire du banc de calibration ;
2. Pré calibration de la caméra permettant d'obtenir le champ visuel de la caméra ;
3. Calibration automatique de la caméra afin d'obtenir un échantillon de points de correspondances 3D et 2D permettant de réaliser une étude statistique de la précision de la calibration de la caméra..

Expérimentation : Méthodes utilisées

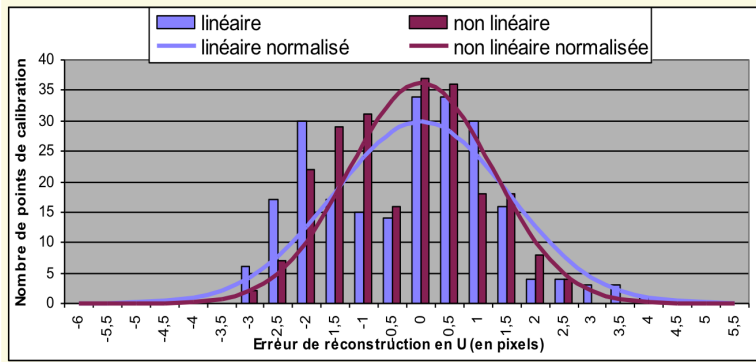
L'objectif est de calculer l'erreur de reconstruction pixellique en u et en v .

L'erreur en u est utilisée.

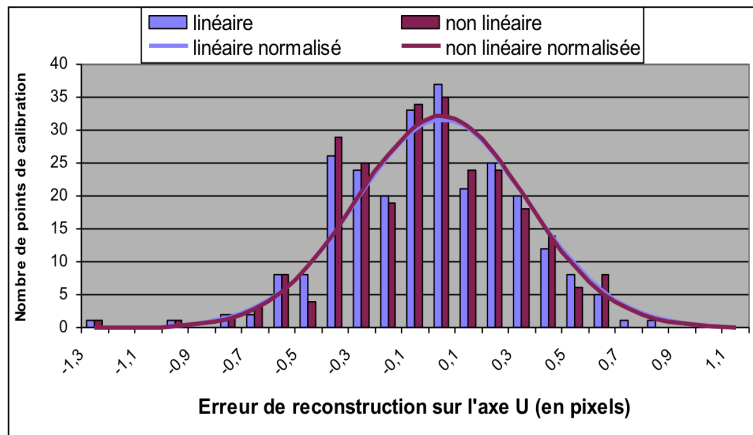
L'estimation des paramètres de la caméra se base sur :

- pour le modèle linéaire (sans distorsions) par la méthode des moindres carrés ;
- pour le modèle non linéaire (incluant la distorsion radiale) par la méthode de Levenberg-Marquard.

Expérimentation : résultat - caméra à tube



Expérimentation : résultat - caméra CCD



Expérimentation : Bilan

L'expérimentation basée sur le banc de calibration apporte plusieurs avantages dont :

- Evaluation rapide et précise de modèles de caméra ;
- Non utilisation d'un objet étalon ;
- Calibration automatique si caméra en mouvement.

Calibration de caméra : diverses méthodes

Il existe plusieurs techniques d'étalonnage : Zhang, Tsai, Heikkilä, Faugeras, etc.

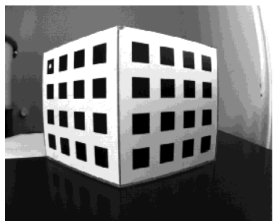
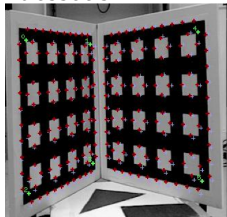
- Le type de cible et le modèle de la caméra utilisé changent ;
- Le système d'équations à résoudre diffère en fonction de ces variantes ;
- L'idée est toujours la même : trouver les paramètres permettant de faire le "mapping" entre les points de l'espace et leur position dans l'image.

Calibration de caméra : diverses méthodes

- Prise en compte de la distorsion des lentilles : - La plupart des techniques incluent des paramètres de distorsion dans leur modèle ;
- L'ajout de ces paramètres rend le système à résoudre non-linéaire.
 - On procède alors en 2 étapes :
1. En supposant la distorsion nulle, on résout le système linéaire ;
 2. Partant de cette estimation initiale, on détermine les paramètres de distorsion à l'aide de techniques d'optimisation non-linéaires.

Calibration d'une caméra : méthode de Zhang

Dans les méthodes précédentes, la calibration exigeait que tous les points ne soient pas co planaires. Les exemples de mires utilisées sont données ci-dessous :



Note : Une telle armada, n'est pas forcément nécessaire !

Calibration d'une caméra : méthode de Zhang

- L'idée est de proposer une méthode de calibration qui permet une calibration en s'affranchissant de la conception de mires complexes ;
- Une grille plane imprimée en N/B sur une feuille de papier et montrée sous plusieurs vues à la caméra suffit. Le mouvement de la caméra ou de la grille n'est pas à connaître ;
 - Les paramètres extrinsèques sont estimés relativement au référentiel lié à la grille de calibration ;
 - Claim to Fame : est un outil basé sur l'algorithme de Zhang et utilisé par MS Research. Matlab Camera Calibration Toolbox utilise également cet algorithme.

Calibration d'une caméra : méthode de Zhang

Caractéristiques de la technique de Zhang :

- Simple, robuste et précise ;
- Utilise une cible plane ;
- Nécessite au moins trois prises de vue (non coplanaires) à cause des paramètres intrinsèques ;
- Prend en compte la distorsion des lentilles.

Liens utiles :

- Zhengyou Zhang (Microsoft research)

<http://research.microsoft.com/en-us/um/people/zhang/calib/> - Janne

Heikkila (Oulu Univ.) <http://www.ee.oulu.fi/jth/calibr/> - Lien

potentiellement intéressant (pour travaux pratiques) :

<http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/links.html>

Questions sur le cours

- Q1 : Dans l'équation (11), quelle conséquence sur l'affichage si $K_u < K_v$?
- Q2 : Que représente la dernière colonne de la matrice T_{hr} ?
- Q3 : Dans l'équation (28), à quoi correspondent les lignes et les Colonnes ?
- Q4 : Quelles est la signification physique de m_{34} ;
- Q5 : Quelle est la condition utilisée pour l'estimation des m_{ij} au sens de la méthode des moindres carrés.

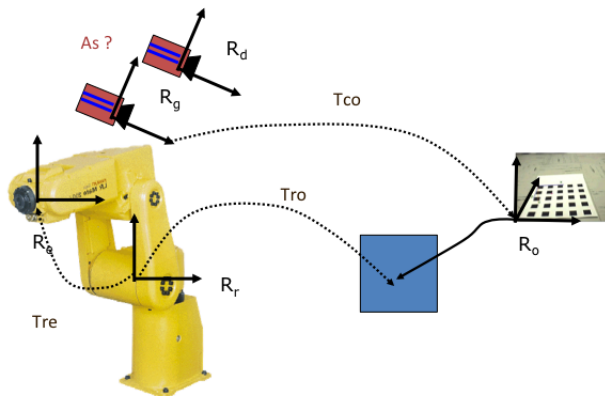
Bibliographie

[Abdel-Aziz 1971] Y.I Abdel-Aziz et H.M. Karara (1971). Direct linear transformation into object space coordinates in close-range photogrammetry. In Proceedings of the ASP/UI Symposium on Close-Range photogrammetry, pages 1-18, Urbana. University of Illinois at Urbana-Champaign.

