

Calibrage de caméra

Michel Roux

Institut Mines Télécom, Télécom ParisTech

16 mars 2023



Acquisition d'une image

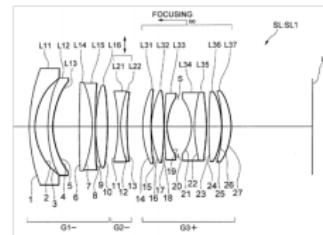
Objectifs de ce cours :

- établir la relation géométrique qui existe entre une scène 3D et son image,



- décrire les applications qui nécessitent le calibrage de la caméra :
 - le recalage d'images,
 - la création d'un panorama à partir de plusieurs images,
 - la stéréovision,
 - ...
- géométrie projective
 - faire des mesures de profondeur avec une seule image,

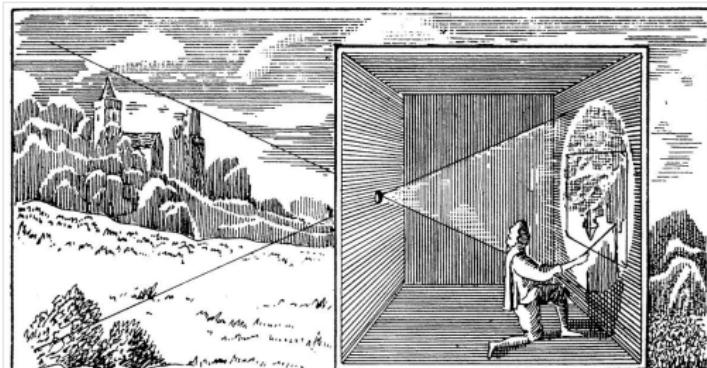
L'appareil photographique



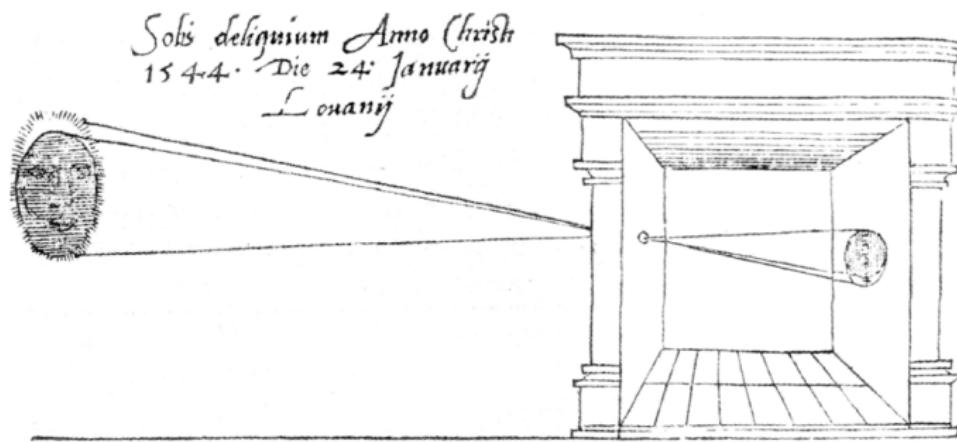
d'où l'utilisation d'un modèle approché de l'acquisition des images

Camera obscura

- IV^e siècle avant J-C en Chine, Mozi, philosophe, Mohist School of Logic.
 - Aristote au IV^e siècle avant J-C ne trouve pas les réponses aux questions qu'il se pose à ce sujet.
 - Ibn al-Haythan (965-1039), Kitab al-Manazir, traité d'optique traduit en latin et en grec.
 - En 1514, Léonard de Vinci : *En laissant les images des objets éclairés pénétrer par un petit trou dans une chambre très obscure tu intercepteras alors ces images sur une feuille blanche placée dans cette chambre. [...] mais ils seront plus petits et renversés.*
 - Surtout utilisé à partir du XVI^e siècle.



Camera obscura



The Great Picture

2006, 34m x 10m, Irvine Californie, hangar à avions, focale : 24m, diamètre : 6mm,
exposition : 35mn.

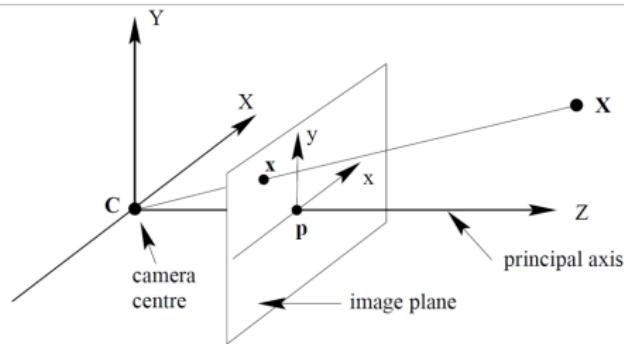


Projection perspective : le modèle sténopé

En pratique, nous considérons des images acquises avec un objectif à petite ouverture, donc une grande profondeur de champs, pour lesquelles le modèle de la "Camera Obscura" est suffisant. Il s'agit du modèle sténopé.

