



Eléments de stéréo-restitution

Michel Roux

GET-ENST, département TSI, 46 rue Barrault, 75634 Paris Cedex 13
Tél. 01 45 81 81 28, Fax : 01 45 81 37 94, Email : michel.roux@enst.fr

1 Stéréovision

1.1 Une image : modèle d'acquisition

Le modèle géométrique pour l'acquisition des images repose sur la propriété du centre optique : chaque point m de l'image A est l'intersection du plan image (plan rétinien à la distance focale) avec une droite passant par le centre optique C . Le point 3D M de la scène dont est issu le point image m se trouve sur la droite (m, C) .

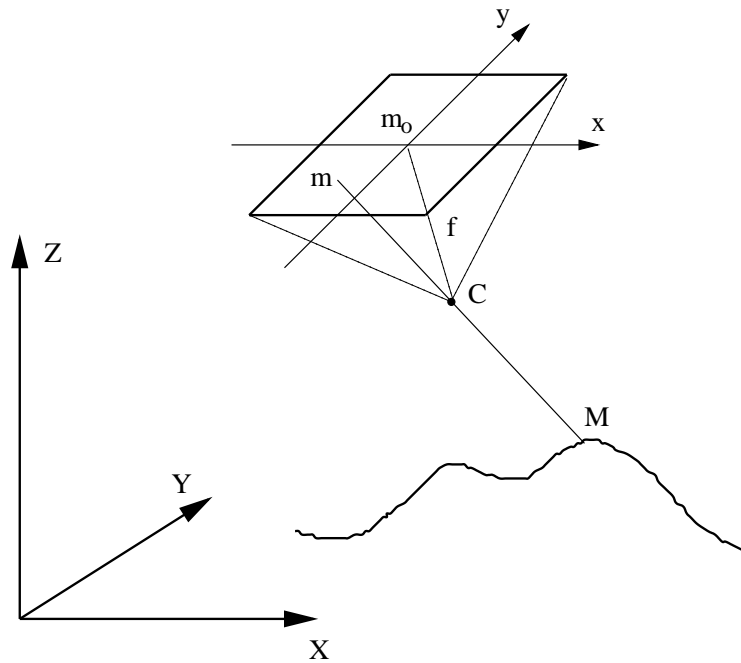


FIG. 1 – Le modèle géométrique d'une caméra

La relation entre ces points s'écrit comme :

$$\begin{bmatrix} X_M \\ Y_M \\ Z_M \end{bmatrix} = \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} + \alpha \cdot R \begin{bmatrix} x \\ y \\ -f \end{bmatrix} \quad (1)$$

où R est une matrice de rotation 3D et α un coefficient d'homothétie.

L'élimination du facteur d'échelle α de ces équations conduit aux équations de colinéarité :

$$\begin{aligned} x_m &= -f \frac{r_{11}(X_M - X_C) + r_{12}(Y_M - Y_C) + r_{13}(Z_M - Z_C)}{r_{31}(X_M - X_C) + r_{32}(Y_M - Y_C) + r_{33}(Z_M - Z_C)} \\ y_m &= -f \frac{r_{21}(X_M - X_C) + r_{22}(Y_M - Y_C) + r_{23}(Z_M - Z_C)}{r_{31}(X_M - X_C) + r_{32}(Y_M - Y_C) + r_{33}(Z_M - Z_C)} \end{aligned}$$

Ces équations expriment le fait que les 3 points M , C et m sont sur la même droite.

