

# Géométrie projective, analyse numérique et vision par ordinateur

Roger Mohr<sup>1</sup>, Matthijs Douze<sup>2</sup>, Peter Sturm<sup>2</sup>

<sup>1</sup> ENSIMAG – INPG, Campus Universitaire, 38402 St Martin d’Hères Cedex

<sup>2</sup> INRIA Rhône-Alpes, 655 Avenue de l’Europe, 38330 Montbonnot St Martin

Roger.Mohr@imag.fr, Matthijs.Douze@inrialpes.fr, Peter.Sturm@inrialpes.fr

## 1 Introduction et motivation

### 1.1 Objectifs

Les objectifs de cet article sont multiples. Ils pourraient être simplement scientifiques, en illustrant comment les mathématiques du 19<sup>e</sup> siècle, couplées avec les outils algébriques et numériques récents, ont permis de revisiter le traitement des images multi-caméra. À partir de celles-ci, un ordinateur peut percevoir en trois dimensions la scène observée, c’est-à-dire la « reconstruire » virtuellement (cf. la figure 1).

Pour les professeurs de classes préparatoires scientifiques, c’est l’illustration d’une démarche d’ingénieur : modélisation à travers un outil mathématique, puis résolution effective en tenant compte des erreurs du monde réel.

Les mathématiciens auront le plaisir de voir un problème concret mis en forme élégamment, avec des mathématiques un peu oubliées : la géométrie projective. Ils apprécieront que la démarche d’ingénieur va bien au delà de la mise en équations. En particulier nous explorerons les méthodes numériques nécessaires ; elles sont mieux connues des physiciens que des mathématiciens.

Le dernier objectif de cet article est de fournir pour l’avenir un sujet de TIPE motivant qui pourrait illustrer la « dualité fondamental – appliqué ». Il a été expérimenté avec des jeunes, et nécessite la mise en œuvre d’instruments informatiques : bibliothèques ou logiciels de calcul, et outil de mise en correspondance dans des images (contribution d’un des laboratoires de l’ENSIMAG).

Le problème concret que nous essayons de résoudre dans cet article est la restitution en trois dimensions de la forme d’un objet, à partir de deux ou plusieurs photographies, dont les caractéristiques de prise de vue sont *inconnues* (réglage du zoom ou de la mise au point, position des points de vue, orientation de l’appareil photo lors de la prise de vue). La figure 1 illustre cette problématique.



FIG. 1 – Deux images d’une scène et une reconstruction 3D obtenue à partir de six images au total. Les pyramides représentent les positions reconstruites et le champ de vue des caméras utilisées.

## 1.2 Plan de l'article

La partie 2 aborde des notions de géométrie projective, présentées essentiellement en notations algébriques. Les lecteurs intéressés par cette formulation pourront se reporter à [5], un ouvrage moderne et très complet sur ce sujet. Nous nous limiterons aux outils nécessaires à l'application présentée par la suite.

La partie 3 traite de la contrainte liant deux vues perspectives d'une même scène. Elle se résume ainsi : le monde tridimensionnel se projette dans deux images bidimensionnelles ; c'est une application d'un espace de dimension 3 dans un espace de dimension  $2 \times 2 = 4$ , les points projetés sont donc liés par une contrainte de dimension 1 que nous expliciterons et dont nous indiquerons l'importance pour les applications. L'ouvrage de référence sur ce sujet est [3]. Nous renvoyons le lecteur francophone au site de Peter Sturm<sup>1</sup>, où il trouvera sa thèse [6] qui discute et prolonge le sujet.

La résolution numérique de l'équation posée est délicate. La section 4 évoquera des méthodes numériques où les outils de l'algèbre linéaire sont mis à contribution.

Nous terminerons cette lecture introductive par des indications qui permettront aux professeurs intéressés de motiver leurs élèves pour un futur TIPE lorsque la thématique le permettra.

## 1.3 Quelques aspects historiques

Les mathématiciens de la Grèce antique avaient des connaissances fines sur les proportions en géométrie. Ils avaient généralisé le théorème de Thalès aux faisceaux de droites (ensembles de droites concourantes), menant à la notion de *birapport*, l'invariant de base de la géométrie projective. Sur la figure 2, le birapport  $(A, B; C, D)$  est égal à  $(A', B'; C', D')$  :

$$(A, B; C, D) = \frac{\frac{\overline{AC}}{AD}}{\frac{\overline{BC}}{BD}} = \frac{\frac{\overline{A'C'}}{A'D'}}{\frac{\overline{B'C'}}{B'D'}} = (A', B'; C', D')$$

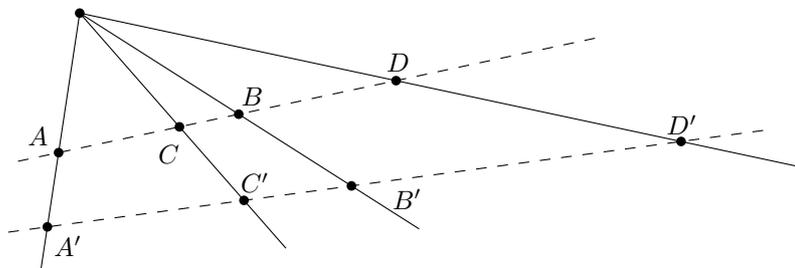


FIG. 2 – Intersection de 4 droites en faisceau avec deux droites non parallèles, définissant deux birapports égaux.

Le père de la géométrie projective est le mathématicien Girard Desargues (1591-1661), mais sa contribution était mal comprise par ses contemporains. La redécouverte d'un de ses manuscrits deux siècles plus tard permit de rendre hommage à ses découvertes.

Les mathématiciens de la révolution, dont Gaspard Monge, ont développé cette géométrie pour des besoins militaires (acquérir les dimensions des places fortes au moyen de plusieurs vues en perspective). Ils ont mis en place les méthodes analytiques permettant d'obtenir des résultats numériques.

La photographie a automatisé la projection perspective manuelle utilisée par Monge, constituant une nouvelle discipline (la *photogrammétrie*) pour laquelle existent encore des départements dans quelques universités.