Dans ce modèle, les rayons optiques sont des droites qui relient les points 3D de la scène aux pixels de l'image en passant par le centre optique.

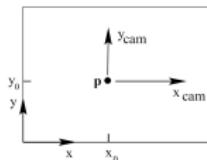
L'essentiel de la suite repose sur la colinéarité des vecteurs Cx et CX .



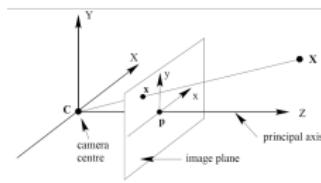
Les espaces utilisés : liaison entre espace 3D et espace 2D

- l'espace objet, c'est un espace 3D dans lequel est décrite la scène, un point 3D M y est repéré par ses coordonnées cartésiennes, données en mm,
- l'espace image, un pixel y est repéré par ses deux indices de ligne et de colonne, donnés en pixels,
- l'espace caméra, il permet de faire la liaison entre les deux espaces précédents, mais c'est là où tout se passe : projection de la scène 3D sur un plan.

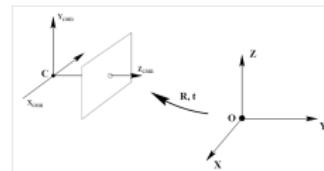
Espaces / transformations



espace image



espace caméra



espace objet

- espace image : translation et homothétie 2D
- espace caméra : projection 3D / 2D (non-linéaire)
- espace objet : translation et rotation 3D

Comment simplifier les calculs ?

Que voulons-nous ?

- notation simple et concise,
- unification des transformations : translations, rotations, projections
- possibilité de concaténer des transformations,
- et si possible, rendre les transformations linéaires.

Une solution : la géométrie projective et les coordonnées homogènes

Intérêt de la géométrie projective

- principe de dualité : points et droites sont équivalents,
- coordonnées homogènes,
- algèbre linéaire,
- géométrie euclidienne comme cas particulier
- compréhension de la structure d'ensemble de droites, de rotations, ...
- beaucoup de capteurs réalisent une projection du monde 3D sur un plan,
- applications multiples en robotique, vision par ordinateur, synthèse d'image, ...

Espace projectif

Définitions :

- \mathcal{P}^n : espace projectif de dimension n ,
- point x est représenté par un vecteur de $n+1$ coordonnées :
 $x = (x_1, x_2, \dots, x_{n+1})^t$, avec au moins une coordonnée x_i non nulle
(coordonnées homogènes),
- $(x_1, x_2, \dots, x_{n+1})^t$ et $(y_1, y_2, \dots, y_{n+1})^t$ représentent le même point
 $\iff \exists \lambda \neq 0, \forall i, x_i = \lambda y_i$;
- point pour lequel $x_{n+1} = 0$: point à l'infini

Espace projectif = espace quotient :

- $\mathcal{P}^n = (E^{n+1} \setminus \{0\}) / \sim$ avec E^{n+1} espace vectoriel de dimension $n+1$
et $x \sim x' \iff \exists \lambda \neq 0, x = \lambda x'$
- vecteur non nul = droite passant par l'origine de E^{n+1}
- deux vecteurs sont équivalents s'ils définissent la même droite
- point projectif (élément de \mathcal{P}^n) : classe d'équivalence de vecteurs.

Projection canonique et bases

Projection canonique :

$$\begin{aligned} p : E^{n+1} \setminus \{0\} &\longrightarrow \mathcal{P}^n \\ \vec{x} &\longmapsto p(\vec{x}) = x \end{aligned}$$

p n'est pas bijective !

Base projective :

- base $\vec{a}_i, i = 1 \dots n + 1$ de E^{n+1}
- points projectifs associés : $m_i = p(\vec{a}_i)$
- ne détermine pas les coordonnées projectives d'un point x ($n + 1$ paramètres),
- l'ajout d'un $(n + 2)^{\text{ème}}$ point dans la base $m_{m+2} = p(\sum \vec{a}_i)$ permet de n'avoir plus qu'un paramètre.