Tsai, 1987R. Y. Tsai (1987). A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. IEEE Journal of Robotics and Automation, RA-3(4) :323-344.

Zhang, 1999Z. Zhang (1999). Flexible camera calibration by viewing a plane from unknown orientations. in International Conference on Computer Vision, volume 1, page 666, Corfu, Greece.

Contexte



Notations Mathématiques

Notations de base :

Scalars, vectors and points $\in \mathbb{R}^2$: in lowercase

Matrices and points $\in \mathbb{R}^3$: in uppercase

M : Matrice

v : Vecteur

$x \cdot y$: Produit scalaire entre les vecteurs x et y

$x \times y$: Produit vectoriel entre les vecteurs x et y

Notations Mathématiques

Géométrie euclidienne et projective :

- \sim : Egalité à un facteur d'échelle près
- p : Point 2D exprimé en coordonnées métriques ;
- m : Point 2D exprimé en coordonnées pixelliques ;
- \tilde{p} : p en coordonnées homogènes
- P : Point 3D
- \tilde{P} : P en coordonnées homogènes
- d : Droite de l'espace
- π : Plan de l'espace
- R : Matrice de rotation
- A^{-1} : Matrice inverse
- A^{+} : Matrice pseudo inverse



Reconstruction 3D

- Objectif : Permettre au robot de poursuivre sa tâche

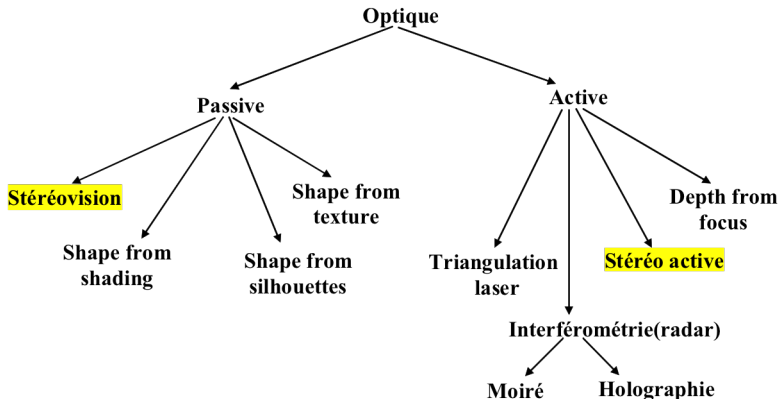
Pré-conditions : :

- Calibrer les capteurs/environnement du robot,
- Communiquer au système la classe d'objet à modéliser,
- Télé commander les capteurs pour effectuer les mesures.

Post conditions

Compléter la connaissance(BD3D) du robot sur son environnement.

Systèmes optiques de reconstruction 3D



Objectif

Etudier la reconstruction 3D basée sur les stéréovisions passive et active.

Stéréovision passive - Problématique

Hypothèses :

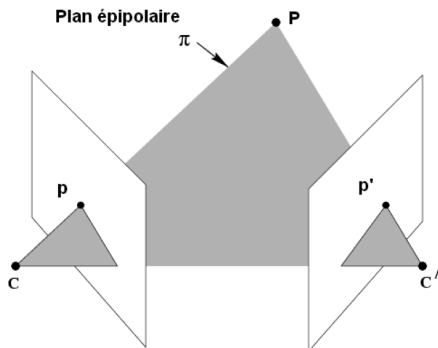
- 2 caméras de centre C et C' ,
- p et p' projections de P ,
- C, C', p, p' et P sont coplanaires,
- à chaque caméra est associé un repère

Données :

- MGI des 2 caméras (voir section 2)

Problème :

Calculer coordonnées de P ?



Géométrie épipolaire

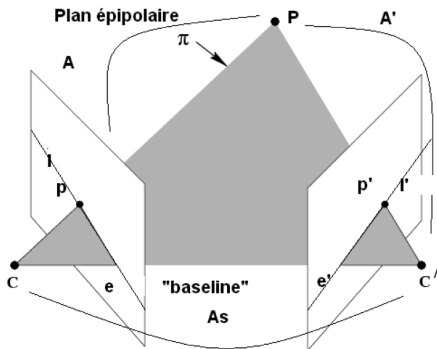
Utilité : inférence 3D



Stéréovision passive - Géométrie

Définitions :

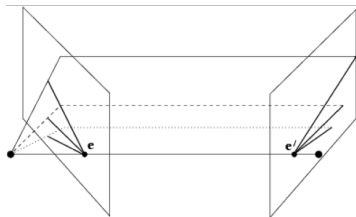
- e et e' sont les épipoles (projections de C et C'),
- baseline intersecte les plans images en les epipoles e et e' ,
- l et l' sont les lignes épipolaires,
- Tout plan contenant baseline est un plan épipolaire,
- A , A' , A_s Mext de caméras.



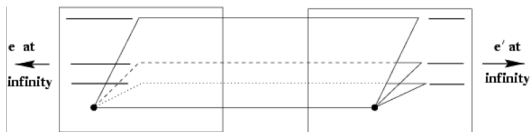
Géométrie épipolaire

Utilité : inférence 3D

Stéréovision passive - Illustration



Stéréovision passive - Illustration 2



Stéréovision passive - calibration

Démarche

- Calibrer chaque caméra/ R_0 (mire),
- Extraire Mext A et A' de chaque caméra,
- Calculer Mext de passage de R_g à R_d : $A_s = A'.A^{-1}$

$$A_s = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} r_1 & t_x \\ r_2 & t_y \\ r_3 & t_z \\ 0 & 1 \end{pmatrix} \quad (1)$$

Stéréovision passive - calibration

Relation entre caméra droite et caméra gauche :

$$\begin{cases} x = (Zr_1 \cdot p + t_x)/(Zr_3 \cdot p + t_z) \\ y = (Zr_2 \cdot p + t_y)/(Zr_3 \cdot p + t_z) \end{cases}$$

avec :

$$p = (x, y)^t \text{ et } p' = (x', y')^t$$

Sachant que la profondeur Z de P est connue ainsi que les paramètres extrinsèques du capteur.