1.2 orientation absolue

Si la position du centre optique C et la distance focale sont connues, les équations de colinéarité deviennent :

$$\begin{aligned} x_m &= \frac{a_{11}X_M + a_{12}Y_M + a_{13}Z_M + a_{14}}{a_{31}X_M + a_{32}Y_M + a_{33}Z_M + a_{34}} \\ y_m &= \frac{a_{21}X_M + a_{22}Y_M + a_{23}Z_M + a_{24}}{a_{31}X_M + a_{32}Y_M + a_{33}Z_M + a_{34}} \end{aligned}$$

L'orientation absolue d'une image consiste alors à déterminer les paramètres a_{ij} . Si N points de contrôle sont disponibles (i.e., des points dont les coordonnées (X, Y, Z) dans l'espace objet et les coordonnées (x, y) dans l'espace image sont connues), nous obtenons un système à $2.N$ équations avec 11 inconnues et 1 degré de liberté. Au moins 6 points de contrôle sont donc nécessaires pour résoudre ce système :

- $a_{34} = 1$: approximation aux moindres carrés,
- $\|a_3\| = 1$: méthode Faugeras-Toscani, avec $\|a_3\| = \sqrt{a_{31}^2 + a_{32}^2 + a_{33}^2}$.

2 Deux images

La perception du relief est capacité relativement complexe du système visuel humain, qui met en jeu différents types d'information : la taille, la forme, l'ombrage, le mouvement, et bien sur l'information fournie par la combinaison des deux images perçues par nos yeux.

La capacité de percevoir le relief à partir de deux images légèrement différentes s'appelle la *stéréovision*. Dans le domaine de la vision par ordinateur, les méthodes développées pour la stéréovision sont fondées sur les mêmes principes géométriques que la vision humaine : un point P se projette en P_A dans l'image A et en P_B dans l'image B , la différence de position relative entre P_A et P_B est appelée *disparité* et est inversement proportionnelle à la distance entre le point P et les plans images.

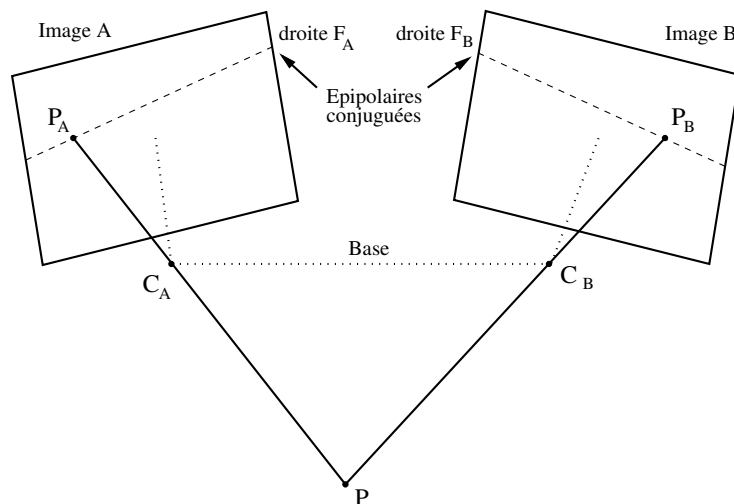


FIG. 2 – Pour la stéréorestitution, deux images sont acquises de deux points de vue différents. Un point P de la scène se projette sur chaque image à l'intersection du plan image et de la droite joignant le point P et le centre optique C_A ou C_B .

2.1 Géométrie épipolaire

Sur la figure 2, les intersections du plan défini par le triplet (P, C_A, C_B) avec les deux plans images définissent une paire de droites conjuguées, appelées droites épipolaires.

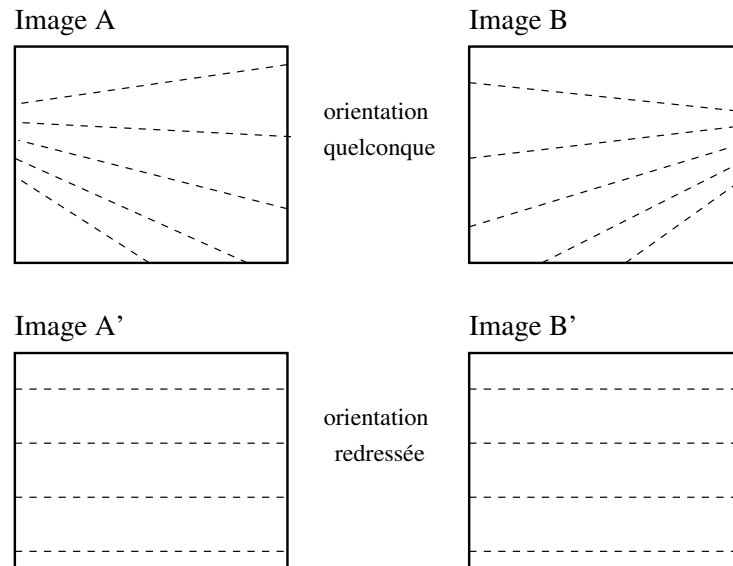


FIG. 3 – Couple d'images en géométrie quelconque (images A et B) et couple d'images en géométrie épipolaire (images A' et B').