Projection canonique et bases

Base canonique :

$$e_i = (0, \dots, 1, \dots, 0)^t$$

pour $i = 1, \dots, n+1$, avec un 1 à la i-ème position,

$$e_{n+2} = (1, \dots, 1, \dots, 1)^t$$

Un point de l'espace projectif peut s'écrire comme combinaison linéaire de $n+1$ points parmi les $n+2$ points de la base :

$$x = \sum_{i=1}^{n+1} x_i e_i$$

Plan projectif \mathcal{P}^2

- point : $x = x_1e_1 + x_2e_2 + x_3e_3$ avec les x_i non tous nuls
- droite d : définie par un triplet de nombres (d_1, d_2, d_3) non tous nuls
- x appartient à $d \iff d^t x = 0 \iff \sum_{i=1}^3 d_i x_i = 0$
- la même équation définit également l'ensemble des droites qui passent par x ,
- dualité entre point et droite
- droite passant par deux points x et y : $z = \alpha x + \beta y$, donc $\text{Det}(z, x, y) = 0$, soit $z^t(x \wedge y) = 0 \implies$ droite $d = x \wedge y$
- point d'intersection de deux droites d et d' : $d \wedge d'$
- deux droite s'intersectent toujours
- $x_3 = 0$: équation de la droite à l'infini

Espace projectif \mathcal{P}^3

- point : $x = \sum_{i=1}^4 x_i$ avec les x_i non tous nuls,
- plan π : défini par quatre nombres non nuls $(\pi_1, \pi_2, \pi_3, \pi_4)$
- équation du plan : $\pi^t x = 0$
- dualité entre point et plan
- droite de \mathcal{P}^3 : définie soit à partir de deux points, soit comme intersection de deux plans
- plan à l'infini obtenu avec l'équation $x_4 = 0$
- tout plan intersecte le plan à l'infini selon une droite à l'infini

Applications linéaires de l'espace projectif

$$H : \mathcal{P}^m \longrightarrow \mathcal{P}^n$$

- linéaire
- définie par une matrice $(m+1) \times (n+1)$,
- de rang plein
- définie à un facteur non nul près
- exemple particulier : changement de base H inversible $(n+1) \times (n+1)$

Exemple de transformation

Translation

$$\begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Exemple de transformation

Rotation autour de x

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

Exemple de transformation

Homothétie

$$\begin{pmatrix} \lambda_x & 0 & 0 & 0 \\ 0 & \lambda_y & 0 & 0 \\ 0 & 0 & \lambda_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Exemple de transformation

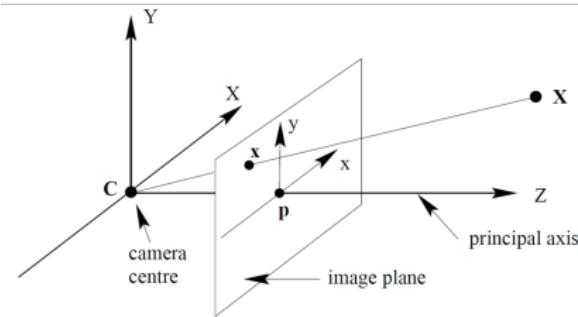
Projection de \mathcal{P}^3 sur \mathcal{P}^2

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

En coordonnées affines : transformation non-linéaire

Projection perspective : le modèle sténopé

Colinéarité des vecteurs Cx et CX dans l'espace caméra.



$$\lambda \begin{pmatrix} x_{cam} \\ y_{cam} \\ f \end{pmatrix} = \begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{pmatrix}$$

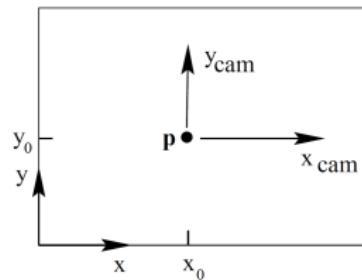
soit, en coordonnées homogènes :

d'où

$$\begin{aligned} x_{cam} &= fX_{cam}/Z_{cam} \\ y_{cam} &= fY_{cam}/Z_{cam} \end{aligned}$$

$$\begin{pmatrix} s \ x_{cam} \\ s \ y_{cam} \\ s \end{pmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{pmatrix}$$

Coordonnées image



$$k_x x_{cam} = x - x_0$$
$$k_y y_{cam} = y - y_0$$

avec k_x, k_y : nombre de pixels par mm

$$\begin{pmatrix} s \cdot x \\ s \cdot y \\ s \end{pmatrix} = \begin{bmatrix} k_x & 0 & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} s \cdot x_{cam} \\ s \cdot y_{cam} \\ s \end{pmatrix}$$

Paramètres intrinsèques

$$x = \begin{pmatrix} s \ x \\ s \ y \\ s \end{pmatrix} = K \begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{pmatrix}$$

K : matrice 3×3 , triangulaire supérieure, matrice de calibrage de la caméra. En posant $f_x = f \ k_x$ et $f_y = f \ k_y$:

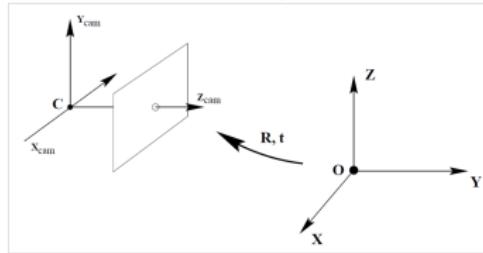
$$K = \begin{bmatrix} f_x & \sigma & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Paramètres intrinsèques de la caméra :

- f_x, f_y : focales en x et y , s'expriment en *pixels* (scaling),
- x_0, y_0 : position du point principal, intersection de l'axe optique et du plan focal (en *pixels*),
- σ : *skew*, suivant l'angle entre les axes x et y , $\sigma = -f_x \cotg(\theta)$, où θ est l'angle entre les axes x et y .