Stéréovision passive - calibration

En éliminant Z dans (11). On obtient l'équation de la ligne épipolaire droite qui représente le lieu où on doit chercher un correspondant à un point dans le repère de l'image gauche .

$$a'x' + b'y' + c' = 0$$

$$a' = (b_y r_{31} - b_z r_{21})x + (b_y r_{32} - b_z r_{22})y + (b_y r_{33} - b_z r_{23})$$

$$b' = (b_z r_{11} - b_x r_{31})x + (b_z r_{12} - b_x r_{32})y + (b_z r_{13} - b_x r_{33})$$

$$c' = (b_x r_{21} - b_y r_{11})x + (b_x r_{22} - b_y r_{12})y + (b_x r_{23} - b_y r_{13})$$

$$\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \cdot \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & b_x \\ -b_y & 0 & b_x \end{pmatrix} = E \cdot p \quad (2)$$

Stéréovision passive - matrice essentielle

La matrice essentielle E possède 5 degrés de libertés : La rotation r et la translation t possèdent toutes deux 3 degrés de liberté mais il subsiste une ambiguïté sur l'échelle globale due à l'utilisation de coordonnées homogènes (ou coordonnées normalisées) pour un point $p = (x, y, 1)^t$ dans le repère de la caméra.

$$(x' y' 1) \cdot E \cdot \begin{pmatrix} x \\ y' \\ 1 \end{pmatrix} = 0 \iff p'^t \cdot E \cdot p = 0 \quad (3)$$

Stéréovision passive - matrice fondamentale

En combinant les transformations caméra-image définies par l'équation du modèle interne de la caméra (cf. section modèle de caméra) et la relation précédente (3), on peut donc enfin définir une relation entre les points des deux images de la paire stéréo dans les repères propres aux images : Soit m un point dans le repère de l'image gauche et p un point dans le repère de la caméra gauche (respectivement m' et p' pour la caméra droite).

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \iff m \cdot K = p \quad (4)$$

Stéréovision passive - matrice fondamentale

On peut donc établir la relation suivante entre points dans le repère image, on définit ainsi la matrice fondamentale que l'on peut considérer comme la version "image" de la matrice essentielle :

$$(4) \iff m'^t \cdot K'^{-1} \cdot E \cdot K^{-1} \cdot m = 0$$

$$\iff m'^t \cdot F \cdot m = 0 \quad (5)$$

F dispose de 9 ddl (5 externes et 4 internes).

Reconstruction 3D par Stéréovision passive

En reprenant l'équation du rayon visuel (cf. section modèle de caméra) et en l'appliquant au cas stéréo, on peut écrire :

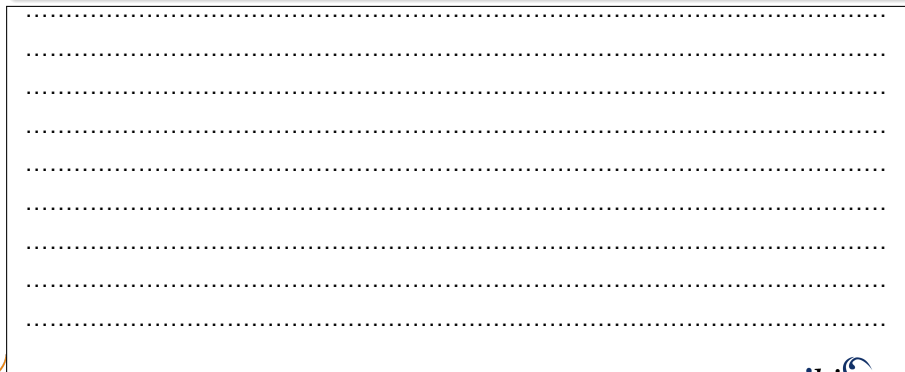
$$\begin{cases} u_1 = \frac{m_{11}^1 X + m_{12}^1 Y + m_{13}^1 Z + m_{14}^1}{m_{31}^1 X + m_{32}^1 Y + m_{33}^1 Z + m_{34}^1} \\ v_1 = \frac{m_{21}^1 X + m_{22}^1 Y + m_{23}^1 Z + m_{24}^1}{m_{31}^1 X + m_{32}^1 Y + m_{33}^1 Z + m_{34}^1} \\ u_2 = \frac{m_{11}^2 X + m_{12}^2 Y + m_{13}^2 Z + m_{14}^2}{m_{31}^2 X + m_{32}^2 Y + m_{33}^2 Z + m_{34}^2} \\ v_2 = \frac{m_{21}^2 X + m_{22}^2 Y + m_{23}^2 Z + m_{24}^2}{m_{31}^2 X + m_{32}^2 Y + m_{33}^2 Z + m_{34}^2} \end{cases}$$

X, Y, Z du point P reconstruit dans le repère de calibrage, se calculent en résolvant ce système de 4 équations linéaires.

Reconstruction 3D par Stéréovision passive

Exercice :

Déterminer les équations du système de 4 équations linéaires ?



Stéréovision active - Problématique

Hypothèses :

- capteur actif composé d'une caméra et d'une source laser,
- p projection de P dans plan image,
- à chaque capteur est associé un repère, (R_0 : référentiel absolu).

Données :

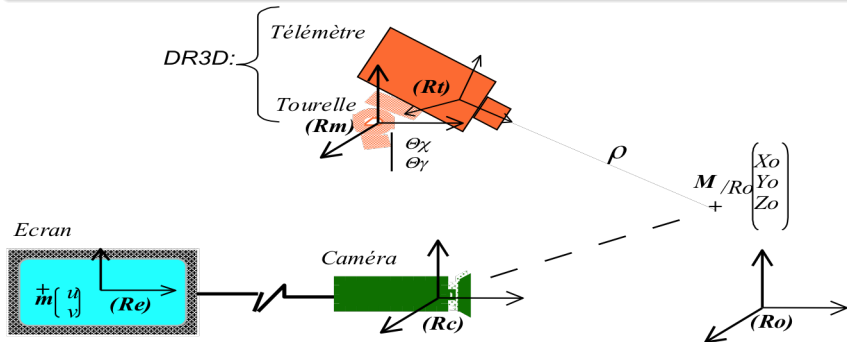
- MGD de la caméra (matrice M : voir section modélisation de caméra), -
- équation du plan lumineux dans repère mire (à déterminer au préalable).

Problème :

Calculer coordonnées de P dans R_0 ?

Capteur actif - Exemple

Capteur actif (DR3D) composé d'une caméra et d'un télémètre laser.