Pour des images non-rectifiées, l'ensemble des droites épipolaires d'une image forment un faisceau qui converge en un point, appelé épipôle, qui est l'intersection de la base (C_A, C_B) avec le plan image. Pour des images redressées en géométrie épipolaire, les droites épipolaires sont parallèles (les épipoles sont à l'infini).

2.2 Orientation relative

Le but de l'orientation relative est de déterminer la position relative des points des deux images. Si nous considérons la figure 4, cette opération consiste à calculer les matrices de rotations R_A et R_B qui permettent de passer respectivement du système de coordonnées image \mathcal{R}_A et \mathcal{R}_B au système de coordonnées de l'espace objet \mathcal{R} .

Si nous considérons deux points conjugués P_A et P_B , alors les 4 points C_A , C_B , P_A et P_B sont coplanaires, ce qui signifie que le produit vectoriel des 3 vecteurs $(C_A C_B, C_A P_A, C_B P_B)$ est égal à 0. Ceci conduit à une équation non-linéaire à 5 inconnues (il y a un degré de liberté parmi les 6 angles de rotation) qui peut être linéarisée à l'aide de la formule de Taylor. Au moins 5 paires de points conjugués sont nécessaires pour résoudre ce système de façon itérative au sens des moindres carrés.

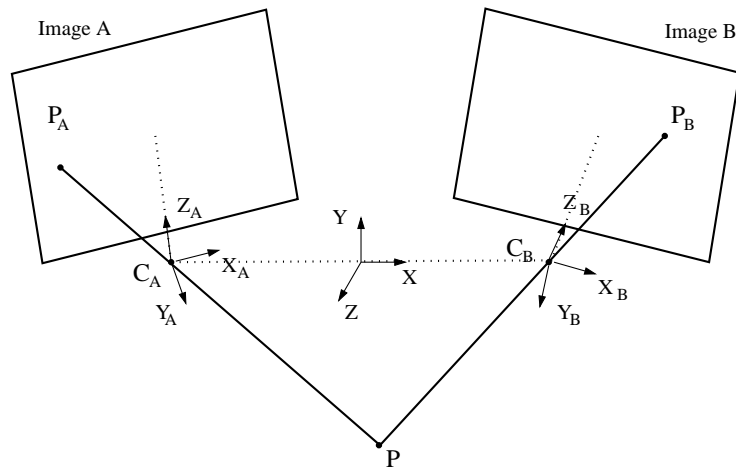


FIG. 4 – Systèmes de coordonnées pour l'orientation relative

2.3 Coordonnées de l'espace objet

Lorsque les paramètres de prise de vue sont connus, il est possible de déterminer la position du point P dans la scène si nous connaissons ses projections P_A et P_B dans les deux images. P est l'intersection des droites (P_A, C_A) et (P_B, C_B) .

Comme la connaissance des paramètres de prise de vue peut être perturbée par du bruit, le point P est alors défini comme le centre du plus petit segment joignant les droites (P_A, C_A) et (P_B, C_B) : P est alors la solution au sens des moindres carrés des 4 équations fournies par les 2 droites (P_A, C_A) et (P_B, C_B) .