Espace objet

Paramètres extrinsèques : rotation et translation 3D



$$\begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{pmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Transformation rigide entre les coordonnées caméra et l'espace objet.

- R : matrice de rotation 3×3
- t : vecteur de translation 3×1

Matrice de projection

Concaténation des matrices

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = K \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = K[R|t]X$$

qui définit la matrice 3×4 de projection d'un espace euclidien 3D vers une image.

$$x = P\mathbf{X}$$

avec

$$P = K[R|t] = K \ R \begin{bmatrix} I | R^T t \end{bmatrix}$$

Le centre de la caméra : $(X, Y, Z)^T = -R^T t$

Caméra projective

Le modèle de caméra pour une projection perspective est une transformation linéaire entre coordonnées homogènes.

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = [P(3 \times 4)] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

point de l'espace objet (en *mm*) → point de l'image (en *pixels*)

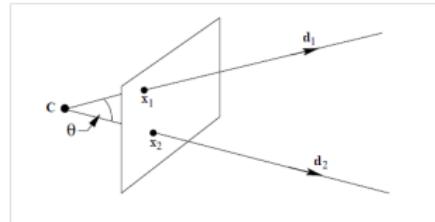
- Le centre de la caméra est le vecteur nul de P
- P a 11 degrés de liberté
- P est de rang 3.

Calibrage de la caméra

Revenons à la matrice K des paramètres intrinsèques.

- Si K est connue, on dit que la caméra est calibrée,
- K permet d'obtenir, dans l'espace caméra (espace objet avec $P = [I|0]$), la direction du rayon lumineux dont est issu un pixel de l'image,
- Une caméra calibrée permet de calculer l'angle entre deux rayons.

$$\begin{pmatrix} sx \\ sy \\ s \end{pmatrix} = \begin{bmatrix} k_x & 0 & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \end{pmatrix}$$



Calibrage d'une caméra : estimation de P

Problème :

- étant données n correspondances $(x_i, X_i)_{i=1\dots n}$ entre des pixels de l'image et des points de l'espace objet,
- calculer la matrice de projection $P = K [R|t]$ telle que $\forall i = 1\dots n, \quad x_i = P X_i.$

La méthode de calibrage de la caméra se décompose en 2 parties :

- calcul de la matrice P à partir des n appariements,
- décomposition de P en K , R et t à partir d'une décomposition QR .

Calibrage d'une caméra : étape 1

$$x_i = P X_i$$

Chaque correspondance donne deux équations :

$$x_i = \frac{p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

$$y_i = \frac{p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}}{p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}}$$

Ce qui se réécrit :

$$\begin{aligned} x_i(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}) &= p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14} \\ y_i(p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34}) &= p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24} \end{aligned}$$

Soit :

$$\left(\begin{array}{ccccccccc} X & Y & Z & 1 & 0 & 0 & 0 & -xX & -xY & -xZ & -x \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -yX & -yY & -yZ & -y \end{array} \right) p = 0$$

avec $p = (p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}, p_{31}, p_{32}, p_{33}, p_{34})$ le vecteur des 12 paramètres à estimer.

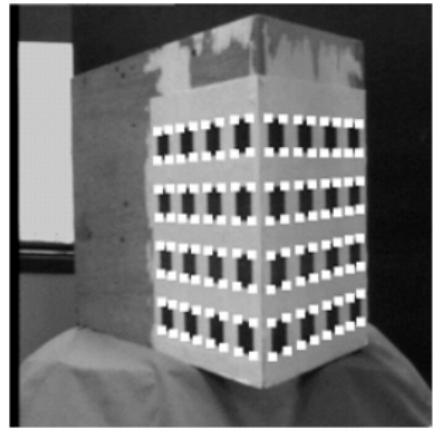
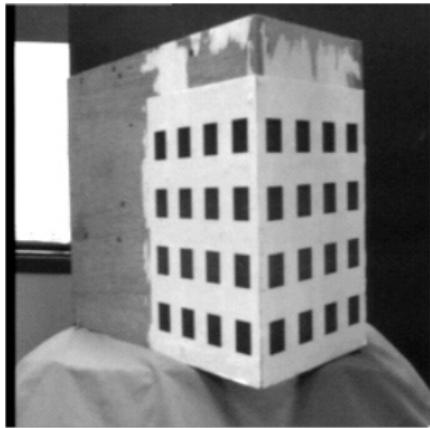
Calibrage d'une caméra : étape 1 (suite)