R_m : Repère lié à la tourelle

R_t : Repère du télémètre

R_c : Repère de la caméra

R_e : Repère global

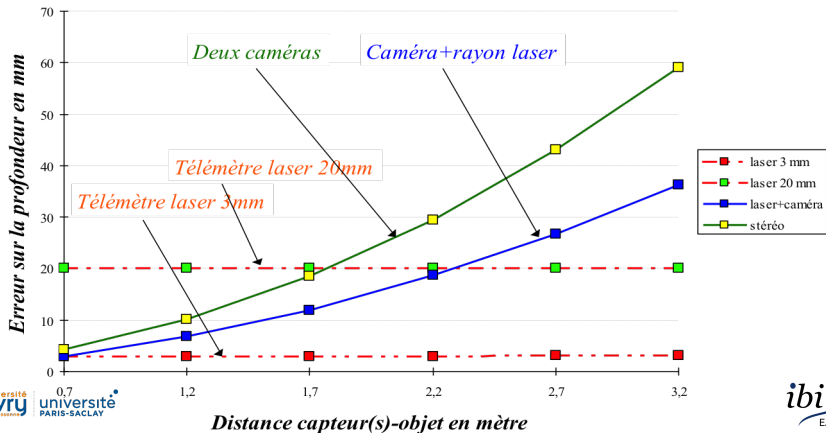
R_e : Repère de visualisation

$$(X_o, Y_o, Z_o) = f_{DR3D}(\theta_x, \theta_y, \rho)$$

$$(u, v) = f_{camera}(X_o, Y_o, Z_o)$$

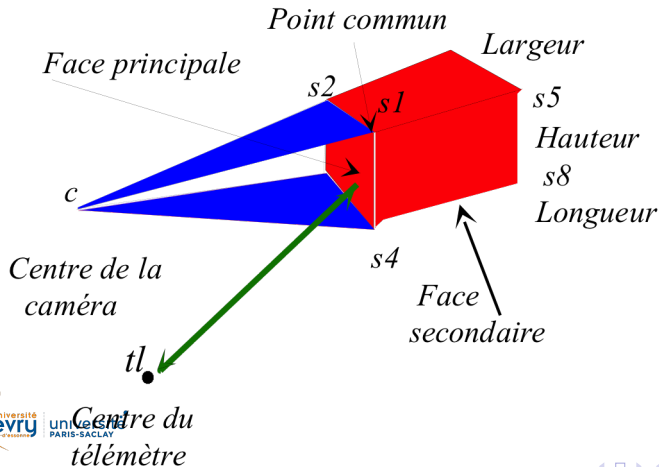
Capteur actif - Evaluation

Erreur statique sur l'inférence de la profondeur en utilisant plusieurs combinaisons de capteurs.



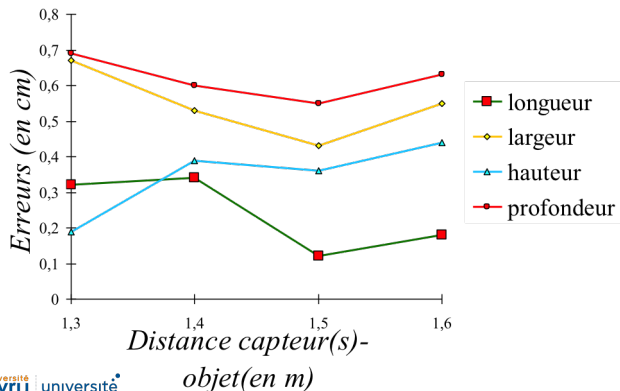
Usage capteur actif - reconstruction 3D d'un polyèdre

La caméra sert à définir la *frustum* contenant le polyèdre. Le télémètre permet de calculer la profondeur.



Evaluation reconstruction 3D d'un polyèdre

Erreur statique sur l'inférence des dimensions et de la profondeur du polyèdre.

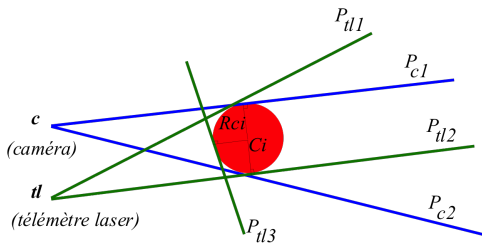


Usage capteur actif - reconstruction 3D d'un cylindre

La caméra sert à définir la *frustum* contenant le cylindre. Le télémètre permet de calculer la profondeur.

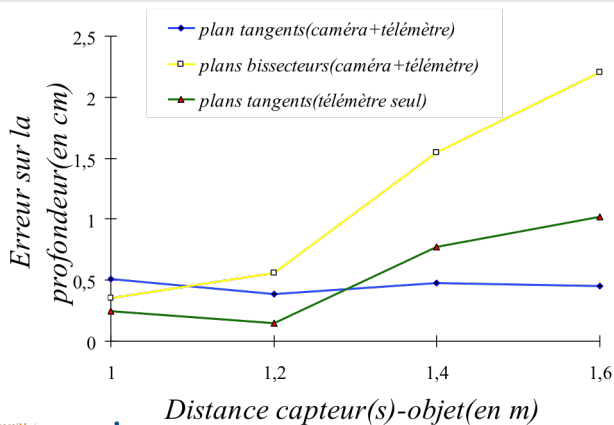
3 méthodes développées :

- Plans tangents (Caméra + Télémètre) : P_{c1}, P_{c2}, P_{tl3} .
- Plans tangents (Télémètre seul) : $P_{tl1}, P_{tl2}, P_{tl3}$.
- Plans bissecteurs (Caméra et Télémètre) : $P_{mc} = P_{c1} \cap P_{c2}$ et $P_{mtl} = P_{tl1} \cap P_{tl2}$



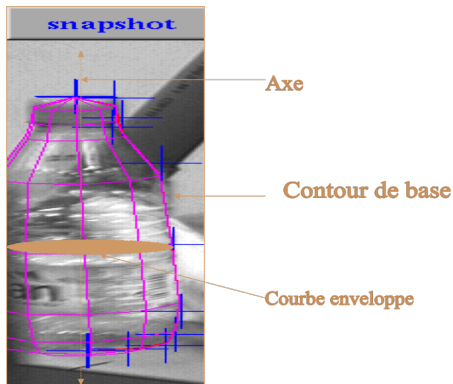
Evaluation reconstruction 3D d'un cylindre

Erreur statique sur l'inférence de la profondeur du cylindre.



Usage capteur actif - reconstruction 3D d'un cylindre généralisé

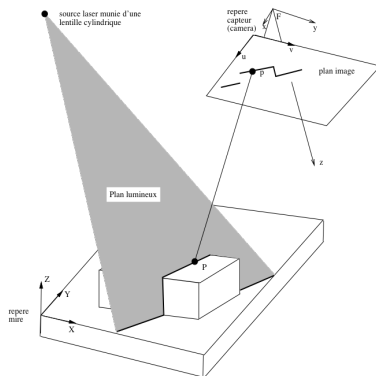
La caméra sert à définir la *frustum* contenant le cylindre généralisé. Le télémètre permet de calculer la profondeur.



Stéréovision active - Géométrie

Définitions :

- R_C : repère caméra,
- (u, v) coord. de p , projection de $P(x, y, z)$ dans plan image,
- $L = (b_1, b_2, b_3)^t$, normale au plan laser,
- R_O , repère mire dont Z est perpendiculaire au plan de pose de la mire,
- M matrice de calibration de la caméra.



Géométrie épipolaire [Horaud 1993]

Utilité : inférence 3D

Stéréovision active - Calibration (L ?)

Exercice : Déterminer l'équation du plan laser L ?