<sup>1</sup><http://perception.inrialpes.fr/people/Sturm/>

Son récent couplage avec l'analyse d'image par ordinateur permet d'automatiser complètement la perception volumétrique automatique d'une scène observée par plusieurs caméras. Ce domaine s'est fortement développé dans la dernière décennie du XX<sup>e</sup> siècle et ses raffinements sont encore à l'étude dans les laboratoires de recherche.

Cet article traite des principes géométriques et numériques de cette technique.

## 2 Introduction à la géométrie projective analytique

Nous nous limiterons le plus souvent dans cette partie à la géométrie du plan. Cependant, les notions introduites se généralisent facilement à des espaces de dimensions supérieures, et quelques allusions feront référence à cette généralisation.

Pour les calculs matriciels, nous identifierons les coordonnées (entre parenthèses) et les vecteurs colonnes (entre crochets). Ainsi,  $(a, b) = \begin{bmatrix} a \\ b \end{bmatrix}$ .

### 2.1 Un peu d'intuition

L'étude de la projection perspective a été initiée par les peintres de la renaissance italienne qui souhaitaient rendre réaliste le relief dans leurs œuvres. Ils utilisèrent nombre de techniques, en particulier le *point de fuite* – la projection du point à l'infini (point où deux droites parallèles se coupent) – qui leur permettait par exemple de représenter en perspective un carré divisé en quatre sous-carrés égaux.

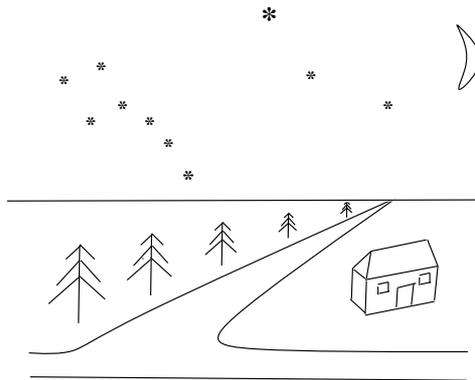


FIG. 3 – Paysage avec horizon

La figure 3 illustre un aspect important d'une projection perspective : les bords de route, parallèles dans le monde tridimensionnel, se rejoignent sur la ligne d'horizon, lieu des points à l'infini du plan horizontal. Cette illustration montre qu'on doit étendre les concepts familiers des géométries euclidienne ou affine pour raisonner et calculer sur les projections perspectives et les images qui en sont issues. Cette section introduit le formalisme algébrique dans lequel sera exprimé, à la section suivante, la projection perspective.

### 2.2 Coordonnées homogènes

Dans le plan affine un point  $P$  est représenté par ses coordonnées  $(X, Y)$ . Le principe des coordonnées homogènes est de représenter ce point par le triplet  $(x, y, t)$ , avec  $t \neq 0$  et tel que  $x/t = X$ ,  $y/t = Y$ . Pour tout scalaire  $\lambda \neq 0$ ,  $\lambda(x, y, t) = (\lambda x, \lambda y, \lambda t)$  représente le même point que  $(x, y, t)$  :

les coordonnées homogènes sont donc définies à un facteur multiplicatif près. En particulier, en posant  $t = 1$ , le point a les coordonnées  $(X, Y, 1)$ .

Formellement,  $\sim$  est la relation d'équivalence suivante :

$$(x, y, t) \sim (x', y', t') \Leftrightarrow \exists \lambda \neq 0 \mid (x, y, t) = \lambda(x', y', t')$$

Le *plan projectif*  $\mathbb{P}^2$  est le quotient de  $\mathbb{R}^3 - (0, 0, 0)$  par la relation  $\sim$ . Le plongement canonique de l'espace euclidien dans l'espace projectif sera développé au § 2.3.

Regardons maintenant comment représenter des points dits à l'infini, en utilisant des coordonnées homogènes. Soit  $P_\alpha = (\alpha X, \alpha Y)$ . En faisant varier  $\alpha$ ,  $P_\alpha$  se déplace le long d'une droite contenant l'origine. Comment raisonner sur la limite de  $\alpha \rightarrow \infty$ ? En coordonnées homogènes,  $P_\alpha$  peut s'écrire  $(\alpha x, \alpha y, t) \sim (x, y, t/\alpha)$ . En passant à la limite on obtient un *point à l'infini*

$$P_\infty \sim (x, y, 0) \sim (X, Y, 0)$$

Ce point à l'infini correspond à la *direction* de la droite engendrée par l'origine et les  $P_\alpha$ . Par ailleurs, des droites qui sont parallèles ont la même direction et sont donc concourantes – en un point à l'infini – concept qui n'existe pas en géométrie affine ou euclidienne.

Remarquons au passage que tout point dont la dernière coordonnée homogène vaut zéro est un point à l'infini.

L'appellation « coordonnées homogènes » vient de ce que toutes les expressions algébriques deviennent homogènes en degré. Ainsi l'équation d'une droite  $aX + bY + c = 0$  devient  $ax + by + ct = 0$  (les trois termes de la somme sont de même degré). On appelle *droite du plan projectif* tout ensemble de points de  $\mathbb{P}^2$  défini par une équation linéaire non nulle, c'est-à-dire avec  $(a, b, c) \neq (0, 0, 0)$ . L'équation  $t = 0$  définit la droite contenant tous les points à l'infini.

Cette propriété d'homogénéité est vérifiée quelque soit le degré. L'hyperbole  $XY = 1$  devient  $xy = t^2$ . Les points  $(1, 0, 0)$  et  $(0, 1, 0)$  y appartiennent : ces points à l'infini correspondent aux directions des axes. L'hyperbole est maintenant connexe : en raccordant les branches à l'infini on en fait une courbe fermée d'un seul tenant (mais sans notion d'intérieur ou d'extérieur).

**Exercice 2.1** : Traduire l'équation de la parabole  $Y = X^2$  en coordonnées homogènes. Vérifier que la droite à l'infini  $t = 0$  est tangente à cette parabole (algébriquement, la tangence se traduit par une solution double pour l'intersection).

**Exercice 2.2** : Montrer que trois points sont alignés si et seulement si leurs coordonnées homogènes sont dépendantes linéairement.