Calcul de p

- Pour ($n \geq 6$) correspondances, on obtient un système à $2n$ équations et 12 inconnues sous la forme : $Ap = 0$, où A est une matrice $2n \times 12$.
- Par régression linéaire, une solution qui minimise $|Ap|$, avec la contrainte $|p| = 1$ est obtenue à partir du vecteur propre associé à la plus petite valeur propre de $A^T A$.
- La solution précédente est utilisée pour initialiser une minimisation non-linéaire portant sur la différence entre les pixels mesurés et les points 3D projetés dans l'image :

$$\min_p \sum_i ((x_i, y_i) - P(X_i, Y_i, Z_i, 1))^2$$

Exemple de calibrage : mire



- Déterminer la position précise des coins
 - détection de contours,
 - estimation des droites,
 - intersection des lignes.
- L'erreur finale entre les points mesurés dans l'image et les points reprojetés est inférieure à 0.02 pixels.

Calibrage d'une caméra : étape 2

Décomposition de P en K , R et t

La première sous-matrice 3×3 de P est le produit d'une matrice triangulaire supérieure K et d'une matrice de rotation R .

- Factorisation $M = KR$ par décomposition "RQ"
- et pour finir :

$$t = K^{-1}(p_{14}, p_{24}, p_{34})$$

- Remarque

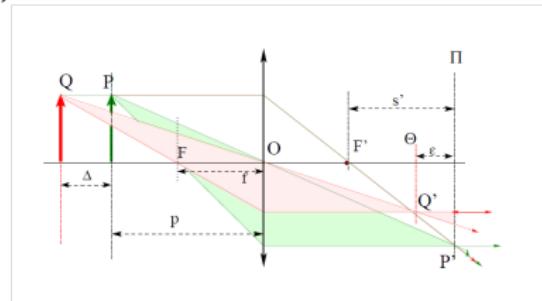
On trouve K sous la forme :

$$K = \begin{bmatrix} f_x & \sigma & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

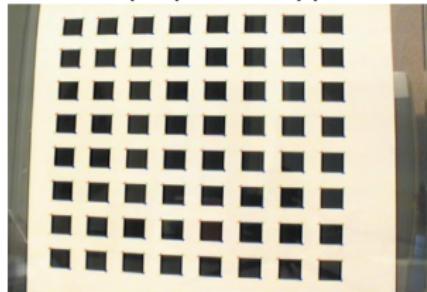
Limitations du modèle de caméra

Le modèle de la “Camera Obscura” ne permet pas de rendre compte de :

- la profondeur de champ (flou artistique) : modèle de lentille mince

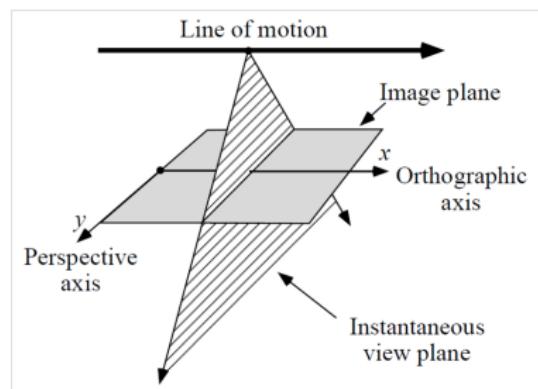


- la distorsion des images due à l'optique de l'appareil.



Caméra à défilement (pushbroom)

Les caméras satellites sont généralement équipées d'une seule barrette de capteurs.



L'acquisition de l'image se fait ligne par ligne.

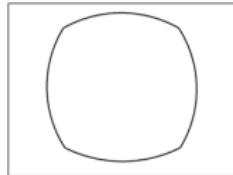
C'est le déplacement du satellite qui engendre la seconde dimension de l'image.

Conséquence : la position et l'orientation du capteur varient au cours du temps lors de l'acquisition de l'image.