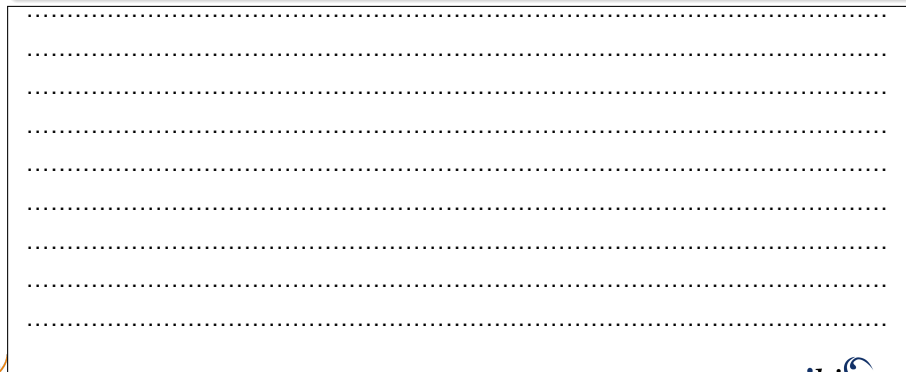
- Démarche :
- Collecter n points P_i de l'objet étalon en fixant $Z_i = H_i$,
 - Les coord. (u_i, v_i, X_i, Y_i) de P_i sont obtenus à partir du modèle de la caméra,
 - P_i appartenant au plan laser, b_3 est connu.
 - En divisant par ce dernier, déterminer l'expression de la nouvelle équation :

Pour n points à différentes hauteurs, on obtient n équations. Les paramètres du plan laser sont la solution au sens des moindres carrés, poser le système d'équations et indiquer la méthode de résolution :

Stéréovision active - reconstruction 3D

Exercice :

Déterminer les coordonnées d'un point P ?



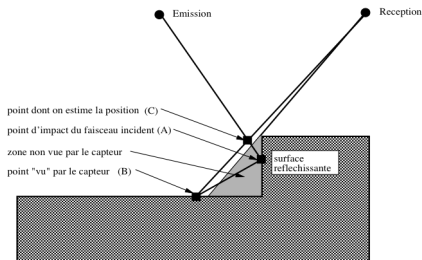
Stéréovision active - Bilan

Avantage :

L'apport principal est d'éviter l'appariement entre primitives géométriques indispensable en stéréovision passive.

Inconvénient :

- instrumentation par une tourelle de la caméra/laser pour une acquisition 3D,



Ambiguïté de points (figure [Horaud 1993])

Stéréovision - Bilan

Avantage :

L'apport principal est d'éviter l'appariement entre primitives géométriques indispensable en stéréovision passive.

Inconvénients :

- appariement (stéréo vision passive) - instrumentation par une tourelle de la caméra/laser pour une acquisition 3D (stéréo vision active) ,

Questions sur le cours

Q1 : Quelle différence y a t il entre les matrices essentielle et Fondamentale ?

Q2 : Quand est ce qu'on peut utiliser l'une ou l'autre en vue d'une reconstruction 3D ?

Q3 : Est il possible d'estimer la matrice fondamentale F ?
Si oui comment ?

Q4 : A quoi sert la rectification d'images

Q5 : Quel est l'apport majeur de la stéréovision active

Bibliographie

Radu Horaud Olivier Monga - Vision par ordinateur : outils fondamentaux (2e édition revue et augmentée). Traité des Nouvelles Technologies, série informatique. Ed. Hermes, 1993, ISBN : 2-86601-481-2,

<http://perception.inrialpes.fr/people/Horaud/livre-fichiersPS/VO-HoraudMonga.pdf>

Olivier Faugeras - Three-dimensional computer vision : a geometric viewpoint. Ed. MIT Press, 1993, ISBN : 0-262-06158-9

R. I. Hartley A. Zisserman - Multiple View Geometry in Computer Vision (Second). Ed. Cambridge University Press, 2004, ISBN : 0521540518

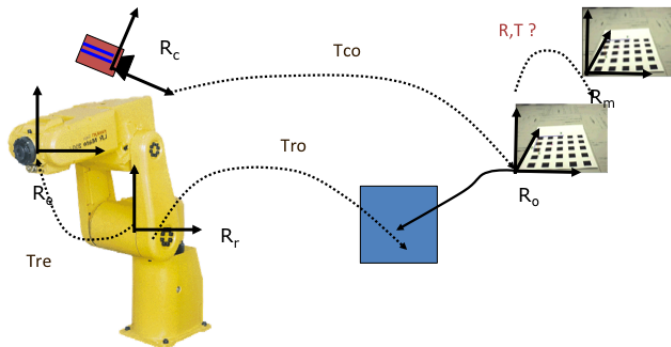
Malik Mallem

Professeur UEVE

24 septembre 2017

Sommaire

Contexte



Notations Mathématiques

Notations de base :

Scalars, vectors and points $\in \mathbb{R}^2$: en minuscules

Matrices and points $\in \mathbb{R}^3$: en majuscules

M : Matrice

v : Vecteur

$x \cdot y$: Produit scalaire entre les vecteurs x et y

$x \times y$: Produit vectoriel entre les vecteurs x et y

Notations Mathématiques

Géométrie euclidienne et projective :

- \sim : Egalité à un facteur d'échelle près
- p : Point 2D exprimé en coordonnées métriques ;
- m : Point 2D exprimé en coordonnées pixelliques ;
- \tilde{p} : p en coordonnées homogènes
- P : Point 3D
- \tilde{P} : P en coordonnées homogènes
- d : Droite de l'espace
- π : Plan de l'espace
- R : Matrice de rotation
- A^{-1} : Matrice inverse
- A^{+} : Matrice pseudo inverse



Calcul de pose

- Objectif : Mise en correspondance des mondes réel et virtuel :
Désigner/agir dans le monde virtuel équivaut à désigner/agir dans le monde réel.

Pré-conditions : :

- Calibrer les capteurs/environnement du robot,
- modéliser les objets à reconstruire,
- effectuer l'appariement 2D-3D ou 3D-3D.

Post conditions

Mise à jour de la BD3D conformément au monde réel.

Calcul de pose

Définition du problème :

Soit un objet réel et son image dans le plan image d'une caméra (modèle sténopé linéaire).

Cet objet est supposé connu. Son modèle géométrique est donc connu. Ce modèle géométrique est appelé objet virtuel. Cet objet virtuel est placé à un endroit (position, orientation) connu dans le repère caméra R_C .

L'image filaire de cet objet virtuel est obtenue par l'application du modèle de la caméra supposé connu.

L'objet réel subi un mouvement non connu.

Le problème consiste à déterminer la nouvelle position et orientation de l'objet réel dans le repère caméra. ?

Solutions

Il s'agit dans un premier temps d'effectuer les correspondances (appariements) entre éléments géométriques de l'objet réel et de son image.

Une fois les correspondants obtenus, la pose peut être estimée. Celle-ci peut être estimée en calculant le mouvement qui permet d'aligner le contour projeté avec le contour image. Ce mouvement permet de mettre à jour la pose prédite. Cependant, ce modèle de solution requière une bonne connaissance a priori de l'environnement.

Il ya deux approches de solutions :

- les solutions analytiques, nécessitant peu de correspondances ;
- les solutions numériques, un nombre important de correspondances afin de minimiser un critère.

Solutions analytiques

Dans ce cas, peu de correspondances (3 à 6) sont utilisées.