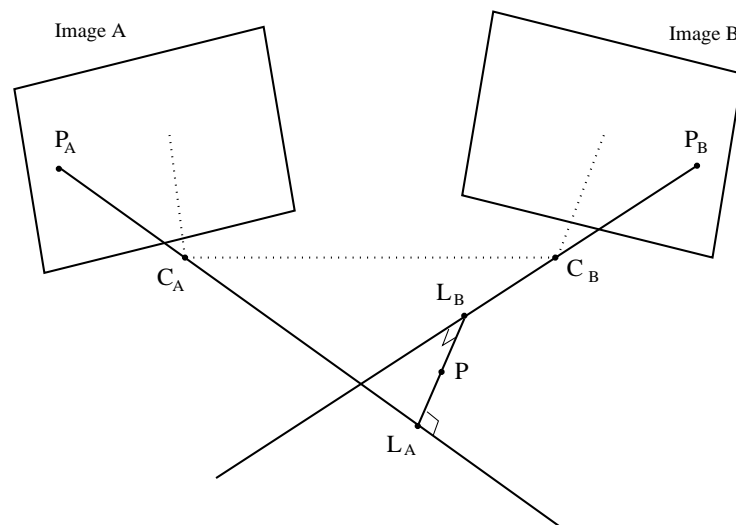


FIG. 5 – Pour deux points conjugués P_A et P_B , l'orientation relative permet de calculer la position du point P dans le système de coordonnées lié au couple d'images. P est défini comme le centre du plus petit segment reliant les droites (P_A, C_A) et (P_B, C_B) .

2.4 Rectification des images

Le but de la rectification des images est reconstruire les images dans un cadre parallèle à la base (C_A, C_B) de la paire stéréo. Cette opération est possible si les paramètres de l'orientation relative sont connus. Si, de plus, les paramètres de l'orientation absolue d'au moins une image sont connus, il est possible de sélectionner l'orientation du plan des nouvelles images par rapport à l'espace objet (par exemple, de choisir une prise de vue verticale). Il est important de noter que dans cette opération, les deux centres optiques gardent la même position

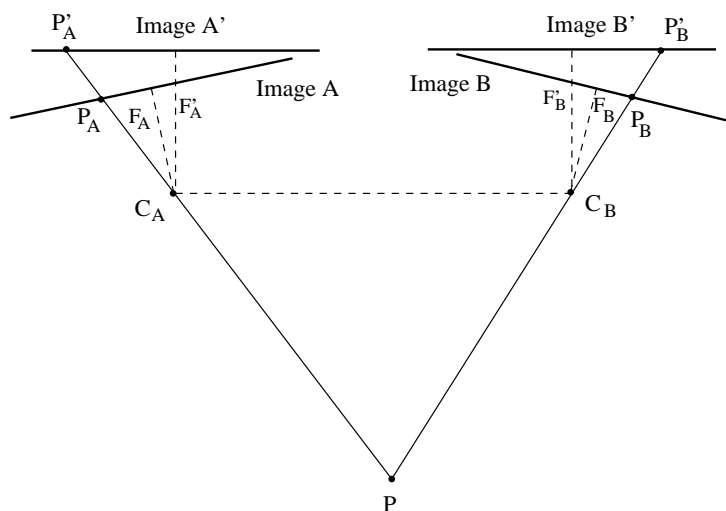


FIG. 6 – Lors de la rectification en géométrie épipolaire, les images sont reconstruites dans deux espaces coplanaires

3 Génération de données 3D

Deux types de données peuvent être obtenues à partir d'images aériennes stéréoscopiques :

Modèle numérique d'élévation (MNE) : l'élévation du terrain est calculée sur une grille régulière géoréférencée. Afin de réduire le volume des données et de rendre leur manipulation plus aisée, il est possible de calculer une triangulation (TIN : *Triangular Irregular Network*) d'un MNE.

Description des bâtiments : chaque bâtiment est décrit par un ensemble de points 3D localisés au niveau des coins du toit. Pour des toits complexes de bâtiments, la description doit préserver la topologie des différentes facettes du toit.

3.1 Approche surface

Dans cette approche, la disparité est calculée pour tous les points (si possible) d'une image : pour chaque point P de l'image A , on recherche le point Q de l'image B qui est le plus *semblable* au point P .

Avec des images en géométrie épipolaire, ce problème de recherche 2D devient un problème 1D puisque deux points conjugués sont nécessairement situés sur la même ligne dans les images A et B .