Distorsion géométrique

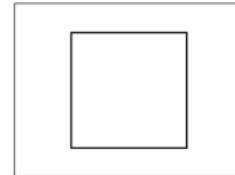


radial distortion



correction

linear image



Distorsion géométrique



image distordue

Distorsion géométrique



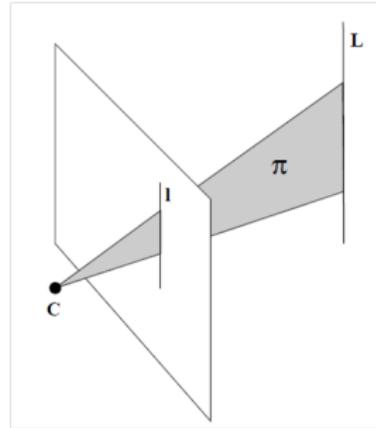
image corrigée

Applications

Le calibrage de caméra va permettre de traiter différentes applications :

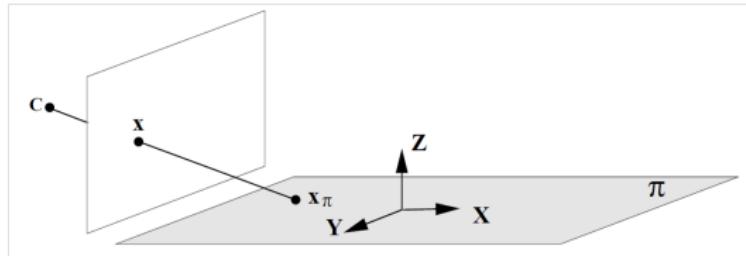
- superposition d'images pour la détection de changement
- création de panoramas,
- corrections de perspective
- redressement de surfaces planes,
- ...

Projection d'une droite, et inversement



- Une droite L de l'espace objet se projette sur la droite l de l'image, intersection du plan focal et du plan π , défini par la droite L et le centre optique C .
- Inversement, l'ensemble des points de l'espace objet, qui se projettent sur la droite l , est le plan $\pi = P^T l$.

Projection d'un plan



- Considérons le plan défini par $Z = 0$,
- Un point X de ce plan se projette en un point x de l'image,
- Soit :

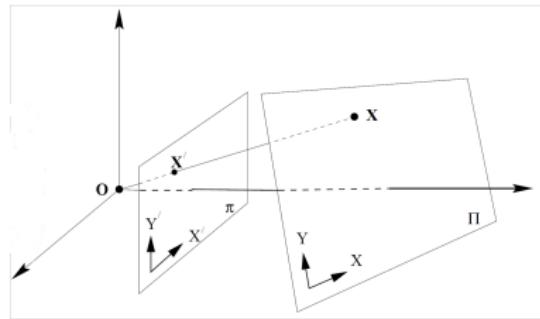
$$x = PX = [p_1 \ p_2 \ p_3 \ p_4] \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = [p_1 \ p_2 \ p_4] \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

qui est l'équation d'une homographie planaire.

Projection d'un plan

Superposition de deux images d'une surface plane acquises de deux positions différentes.

Projection d'un plan



$$x' = Hx$$

où H est une homographie planaire (matrice 3×3 , homogène, non-singulière).

Projection d'un plan



Projection d'un plan



Projection d'un plan

Correction de perspective
Sélection manuelle des points



Image originale



Correction de la perspective

Projection d'un plan

Superposition de deux images d'une même surface plane acquises de deux positions différentes.

- détection de points caractéristiques dans les deux images,
- mise en correspondance des points caractéristiques
- estimation de l'homographie par une méthode de type RANSAC,
- superposition de l'image transformée sur l'autre image.



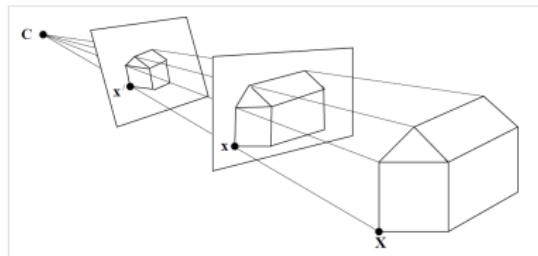
Image de face



Image de côté

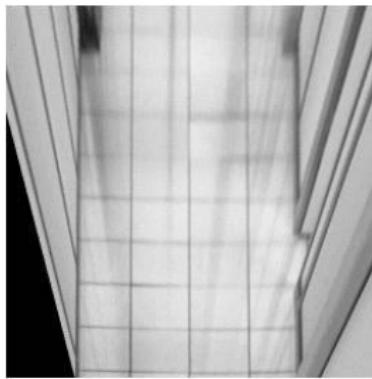


Cône de rayons



- Une image est l'intersection du cône de rayons qui passent par le centre optique et d'un plan.
- Deux images, qui ont le même centre optique, sont reliées par une homographie planaire.

Correction de perspective



Rotation vs. translation

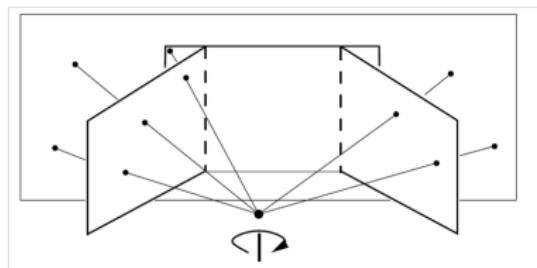
Voici 3 images.