## 2.3 Espace projectif

Nous avons introduit les points du plan projectif  $\mathbb{P}^2$ . La généralisation à  $n$  dimensions est directe : les points de l'espace projectif  $\mathbb{P}^n$  sont les classes d'équivalence de  $\mathbb{R}^{n+1}$  par la relation de proportionnalité  $\sim$ . Complétons la géométrie avec la description des transformations compatibles sur ces espaces.

Les *transformations projectives* sont les applications bijectives correspondantes aux transformations linéaires des coordonnées des points. Ces coordonnées étant définies à un facteur de proportionnalité près, il en est de même pour les matrices de ces transformations. On appelle *homographie* toute transformation projective de l'espace projectif  $\mathbb{P}^n$  sur lui-même.

L'espace affine est plongé dans l'espace projectif, il lui manque simplement les points à l'infini. Nous montrons ici, dans un espace de dimension 2, que les transformations affines se plongent aussi dans l'espace projectif.

Une translation par un vecteur  $(a, b)$  transforme  $(X, Y)$  en  $(X + a, Y + b)$ , dont les plongements sont respectivement  $P(X, Y, 1)$  et  $P'(X + a, Y + b, 1)$ , d'où :

$$P' = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} P$$

La translation devient donc une homographie qui laisse les points à l'infini invariants. Ceci correspond bien à notre intuition : quand nous voyageons en translation, les étoiles (points à l'infini) restent toujours à la même position dans notre référentiel.

Plus généralement, une transformation affine  $(X, Y) \mapsto (aX + bY + c, dX + eY + f)$  correspond à

$$P' = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} P$$

Le lecteur vérifiera aisément que les points à l'infini y restent, donc le plongement de l'espace affine dans l'espace projectif est cohérent.

Nous ne développons pas les propriétés d'invariance de l'espace projectif. Nous avons déjà mentionné (§ 1.3) que le birapport était l'invariant fondamental de cette géométrie.

## 2.4 Espace projectif dual

Soit la droite projective d'équation  $ax + by + ct = 0$ ,  $(x, y, t) \in \mathbb{P}^2$ . Ses coefficients  $(a, b, c)$  sont définis à un facteur multiplicatif près. On peut disposer l'équation sous la forme du produit scalaire

$$(a, b, c) \cdot (x, y, t) = 0$$

Notons la symétrie de l'équation : elle peut être vue soit comme l'équation du faisceau de droites passant par un point, soit comme l'équation des points appartenant à une droite. Cette symétrie mène à la notion d'*espace dual* des droites (et plus généralement des hyperplans). À tout espace projectif  $\mathbb{P}^n$  correspond l'espace dual  $\mathbb{P}^{*n}$  de ses hyperplans. Ainsi,  $\mathbb{P}^{*2}$  est l'espace des droites de  $\mathbb{P}^2$ , et  $\mathbb{P}^{*3}$  celui des plans de  $\mathbb{P}^3$ .

Le *principe de dualité* s'énonce, en deux dimensions :  
*pour toute propriété de géométrie projective établie entre des points, il y a un résultat symétrique dans lequel les rôles des droites et des points sont échangés.*

Les points deviennent des droites, les droites des points, la droite passant par un point devient le point sur une droite, etc.

Ainsi la propriété *deux points distincts définissent une droite unique* (portant les deux points), devient *deux droites distinctes définissent un point unique* (d'intersection). Ce résultat n'est valide que dans l'espace projectif : deux droites parallèles dans l'espace affine ne se coupent pas, alors que dans l'espace projectif elles se coupent en un point à l'infini comme nous avons vu plus haut.

**Le théorème de Desargues.** Le théorème suivant est une jolie illustration du principe de dualité et d'une démonstration basée sur le calcul algébrique. Il établit que si les sommets de deux triangles sont en correspondance perspective, alors leurs côtés sont aussi en correspondance perspective.

**Théorème :** *Soient  $ABC$  et  $A'B'C'$  deux triangles du plan projectif (figure 4). Les droites  $(AA')$ ,  $(BB')$  et  $(CC')$  sont concourantes si et seulement si les points d'intersection des côtés correspondants  $(AB)$  et  $(A'B')$ ,  $(BC)$  et  $(B'C')$ ,  $(CA)$  et  $(C'A')$  sont alignés.*

Voici la démonstration algébrique d'une direction du théorème – la démonstration de la réciproque est « gratuite » grâce au concept de dualité, voir plus bas. Supposons alors que les droites  $(AA')$ ,

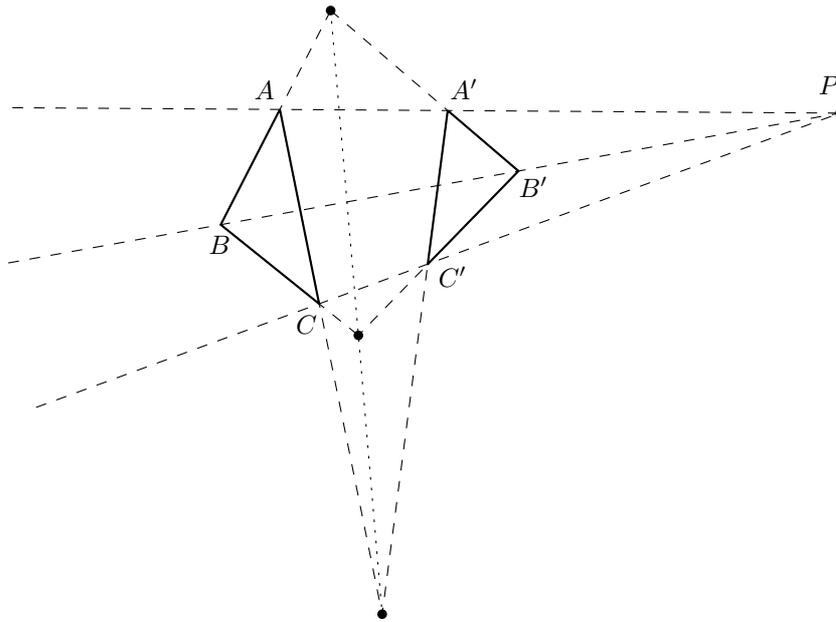


FIG. 4 – Deux triangles vérifiant les hypothèses du théorème de Desargues

$(BB')$  et  $(CC')$  sont concourantes. Soit  $P$  est l'intersection de  $(AA')$ ,  $(BB')$  et  $(CC')$ . Il existe (selon le résultat de l'exercice 2.2) 6 valeurs scalaires  $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$  telles que :

$$\left. \begin{array}{l} \alpha A - \alpha' A' = P \\ \beta B - \beta' B' = P \\ \gamma C - \gamma' C' = P \end{array} \right\} \implies \left\{ \begin{array}{l} \alpha A - \beta B = \alpha' A' - \beta' B' \\ \beta B - \gamma C = \beta' B' - \gamma' C' \\ \gamma C - \alpha A = \gamma' C' - \alpha' A' \end{array} \right.$$