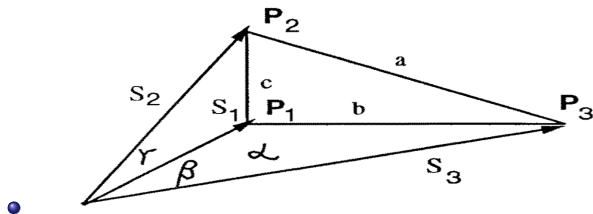
Il existe une famille de méthodes dites PnP :

- $n=3$: [Grunert, 1841] [Fischler et Bolles, 1981],
- $n=4$ correspondances coplanaires : [Hung et al, 1985],
- $n=4$ correspondances non coplanaires : [Horaud et al, 1989],
- pour 3 droites [Dhome et al, 1989].

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Problème : Calcul de la pose de la caméra/ R_o à partir de 3 points $P_i (i = 1..3)$? [Grunert 1841], (cf. figure)



Hypothèses : Longueurs (a, b, c) du triangle $P_1P_2P_3$ connus ;
 Vecteurs unitaires de $O_cP_i = J_i (i = 1..3)$ et focale caméra connus.

Inconnues : $S_i (i = 1..3)$: longueurs des segments $O_cP_i (i = 1..3)$?

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Etape 1 : Calculer des coordonnées 3D de 3 points $P_i (i = 1..3)$?

Démarche :

- Calculer les côtes (a, b, c) du triangle $P_1P_2P_3$;
 - Calculer les coordonnées des points $q_i(u_i, v_i)$ projections des $P_i (i = 1..3)$ dans le plan image situé à une distance f de O_c ,
 - Exprimer ensuite les vecteurs directeurs J_i en fonction des q_i ,
 - Calculer ensuite les angles α, β, γ ,
 - Les longueurs S_i peuvent alors être calculées,
- Connaissant S_i et J_i , on déduit les coordonnées 3D des P_i .

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : Calculer des coordonnées 3D de 3 points $P_i (i = 1..3)$?

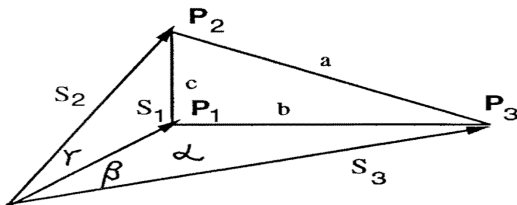
Démarche :

- Calculer les côtes (a, b, c) du triangle $P_1P_2P_3$;
- Calculer les coordonnées des points $q_i(u_i, v_i)$ projections des $P_i (i = 1..3)$ dans le plan image situé à une distance f de O_c ,

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Problème : Calcul de la pose de la caméra/ R_o à partir de 3 points $P_i, (i = 1..3)$? ,(cf. figure)



Hypothèses : Coordonnées 3D de 3 points $P_i, (i = 1..3)$ connues

Inconnues : Rotation R et Translation t ?

Solutions analytiques

Approche P3P : Méthode de Grunert [Haralick et al 1994]

Etape 2 : - Calculer la rotation R et la translation t satisfaisant l'équation suivante : Les 3 points $P_i, (i = 1..3)$ sont exprimés dans R_c et $P_i^j, (i = 1..3)$ sont les mêmes points dans R_o .

- Le système obtenu est sous contraint car nous disposons de 12 ddl et nous n'avons que 9 équations.
- P_i étant coplanaires, il sont supposés contenus dans le plan $Z = 0$,
- L'équation devient :
- Sa formulation matricielle donne un système de la forme $AX = B$,
- X est calculé au sens des moindres carrés,
- Une fois X calculé, le vecteur $r_3 = (r_{13}r_{23}r_{33})^t$ est calculé.

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : - Calculer la rotation R et la translation t satisfaisant l'équation suivante : Les 3 points $P_i, (i = 1..3)$ sont exprimés dans R_c et $P_i^j, (i = 1..3)$ sont les mêmes points dans R_o . Calculer l'expression des nouveaux points :

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : - P_i étant coplanaires, il sont supposés contenus dans le plan $Z = 0$, calculer leur expression :

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : - Sa formulation matricielle donne un système de la forme $AX = B$, calculer ce système d'équation :

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

- Exercice :** - X est calculé au sens des moindres carrés,
- Une fois X calculé, déterminer le vecteur $r_3 = (r_{13} r_{23} r_{33})^t$:

.....

.....

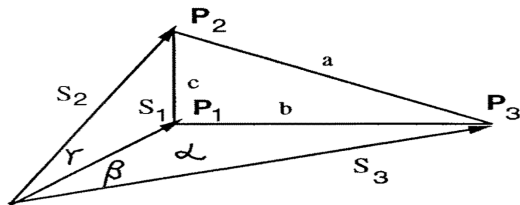
.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Problème : Calcul de la pose de la caméra/ R_o à partir de 3 points $P_i, (i = 1..3)$? , (cf. figure)



Hypothèses : Longueurs (a, b, c) du triangle $P_1P_2P_3$ connus ;
 α, β, γ connus.

Inconnues : $S_i (i = 1..3)$: longueurs des segments $O_c P_i (i = 1..3)$?

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Etape 2 : - Exprimer les relations de cosinus au sein du triangle $P_1P_2P_3$ (Al-Khachi),

- Exprimer S_2 et S_3 par combinaison linéaire de S_1 : $S_2 = uS_1$ et $S_3 = vS_1$,
- substituer dans les relations de cosinus S_2 et S_3 par S_1 ,
- Calculer le coefficient u , le substituer dans l'équation précédente,
- Déterminer les coefficients du polynôme,
- calculer les racines en v du polynôme obtenu,
- Toute solution en v permet de calculer une solution en u .
- calculer ensuite S_1 puis S_2 et S_3 .

Solutions analytiques

Approche $P3P$: Méthode de Grunert [Haralick et al 1994]

Exercice : - Exprimer les relations de cosinus au sein du triangle $P_1P_2P_3$ (Al-Khachi),

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : - Exprimer S_2 et S_3 par combinaison linéaire de S_1 : $S_2 = uS_1$
et $S_3 = vS_1$,

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : - substituer dans les relations de cosinus S_2 et S_3 par S_1 ,

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

Exercice : Calculer le coefficient u , le substituer dans l'équation précédente,,

.....

.....

.....

.....

Solutions analytiques

Approche *P3P* : Méthode de Grunert [Haralick et al 1994]

- Exercice :**
- -calculer les racines en v du polynôme obtenu,
 - -Toute solution en v permet de calculer une solution en u .
 - - calculer ensuite S_1 puis S_2 et S_3 .

.....

.....

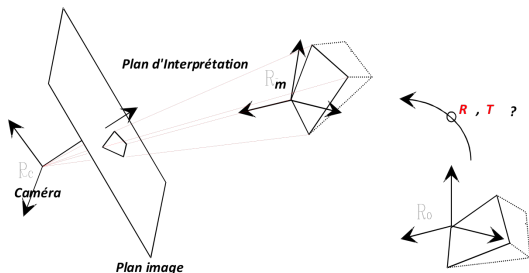
.....