Quel couple d'images permet de produire un panorama et quel couple permet de produire des données 3D ?



Mosaïque d'images : panorama

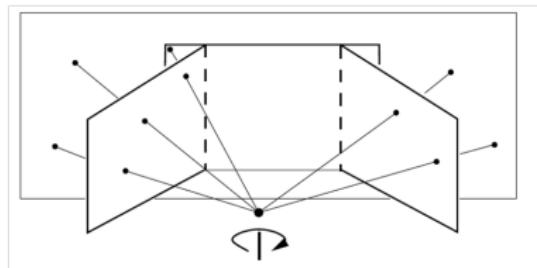
Construction d'un panorama



- Superposition des images après correction géométrique
- Rotation de la caméra autour de son centre optique (ou scène plane).
- Les images sont reliées l'une à l'autre par une transformation homographique.

Mosaïque d'images : panorama

Construction d'un panorama



Estimation des homographies

- $\lambda x' = Hx \implies x' \wedge Hx = 0$
- chaque correspondance (x, x') donne deux équations indépendantes,
- Quatre correspondances suffisent à estimer H , mais il est possible d'en utiliser plus avec une méthode d'estimation robuste à la présence de données aberrantes.

Mosaïque d'images : panorama



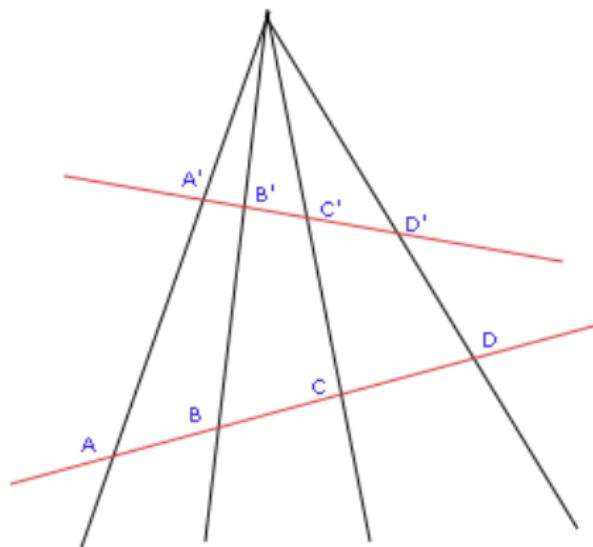
Mosaïque d'images : panorama



Mosaïque d'images : panorama

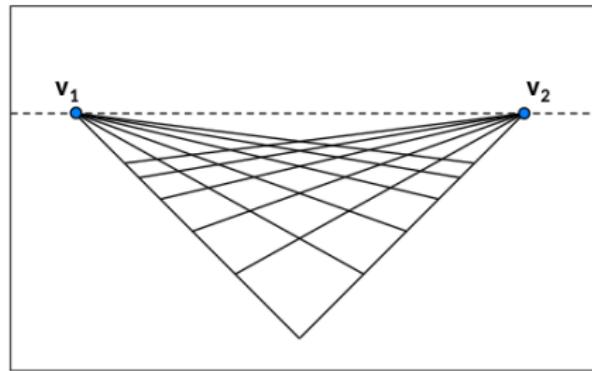


Cross ratio

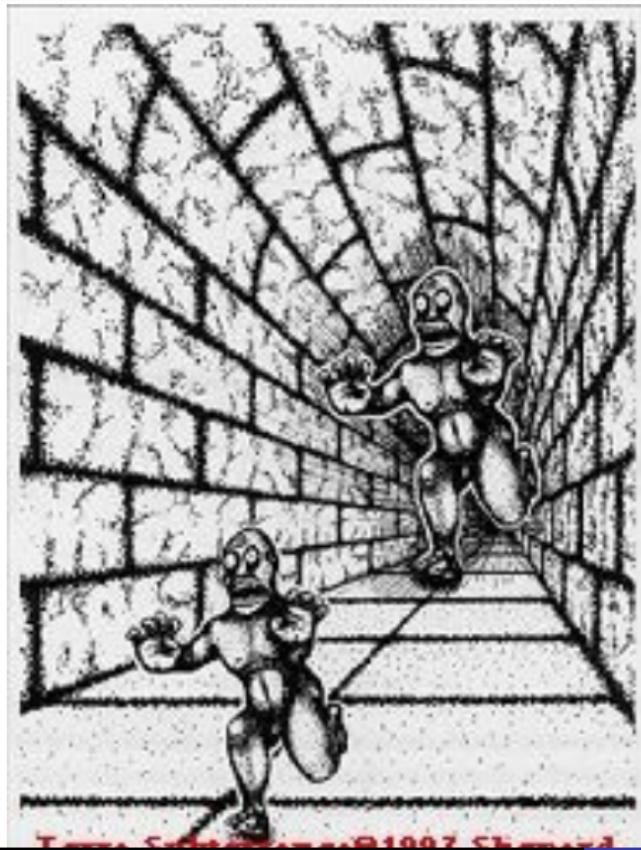


Le bi-rapport $\frac{CA}{CB} \frac{DA}{DB}$ est invariant par projection centrale (projection conique).