La première égalité montre que  $\alpha A - \beta B$ , sur  $(AB)$ , coïncide avec  $\alpha' A' - \beta' B'$ , sur  $(A'B')$ . C'est donc l'intersection des deux droites. Il en va de même pour  $\beta B - \gamma C$  et  $\gamma C - \alpha A$ . Or,

$$(\alpha A - \beta B) + (\beta B - \gamma C) + (\gamma C - \alpha A) = 0$$

Donc les trois points sont alignés. Ceci démontre la première implication du théorème. La démonstration de la réciproque est le sujet de l'exercice suivant.

**Exercice 2.3 :** *Démontrer la réciproque du théorème de Desargues. Conseil : utiliser la démonstration ci-dessus, en échangeant les points avec les droites et vice-versa, ainsi que la sémantique des notions « alignés » et « concourantes » ainsi que « droite engendrée par deux points » et « point d'intersection de deux droites ».*

## 2.5 Règles de calcul en géométrie projective

Nous aurons besoin de quelques règles de calcul dans le plan projectif pour établir le résultat de la partie 3.

Le produit scalaire «  $\cdot$  » est utilisé entre un point de l'espace projectif  $\mathbb{P}^2$  et une droite (c'est-à-dire un point de  $\mathbb{P}^{*2}$ ). Nous noterons «  $\otimes$  » le produit vectoriel. Rappelons que pour deux vecteurs  $\vec{V}, \vec{U}$  de l'espace euclidien de dimension 3,  $\vec{V} \otimes \vec{U}$  est un vecteur orthogonal au plan défini par  $\vec{V}$  et  $\vec{U}$ . En notant  $\vec{U} = (u_1, u_2, u_3)$  et  $\vec{V} = (v_1, v_2, v_3)$ ,

$$\vec{U} \otimes \vec{V} = \begin{bmatrix} v_2 u_3 - v_3 u_2 \\ v_3 u_1 - v_1 u_3 \\ v_1 u_2 - v_2 u_1 \end{bmatrix}$$

L'application  $\vec{U} \mapsto \vec{U} \otimes \vec{V}$ , linéaire, est définie par la matrice antisymétrique de rang 2 :

$$\begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}$$

En identifiant  $\mathbb{R}^3$  à  $\mathbb{P}^2$ ,  $\otimes$  peut être utilisé entre deux points  $M, N \in \mathbb{P}^2 : l = M \otimes N \in \mathbb{P}^{*2}$  est tel que  $l \cdot M = 0$  et  $l \cdot N = 0$ ;  $l$  est donc la droite passant par ces deux points.

La formulation duale de cette écriture est : soient  $l$  et  $l'$  deux points de  $\mathbb{P}^{*2}$  (donc deux droites de  $\mathbb{P}^2$ ), et considérons  $M = l \otimes l'$  comme un point de  $\mathbb{P}^2$ ;  $l \cdot M = l' \cdot M = 0$ , donc  $M$  est le point d'intersection des deux droites.

**Exercice 2.4** Soit une homographie dans  $\mathbb{P}^n$  de matrice  $A$ . Montrer qu'elle se traduit par une homographie de matrice  $(A^{-1})^\top$  dans  $\mathbb{P}^{*n}$ .

**Exercice 2.5** Soient  $A, B, C, D$  les sommets consécutifs d'un parallélogramme du plan affine qui se transforment en  $A', B', C', D'$  par une homographie. Montrer que l'image par cette homographie de la droite à l'infini du plan est

$$((A' \otimes B') \otimes (C' \otimes D')) \otimes ((A' \otimes D') \otimes (C' \otimes B')) \in \mathbb{P}^{*2}$$

*Indications* : Le point  $(A \otimes B) \otimes (C \otimes D)$  est l'intersection à l'infini des parallèles  $(AB)$  et  $(CD)$ , donc le point à l'infini dans la direction  $\vec{AB}$ . La droite à l'infini est définie par deux points à l'infini.

Dans un dessin en perspective, cette règle permet de trouver l'horizon si on a déjà dessiné un parallélogramme horizontal.

## 2.6 Base d'un espace projectif

Pour faire de la géométrie analytique dans un espace affine, on se lie à un repère arbitraire. En dimension  $n$ , le repère est défini par  $n + 1$  points  $(P_0, \dots, P_n)$  en configuration générale (ne se trouvant pas dans un hyperplan ou bien un sous-espace de dimension  $n - 1$ ) :  $P_0$  est l'origine et  $P_i$  est l'extrémité du  $i^{\text{ième}}$  vecteur directeur. Pour deux repères  $\mathcal{R}_1$  et  $\mathcal{R}_2$ , il existe une unique transformation affine telle que l'image de  $\mathcal{R}_1$  soit  $\mathcal{R}_2$ .

Cette propriété de l'espace affine se transpose dans l'espace projectif. Nous avons jusqu'à présent fixé la base projective à partir de celle de l'espace vectoriel affine plongé.

Dans un espace projectif  $\mathbb{P}^n$ , on appelle *base* toute séquence de  $n + 2$  points  $(P_0, P_1, \dots, P_{n+1})$  linéairement indépendants dans  $\mathbb{P}^n$  (dans le plan, ce sont 4 points dont aucun groupe de 3 n'est aligné). On peut montrer que l'image de ces  $n + 2$  points détermine une (seule) homographie de  $\mathbb{P}^n$ . La démonstration est technique, mais un raisonnement sur les degrés de liberté permet de toucher du doigt le résultat : l'homographie est définie par une matrice  $(n + 1) \times (n + 1)$  à un facteur multiplicatif près, soit  $(n + 1)^2 - 1 = n^2 + 2n$  degrés de liberté ; l'image de chaque point fournit  $n + 1$  équations à un facteur près, donc les  $n + 2$  points fournissent  $n^2 + 2n$  contraintes.