.....

Solutions analytiques

Approche basée sur les segments de droites [Dhome et al 1989]

Problème : Calcul de la pose de la caméra/ R_o à partir de 3 droites $L_i (i = 1..3)$ et de leurs images $l_i (i = 1..3)$? (cf. figure)



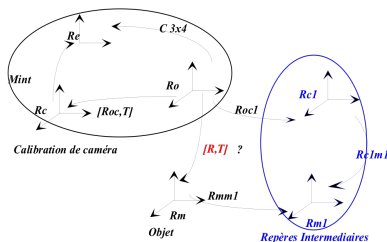
Méthode : Calculer Rotation R et Translation t ?

- avec 3 angles ayant 3 équations non linéaires (la résolution est difficile),
opter pour 2 angles en définissant des repères intermédiaires.

Solution analytique [Dhome et al 1989]

Problème

Les repères intermédiaires sont ainsi définis : (cf. figure)



Méthode :

Pour toute arête 3D $L_i, i = 1, 2, 3,$

Pour tout sommet des segments 2D $l_i, i = 1, 2,$ calculer l'équation du rayon visuel (cf. rappel ci-après) passant par P et son image p

L'application du modèle de la caméra à un point $\tilde{P} = (X, Y, Z)^t$ permet de calculer son image $\tilde{p} = (su, sv, s)^t$:

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

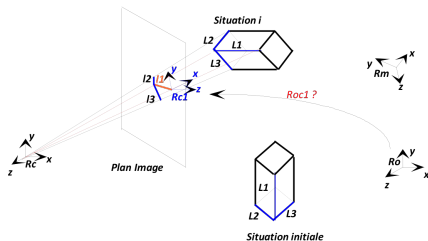
$$\Leftrightarrow \begin{cases} u = (P^t \cdot m_1 + m_{14}) / (P^t \cdot m_3 + m_{34}) \\ v = (P^t \cdot m_2 + m_{24}) / (P^t \cdot m_3 + m_{34}) \end{cases} \Leftrightarrow$$

$$\begin{cases} P^t \cdot n_u + a_u = 0 \\ P^t \cdot n_v + a_v = 0 \end{cases}$$

avec :

$$n_u = m_1 - u \cdot m_3; n_v = m_2 - v \cdot m_3; a_u = m_{14} - u \cdot m_{34}; a_v = m_{24} - u \cdot m_{34};$$

Solution analytique [?]

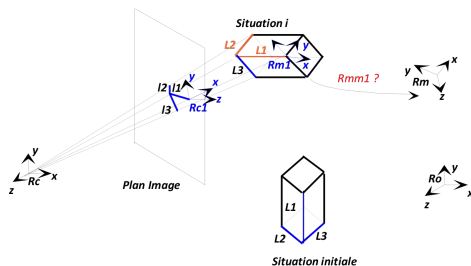


Définition de R_{c1}

$$R_{oc1} = \begin{pmatrix} (nu_{11} \times nv_{11}) & N_1 & (nu_{11} \times nv_{11}) \times N_1 \end{pmatrix}$$

$$\text{avec : } N_1 = (nu_{11} \times nv_{11}) \times (nu_{12} \times nv_{12}) \quad (2)$$

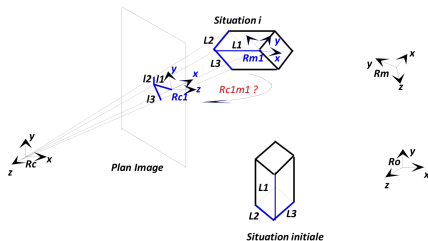
Solution analytique [?]



Définition de R_{m1}

$$R_{mm1} = \left(V_1 \quad (V_1 \times V_2) \times V_1 \quad (V_1 \times V_2) \right) \quad (3)$$

Solution analytique [?]



Méthode :

- Les axes X de (R_{c1}) et (R_{m1}) appartient au plan d'interprétation de l_1 ,
- L'axe Y de (R_{c1}) est le vecteur normal du plan d'interprétation de l_1 .

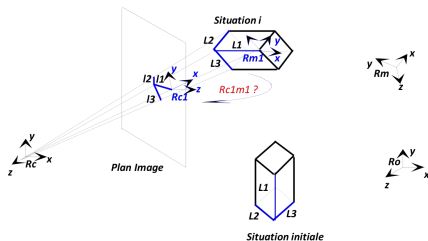
Donc :

$$\text{rotation } (y_{c1}, \theta_y) \implies x_{c1} \parallel x_{m1}$$

$$\text{rotation } (x_{c1}, \theta_x) \implies y_{c1} \parallel y_{m1}$$

Un plan d'interprétation contient le segment L_j et son image l_j .

Solution analytique [?]



Méthode :

- Les axes X de (R_{c1}) et (R_{m1}) appartient au plan d'interprétation de l_1 ,
- L'axe Y de (R_{c1}) est le vecteur normal du plan d'interprétation de l_1 .

Donc :

$$\text{rotation } (y_{c1}, \theta_y) \implies x_{c1} \parallel x_{m1}$$

$$\text{rotation } (x_{c1}, \theta_x) \implies y_{c1} \parallel y_{m1}$$

Un plan d'interprétation contient le segment L_j et son image l_j .

Solution analytique [?] : Résolution

Détermination de $[R_{c1m1}] = R_{\theta_y} R_{\theta_x}$:

$$N_{i(Rc1)} \cdot V_{i(Rc1)} = 0 (i = 1..3)$$

$$\iff [R_{c1o} \cdot N_{i(Ro)}] \cdot [R_{c1m1} \cdot R_{m1m} \cdot V_{i(Rm)}] = 0(4)$$

- $R_{m1m} \cdot V_1$ est invariant en fonction de θ_x ;
- Aussi, l'équation donnée par V_1 n'est pas utilisée pour le calcul de θ_x .
- V_2 et V_3 sont utilisés. Ils doivent être différents et non parallèles à L_1 .

L'équation (4) permet d'aboutir à un polynôme de degré 8,

-Sa résolution se fait par le calcul des valeurs propres de la matrice

Compagnon :

$$R = R_{om} = R_{oc1} \cdot R_{c1m1} \cdot R_{m1m}$$

Solution analytique [?] : Résolution

Détermination de t , translation de R_o à R_m

$$N_{i(Rc)} \cdot (t + R_{cm} \cdot M_{Rm}) = 0 (i = 1..3)$$

$$\iff N_{i(Ro)} \cdot (t + R_{om} \cdot M_{Rm} - C_{Ro}) = 0 (5)$$

Il s'agit de la résolution d'un système linéaire.

Solutions numériques : nb de correspondances est important > 6

Diverses approches

- [Ganapathy 1984] : solution linéaire,
- [Yuan 1989] : séparation translation et rotation,
- [Kuman et all 1989] : correspondance de triplet et linéarisation de la fonction erreur,
- [Liu 1990] : méthode itérative pour faible amplitude angulaire,
- [Lowe1991] : critère quadratique et méthode de Newton, ...