Une carte de disparité est créée dans le même espace de coordonnées que l'image A : pour chaque point, la disparité est la différence d'abscisse entre les points P et Q .

Parties cachées : comme un point de l'image A peut ne pas être vu dans l'image B , il n'est pas possible en général de calculer une carte de disparité complètement dense. Des systèmes multi-vues peuvent être envisagés pour résoudre ce problème.

3.1.1 Corrélacion croisée

Les techniques fondées sur la corrélation des intensités des images ont été utilisées dans de nombreuses applications commerciales de la stéréo-photogrammétrie, mais elles sont également l'une des plus anciennes méthodes développées en vision par ordination.

Le principe de la corrélation croisée pour mesurer la similarité entre deux points est de comparer l'intensité dans des fenêtres rectangulaires autour de chaque point.

Différentes mesures de similarité peuvent être envisagées :

- Différence quadratique
- Corrélation croisée

Le coefficient de corrélation centré et normalisé s'écrit de la façon suivante :

$$\mathcal{C}(P, Q) = \frac{\sum_{k=-K}^{k=K} \sum_{l=-L}^{l=L} (A(x_P+k, y_P+l) - \overline{A(x_P, y_P)}) \times (B(x_Q+k, y_Q+l) - \overline{B(x_Q, y_Q)})}{(2K+1) \cdot (2L+1) \sigma_A(x_P, y_P) \sigma_B(x_Q, y_Q)}$$

où $\overline{A(x_P, y_P)}$ et $\sigma_A(x_P, y_P)$ sont respectivement l'intensité moyenne et

l'écart-type de l'image A au point P (même notation pour l'image B) :

$$\overline{A(x_P, y_P)} = \frac{\sum_{k=-K}^{k=K} \sum_{l=-L}^{l=L} A(x_P+k, y_P+l)}{(2K+1).(2L+1)}$$

$$\sigma_A(x_P, y_P) = \frac{\sqrt{\sum_{k=-K}^{k=K} \sum_{l=-L}^{l=L} (A(x_P+k, y_P+l) - \overline{A(x_P, y_P)})^2}}{(2K+1).(2L+1)}$$

Avec cette définition du coefficient de corrélation croisée, nous avons les propriétés suivantes :

- $-1 \leq \mathcal{C}(P, Q) \leq +1$
- $\mathcal{C}(P, Q) = \pm 1 \iff \exists a, b \quad \forall k, l \quad B(x_Q+k, y_Q+l) = a \times A(x_P+k, y_P+l) + b$
- Information mutuelle

Pour chaque point P de l'image A , nous choisissons sur la même ligne épipolaire le point Q de l'image B qui maximise le coefficient de corrélation croisée $\mathcal{C}(P, Q)$.

3.1.2 Contraintes géométriques

Plusieurs contraintes géométriques peuvent être utilisées pour réduire l'espace de recherche de points conjugués.

- unicité : pour des objets opaques, un point de l'image A n'a au plus qu'un correspondant dans l'image B , et vice-versa,
- ordre : toujours pour des objets opaques, sur une ligne épipolaire, si le point P' est à la droite du point P , alors le correspondant de P' est à la droite du correspondant du point P .
- continuité : pour des paysages relativement peu accidentés (zones rurales), la carte de disparité ne doit pas comprendre de forte variation de disparité, en zone urbaine le gradient de disparité correspondent généralement à des parties cachées.

3.2 Approche objets

Dans cette approche, nous commençons par détecter des primitives dans les deux images que nous essayons ensuite de mettre en correspondance. La géométrie épipolaire est également utilisée pour contraindre la recherche des appariements.

Différents types de primitives peuvent être envisagées : points, segments régions ou primitives de plus haut niveau comme les coins ou les carrefours. Le choix des primitives va dépendre de l'application considérée.

Les mêmes contraintes peuvent être utilisées que pour l'approche surface (géométrie épipolaire, unicité, ordre, continuité, gradient de disparité), mais doivent être adaptées à chaque type de primitives.