Le parallélogramme et son image de l'exercice 2.5 étaient des bases, donc ils définissaient complètement une homographie.

**Exercice 2.6** Soit un parallélogramme dont on connaît l'image par une homographie. Construire l'image des milieux de ses côtés en utilisant seulement une règle (la technique était déjà connue de Léonard de Vinci).

Une fois la base choisie, on y associe habituellement les coordonnées homogènes : l'origine  $P_0 = (0, \dots, 0, 1)$ , le point à l'infini sur la  $i^{\text{ième}}$  direction  $P_i = (0, \dots, 1, 0, \dots, 0)$ , le *point unitaire*  $P_{n+1} = (1, 1, \dots, 1)$ .

### 3 Stéréovision

Nous explicitons ici la contrainte liant deux vues d'une même scène. Celle-ci sera précieuse pour la phase d'estimation de la section suivante.

Nous adopterons la notation 2D pour « bidimensionnel », 3D pour « tridimensionnel ».

#### 3.1 Projection perspective

La projection perspective est la projection, réalisée par des rayons lumineux, d'une image du monde observé à travers un orifice infiniment petit (le *centre optique*) sur un plan (le *plan image*). Géométriquement (figure 5), c'est une projection centrale par rapport au centre optique  $O$  d'un point 3D  $M$  sur un plan  $I$ .

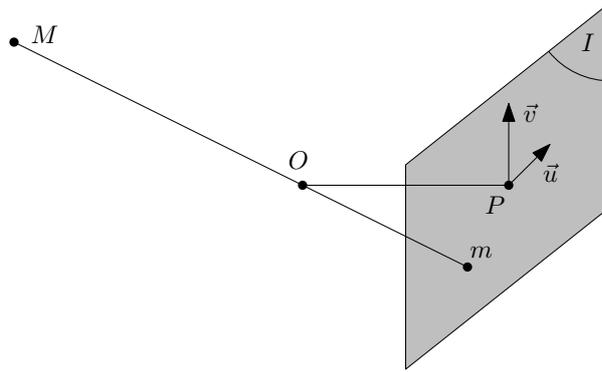


FIG. 5 – Projection perspective d'un point 3D.

Notons  $P$  la projection orthogonale de  $O$  sur  $I$ . Soit un repère affine orthonormal 3D  $(O, \vec{u}, \vec{v}, \vec{OP})$ , notons  $(X, Y, Z)$  les coordonnées de  $M$ . Dans  $(P, \vec{u}, \vec{v})$  appelons  $(x, y)$  les coordonnées de  $m$  sur le plan image. Alors,  $(x, y) = (X/Z, Y/Z)$ .

En plongeant les deux repères dans des espaces projectifs, les coordonnées de  $M$  et  $m$  deviennent  $(x, y, 1)$  et  $(X, Y, Z, 1)$ . L'égalité devient :

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

L'application  $M \mapsto m$  est une transformation linéaire des coordonnées homogènes. La projection centrale est donc une transformation projective, représentée par une *matrice de projection* de dimensions  $3 \times 4$ . Cette propriété est indépendante des systèmes de coordonnées choisis en 2D et 3D.

Pour faire de la stéréovision analytique, les matrices de projection associées aux différentes vues doivent bien sûr être exprimées par rapport au même système de coordonnées en 3D.

**Exercice 3.1 :** *Considérons le cas particulier suivant : deux projections perspectives d'un ensemble de points 3D qui sont coplanaires. Montrer qu'il existe une homographie de  $\mathbb{P}^2$  entre les deux ensembles de projetés des points 3D.*

Un appareil photo, même bien focalisé, ne réalise pas exactement une projection centrale : les optiques produisent des distorsions qui courbent les droites sur les images (l'effet est le plus fort sur des optiques de mauvaise qualité à focale courte). Des corrections peuvent être apportées aux

images pour annuler les distorsions. Dans la suite de cet exposé, nous supposons les projections perspectives parfaites.

Avant de passer à la suite, contemplons une construction du 16<sup>e</sup> siècle, développée par Dürer afin de réaliser des dessins perspectifs (cf. la figure 6). Le point d'attache d'un fil sur un mur joue le rôle du centre optique. Le fil, si tendu entre ce point d'attache et un point 3D de l'objet à dessiner, peut être assimilé à un rayon lumineux. Quant au plan image, il est représenté par un cadre permettant de mesurer les coordonnées 2D du point de passage du fil. Il suffit ensuite de dessiner des points avec ces coordonnées sur une planche afin de produire un dessin perspectif !

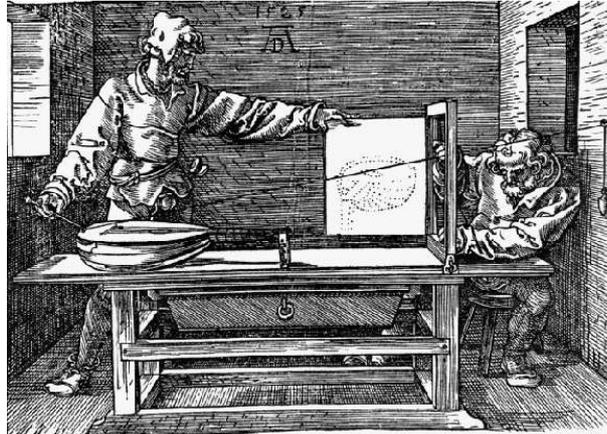


FIG. 6 – Étude de la projection perspective (A. Dürer).

### 3.2 Reconstruction 3D avec un système de stéréovision étalonné

Considérons le problème énoncé dans l'introduction de cet article et illustré sur la figure 1 : à partir de deux projections perspectives d'une scène, comment restituer la scène en trois dimensions ? Supposons pour le moment que les deux images aient été prises par un système stéréovision calibré, c'est-à-dire deux caméras dont nous connaissons les matrices de projection  $P$  et  $P'$  par rapport au même système de coordonnées 3D.