Diverses méthodes de résolution :

- Moindres carrés,
- Levenberg-Marquardt,
- Filtrage de Kalman,
- Filtrage particulaire,

...

Solution : Le système représentant le rayon visuel d'un modèle *pinhole* (8) est rappelé ci-après :

$$\begin{cases} P^t \cdot n_u + a_u = 0 \\ P^t \cdot n_v + a_v = 0 \end{cases}$$

Après dérivation (cas de faible amplitude) :

$$\begin{cases} dnu \cdot P + nu \cdot dP + da_u = 0 \\ dnv \cdot P + nv \cdot dP + da_v = 0 \end{cases}$$

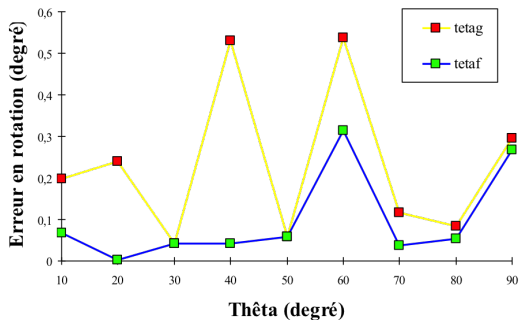
La substitution de dP , et le produit mixte de vecteurs donnent :

$$\begin{cases} nu \cdot T + (P \times nu) \cdot R + dnu \cdot P + da_u = 0 \\ nv \cdot T + (P \times nv) \cdot R + dnv \cdot P + da_v = 0 \end{cases}$$

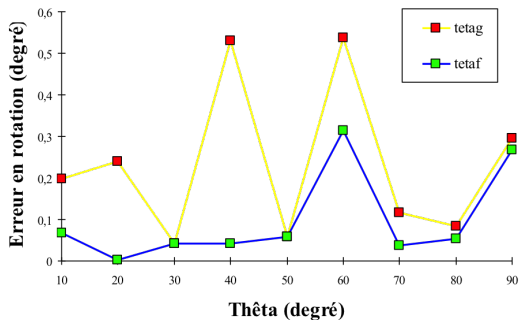
Le système obtenu peut être résolu par une méthode numérique :

$$A \cdot X = B \text{ où } X = (t_x, t_y, t_z, R_x, R_y, R_z)^t$$

Expérimentation : résultat - Méthodes analytique et numérique Faible amplitude



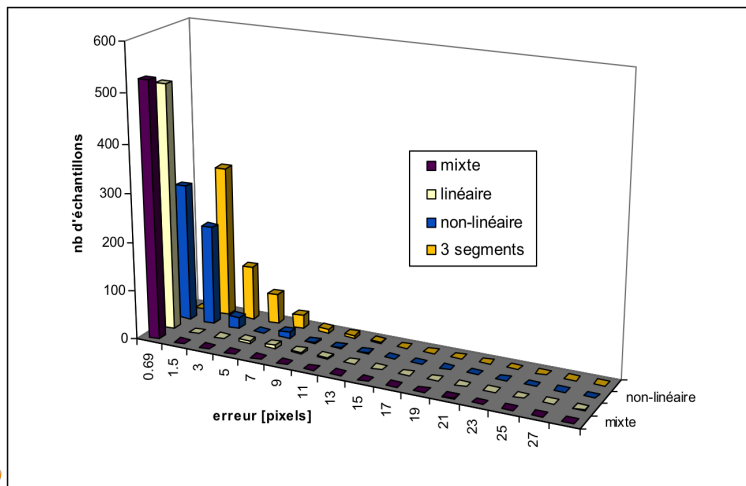
Expérimentation : résultat - Méthodes analytique numérique Faible amplitude



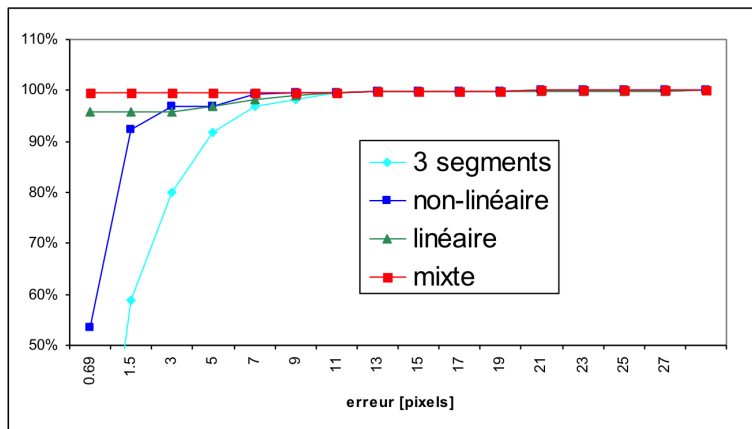
Recalage mixte - algorithme

1. Appliquer la méthode des trois segments,
2. Utiliser la méthode linéaire pour affiner le recalage,
 - si le critère d'erreur est suffisamment petit : le résultat est considéré bon, fin.
 - sinon : position singulière, continuer,
3. Appliquer la méthode non-linéaire,
 - si elle ne converge pas : pas de solution, fin.
 - sinon : sauver le résultat, continuer,
4. Appliquer à nouveau la méthode linéaire,
 - si le critère d'erreur est suffisamment petit : le résultat est considéré bon, fin.

Expérimentation : résultat - Méthode mixte



Expérimentation : résultat - Méthode mixte



Questions sur le cours

Q1 : Quelle différence y a t il entre les méthodes analytiques et les méthodes numériques ?

Q2 : Est ce que la calibration des capteurs est toujours nécessaires Pour effectuer le recalage ?

Q3 : Quels sont les avantages et limitations du recalage fin ?

Q4 : Quelle méthode de recalage est la plus appropriée si Le critère prioritaire est le temps réel ?

Q5 : Quel est l'apport majeur du calcul de pose ?

Bibliographie

- [Dhome et al., 1989] M. Dhome, M. Richetin, et J.-T. Lapreste (1989). Determination of the attitude of 3d objects from a single perspective view. IEEE Trans. Pattern Anal. Mach. Intell., 11(12) :1265-1278.
- DeMenthon et Davis, 1992D. DeMenthon et L.S. Davis (1992). Exact and approximate solutions of the perspective-three-point problem. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(11) :1100-1105.
- Haralick et al, 1994Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem. International Journal of Computer Vision, 13, 3, 331-356 (1994) Kluwer Academic Publishers. Manufactured in The Netherlands.

Bibliographie

[Lu et al., 2000] C.-P. Lu, G.D. Hager, et E. Mjolsness (2000). Fast and globally convergent pose estimation from video images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(6) :610-622.

Lowe, 1992D. Lowe (1992). Robust model-based motion tracking through the integration of search and estimation. *International Journal of Computer Vision*, 8(2) : 113-122.

R.L. Carceroni, C.M. Brown 1997 Numerical Methods for Model Based Pose Recovery University of Rochester, Technical Report 659, August 1997