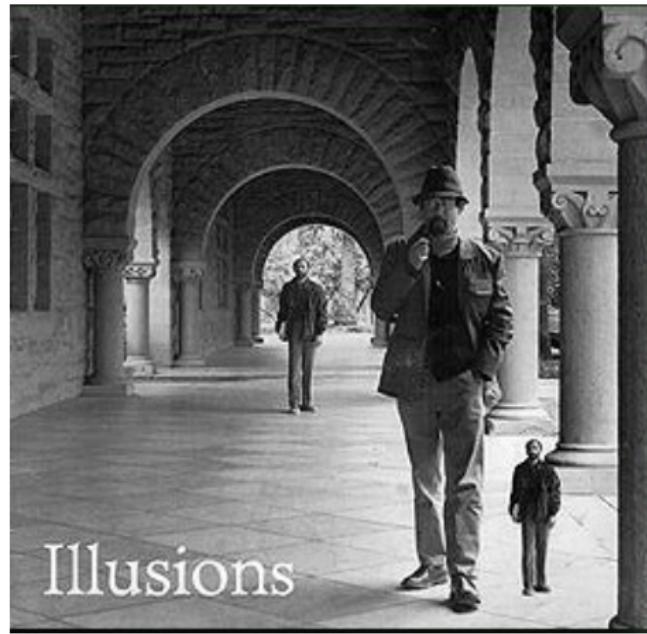
Vanishing points



Vanishing points



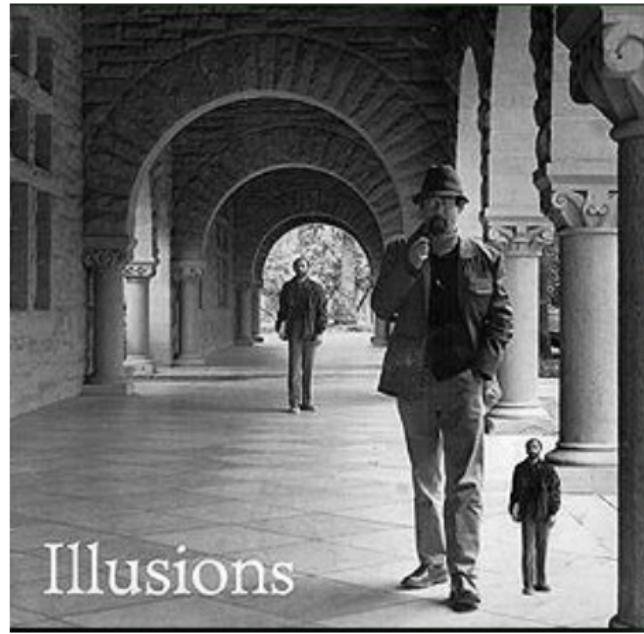
Vanishing points



Illusions

Le bi-rapport $\frac{CA}{CB} \frac{DA}{DB}$ est invariant par projection centrale (projection conique).

Vanishing points



Illusions

Conclusion

- La matrice de calibrage, ou matrice intrinsèque, $K(3 \times 3)$ dépend uniquement des paramètres internes de la caméra,
- la matrice de projection, $P(3 \times 4)$ dépend de K et des paramètres de position et d'orientation de la caméra,
- les coordonnées homogènes permettent de rendre linéaires les équations du calibrage de caméra.
- une relation homographique existe entre :
 - 2 images acquises en rotation autour du centre optique,
 - entre les points d'un plan acquis dans deux images quelconques,
- il suffit de 4 correspondances pour estimer une homographie (points non-alignés).

Bibliographie

- Olivier Faugeras , *Three-Dimensional Computer Vision A Geometric Viewpoint*, MIT Press, Cambridge, 1993.
- Kyoji Narai, Yoshiya Matsui, *Fundamentals of Practical Aberration Theory*, World Scientific, 1993.
- José-Philippe Pérez, *Optique géométrique, ondulatoire et polarisation*, Masson, Paris, 4ème édition, 1994.
- Z. Zhang, *A flexible new technique for camera calibration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11) :1330-1334, 2000.
- Michel Dhome, *Perception visuelle par imagerie vidéo*, Traité IC2, série Signal et image, Hermès-Lavoisier, Paris 2003.
- Richard Hartley, Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, 2004.