Nous formulons la reconstruction 3D de la scène par le calcul des coordonnées 3D des points particuliers ou caractéristiques de la scène, à partir de leurs projetés dans les deux images. Avant de passer au calcul proprement dit, contemplons deux questions fondamentales de la vision *par ordinateur*, c'est-à-dire d'un traitement automatique des images :

1. comment définir des « points particuliers » ou bien, comment les trouver automatiquement dans les images ? En regardant les images de la figure 1, on aura envie de sélectionner des points correspondants aux extrémités de la maison ou de ses fenêtres par exemple. Or, ce raisonnement implique que nous avons déjà reconnu l'objet sur les images comme une maison. Notre système visuel humain est très performant en reconnaissance d'objets, grâce à une évolution et un apprentissage au travers de millions d'années. La situation est différente en vision par ordinateur – une image digitale est tout d'abord un grand ensemble de chiffres (les niveaux de gris des pixels) et reconnaître des objets à partir de ces données est tout sauf trivial. Il s'agit là d'un domaine de recherche à part entière, en pleine floraison.

En l'absence d'une telle reconnaissance d'objets, nous pouvons néanmoins donner des définitions plausibles de points particuliers. Celles habituellement adoptées considèrent chaque pixel d'une image ; si la région autour du pixel exhibe un fort contraste, dans au moins deux directions, le pixel est déclaré comme étant un point particulier. Ce type d'approche s'implémente aisément sur ordinateur.

2. comment identifier des points homologues, c'est-à-dire des points particuliers qui sont les projetés du même point de la scène ? Ce problème est aussi appelé la mise en correspondance. Des centaines d'approches existent dans la littérature. Elles ont pour principe de base commun une comparaison des régions autour de points particuliers, c'est-à-dire des niveaux de gris des pixels dans un voisinage des points particuliers. Des régions similaires indiquent évidemment qu'il peut s'agir de points homologues.

Cette définition est intuitive et apparemment simple. Néanmoins, elle entraîne a priori une exploration combinatoire de toutes les paires de points particuliers dans les deux images, processus coûteux en temps de calcul. De plus, il y a toujours un certain risque de faire des erreurs – dans les images sur la figure 1 par exemple, plusieurs coins de fenêtres se ressemblent visuellement. Comment trouver les points homologues correctes ? Une réponse partielle est donnée au §3.3. Même des points homologues correctes sont sujets à des imprécisions que l'on dénomme habituellement par « bruit » : les niveaux de gris sont sujet à des bruits électroniques inévitables et les images digitales sont discrètes, ne permettant pas de retrouver des positions de points homologues mathématiquement exactes.

Supposons que les problèmes ci-dessus ont été résolus et que nous disposons de paires de points homologues. Soient  $m$  et  $m'$  une telle paire. Nous pouvons maintenant calculer les coordonnées du point 3D original  $M$ , en s'appuyant sur l'égalité (1), adaptée à nos choix :  $m \sim PM$  et  $m' \sim P'M$ . On peut en déduire que  $m \otimes (PM) = 0$  et  $m' \otimes (P'M) = 0$ , où 0 représente bien entendu un vecteur de longueur 3, ne contenant que des zéros. Nous disposons donc de 6 équations linéaires et homogènes en les coordonnées de  $M$ . Il faut souligner que, à cause du bruit mentionné ci-dessus, ce système d'équations n'admet en général aucune solution exacte. Il faut donc trouver une solution représentant le meilleur compromis ; la méthode de choix est la résolution « aux moindres carrés », détaillée au §4.2.

La partie droite de la figure 1 montre le résultat d'une reconstruction 3D, où en plus des points 3D, des parties planes de la scène ont été déterminées automatiquement (non détaillé ici).

### 3.3 La géométrie épipolaire

La mise en correspondance, telle qu'expliquée au paragraphe précédent, requiert a priori une recherche combinatoire de points homologues. Il se trouve que la complexité de cette recherche peut être réduite de manière drastique, en exploitant une relation géométrique fondamentale, révélée par la suite, que doit admettre toute paire de points homologues correcte.

Sur la figure 7,  $M$  se projette par rapport à  $O$  (respectivement  $O'$ ) sur le plan image  $I$  (resp.  $I'$ ) en  $m$  (resp.  $m'$ ). Le plan épipolaire ( $OO'M$ ) coupe les deux plans images en deux droites épipolaires conjuguées. Cette situation est aussi illustrée pour des images réelles sur la figure 8.

Lorsque  $M$  se déplace, les plans épipolaires ( $O, O', M$ ) forment un faisceau (ensemble de plans contenant une même droite) s'appuyant sur la droite ( $OO'$ ). L'intersection du faisceau avec  $I$  est un faisceau de droites épipolaires se coupant au point  $e$ , l'épipole, intersection de ( $OO'$ ) avec  $I$ . Appelons  $e'$  l'épipole sur  $I'$ .

Deux droites épipolaires conjuguées sont en projection perspective l'une de l'autre, il existe donc une homographie de  $\mathbb{P}^{*2}$  mettant ces droites en correspondance. En revanche,  $m'$  n'est pas la projection sur  $I'$  du point  $m$ . Dans la suite, nous explicitons la relation entre les deux points.

### 3.4 La matrice fondamentale

La droite épipolaire  $l$  passant par  $m$  et  $e$  est  $l = m \otimes e$ . L'application  $m \mapsto l$  est linéaire (§ 2.5) ; notons  $C$  sa matrice  $3 \times 3$  de rang 2 :  $l = Cm$ .

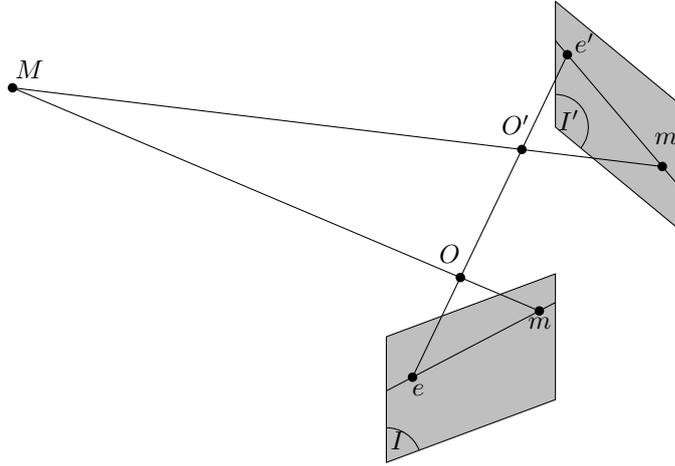


FIG. 7 – Projection stéréoscopique d'un point

Nous avons vu que l'application qui fait correspondre à une ligne épipolaire  $l$  sa conjuguée  $l'$  est une homographie de l'espace dual  $\mathbb{P}^{*2}$ . Notons  $A$  sa matrice, de rang 3 :  $l' = Al$ .

Le point  $m'$  appartient à la droite  $l'$ , donc  $m' \cdot l' = 0$ .

En combinant :  $0 = m' \cdot l' = m'^T l' = m'^T ACm$  l'équation de la contrainte qui lie, de manière symétrique, deux projections d'un point est donc de la forme :

$$m'^T Fm = 0 \quad (2)$$

On appelle  $F = AC \in \mathbb{R}^{3 \times 3}$  la *matrice fondamentale* de la paire d'images [2]. Elle est de rang 2 et définie à un facteur de proportionnalité près.

Comme suggéré plus haut, la connaissance de la matrice fondamentale permet de diminuer la combinatoire dans la mise en correspondance. En effet, pour tout point particulier  $m$  dans une image, il suffit de chercher son homologue le long de la droite épipolaire donnée par  $Fm$ . La mise en correspondance ne devient alors non seulement plus efficace mais aussi moins sujette à des erreurs.

On doit maintenant se poser la question comment obtenir la matrice fondamentale. La connaissance des matrices de projection associées deux images permet de calculer  $F$  (voir [3]). Or, les matrices de projection ne sont en pratique que connues lorsque l'on utilise un système stéréo, consistant de caméras rigidement attachées l'une à l'autre. Une situation beaucoup plus générale consiste à acquérir des images par une seule caméra, au cours d'un déplacement inconnu et a priori quelconque. C'est par exemple le cas lorsque nous prenons des photographies d'un édifice. La détermination de la matrice fondamentale dans cette situation générale est décrite dans la section suivante.

## 4 Détermination de la matrice fondamentale

La détermination de la matrice fondamentale  $F$  est la première étape de la reconstruction 3D à partir d'images. Typiquement, l'algorithme consiste à :

1. détecter sur les images  $I$  et  $I'$  des points particuliers ;
2. mise en correspondance (bien entendu sans utilisation de la contrainte donnée par la matrice fondamentale à ce stade). Ne retenir que des couples de points homologues « fiables », c'est-à-dire pour lesquels le point homologue trouvé est le seul point particulier dans l'autre image avec un voisinage très similaire ;
3. estimer la matrice  $F$  à partir de ces couples en utilisant l'équation (2).

Les étapes 1 et 2 d'analyse d'images et de mise en correspondance constituent un domaine de recherche à part. Nous retiendrons simplement que, quelle que soit la sophistication de l'algorithme, les couples de points détectés contiennent toujours des erreurs. Le premier type est le bruit, expliqué plus haut. Le second type d'erreurs est dû aux échecs de la mise en correspondance. Par exemple, sur la figure 8, il est difficile de distinguer les coins des différentes fenêtres par une analyse locale des images. Les couples de points sont alors inexploitable.

L'étape 3 de l'algorithme doit résister à ces entrées contaminées, c'est-à-dire que l'estimation de  $F$  doit être *robuste*. Nous traitons ce sujet dans la suite de l'article.

Remarquons qu'une fois la matrice fondamentale estimée, elle peut être utilisée pour refaire une mise en correspondance complète, qui en général va fournir beaucoup plus de points homologues correctes et ultérieurement une reconstruction 3D plus détaillée.

#### 4.1 Estimation de la matrice fondamentale – « Des maths aux maths appli »

La matrice  $F$  a 9 éléments, mais elle est singulière et définie à un facteur près, donc il y a 7 degrés de liberté.

D'après l'équation (2), un couple de points  $(m, m')$  fournit une contrainte linéaire sur  $F$  :

$$xx'F_{11} + xy'F_{12} + xF_{13} + yx'F_{21} + yy'F_{22} + yF_{23} + x'F_{31} + y'F_{32} + F_{33} = 0 \quad (3)$$

en notant  $F = [F_{ij}]_{i,j=1..3}$ ,  $(x, y, 1)$  les coordonnées de  $m$  et  $(x', y', 1)$  celles de  $m'$ .

Pour  $n$  couples  $(m_i, m'_i)$ , les contraintes s'écrivent matriciellement  $Af = 0$ , où  $A \in \mathbb{R}^{n \times 9}$  et  $f$  est un vecteur colonne des éléments de  $F$ .

Huit paires de points suffisent pour estimer  $F$  de cette manière et sept si en plus des équations du type (3) nous prenons en compte le fait que  $F$  doit être singulière.

#### 4.2 Résolution d'un système surdéterminé

Dans la pratique, l'ingénieur tient compte du premier type d'erreurs, le bruit, qui entache tout échantillon de 7 couples de points. Pour cela, il utilise  $n > 7$  couples pour estimer  $F$ , dans l'espoir que « les erreurs s'annulent ».

Le plus simple est alors de minimiser un critère de moindres carrés, qui s'écrit dans ce cas :

$$\text{Min}_{f \text{ t.q. } \|f\|=1} \|Af\|^2$$

la contrainte  $\|f\| = 1$  lève l'ambiguïté sur l'échelle de  $F$  et supprime la solution triviale  $f = 0$ .

Pour la résolution, supposons  $A$  de rang plein (donc  $n \geq 9$ ). Diagonalisons la matrice symétrique  $A^T A : A^T A = V \Lambda V^T$ , où  $V \in \mathbb{R}^{9 \times 9}$  est orthonormale et  $\Lambda$  diagonale. Alors :

$$\|Af\|^2 = f^T A^T A f = (V^T f)^T \Lambda (V^T f) = \sum_{i=1..9} \lambda_i \phi_i^2$$

où  $\lambda_1 \geq \dots \geq \lambda_9 > 0$  sont les valeurs propres et  $\phi_i$  les éléments du vecteur  $V^T f$ . La contrainte sur la norme de  $f$  devient  $\|V^T f\|^2 = \|f\|^2 = \sum_{i=1..9} \phi_i^2 = 1$ . Clairement,  $\|Af\|$  est minimal pour  $V^T f = (0, \dots, 0, 1)$ , donc  $f = v_9$ , le vecteur propre associé à la plus petite valeur propre de  $A^T A$ .

L'ingénieur obtient donc une estimation  $F_0$  de la matrice fondamentale, mais il constate qu'elle ne satisfait pas à la contrainte  $\text{rang}(F) = 2$ . C'est ennuyeux : ses lignes épipolaires ne sont pas concourantes. Il cherche alors à ajuster un peu son résultat, en optimisant l'expression

$$\text{Min}_{F \text{ t.q. } \text{rang}(F)=2} \|F - F_0\|_2 \quad (4)$$

où  $\|\cdot\|_2$  est la norme 2 sur les matrices :  $\|M\|_2 = \sup_{\|x\|=1} \|Mx\|$ .

L'optimisation de ce type d'expression fait intervenir la décomposition en valeurs singulières (*singular value decomposition*, SVD). Pour les matrices carrées, elle stipule que tout  $M \in \mathbb{R}^{n \times n}$  peut s'écrire  $M = U\Sigma V^\top$ , où  $U$  et  $V$  sont des matrices orthonormales et  $\Sigma$  une matrice diagonale dont les éléments (les *valeurs singulières*) vérifient  $\sigma_1 \geq \dots \geq \sigma_n \geq 0$ . Parmi les qualités de cette décomposition citons qu'elle est numériquement facile à calculer, et que  $\Sigma^2$  et  $V$  sont les matrices de valeurs et vecteurs propres de  $M^\top M$ .

Pour optimiser (4), on calcule la SVD de  $F_0$  :  $F_0 = U\Sigma_0 V^\top$ . Notons  $\Sigma = U^\top F V$ . On a :

$$\|F - F_0\|_2 = \sup_{\|x\|=1} \|U(\Sigma - \Sigma_0)V^\top x\| = \|\Sigma - \Sigma_0\|_2$$

comme  $U$  et  $V$  sont orthogonaux et conservent la norme, cela revient à trouver  $\Sigma$  de rang 2 le plus proche de  $\Sigma_0$  : il suffit pour cela d'annuler le plus petit coefficient de  $\Sigma_0$ .

La solution à (4) est donc :

$$F = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} V^\top$$

La figure 8 montre un résultat de cette approche.

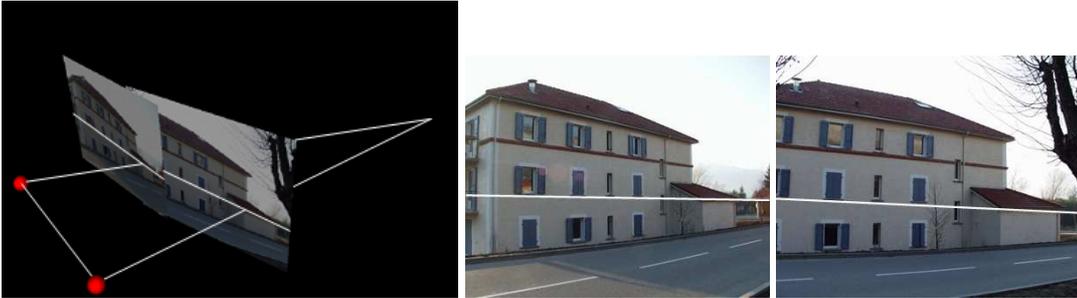


FIG. 8 – Résultat d'une estimation automatique de la matrice fondamentale. A gauche, illustration de la situation en 3D – le plan épipolaire coupe les plans image en les droites épipolaires (NB : Pour une visualisation plus aisée, les plans image sont ici situés *devant* les centres optiques). A droite, ces deux lignes épipolaires conjuguées affichées sur les images.

### 4.3 Stabilisation de la résolution – « Des maths appli à l'ingénierie »

L'ingénieur qui applique cette méthode se rend compte que les résultats sont toujours assez médiocres.

Analysons le problème. Dans une image de  $1024 \times 1024$  Pixels, un point a des coordonnées de l'ordre de 500. Dans l'équation (3) on trouve donc des valeurs telles que  $xx' \approx 500^2$  qui se reportent dans la matrice  $A$ . Cela conduit à des coefficients de l'ordre de  $500^4 \approx 5 \cdot 10^{10}$  dans  $A^\top A$ . À côté de ces valeurs énormes, nous avons toujours le coefficient 1 associé à  $f_{33}$  dans  $A$ . Cette inhomogénéité des éléments de  $A$  a deux conséquences :

- elle provoque une instabilité numérique de la résolution. En informatique, ajouter deux nombres d'ordres de grandeur différents cause une perte de précision sur le plus petit des deux dont le résultat final se ressent ;
- il y a du bruit sur les coefficients  $x, y, x', y'$  qui apparaissent dans des termes de degré 2 à l'équation (3). Les erreurs correspondant aux pixels de coordonnées les plus élevées (dans le coin inférieur droit) ont donc plus de poids dans le critère à minimiser. Les erreurs « ne s'annulent pas ».

L'ingénieur a une solution pratique à ces ennuis numériques : il effectue un changement de coordonnées de  $[0, 1023] \times [0, 1023]$  à  $[-1, 1] \times [-1, 1]$ . Les résultats deviennent enfin satisfaisants.

Le mathématicien se lasse cependant de ces approximations et ajustements. Il reformule le problème de la manière suivante : supposons que les erreurs de localisation des points  $(m_i, m'_i)_{i=1..n}$  suivent toutes indépendamment la même distribution gaussienne isotrope d'écart-type  $\sigma$  (c'est réaliste). La manière rigoureuse de poser le problème d'optimisation est alors :

$$\text{Min}_{\widehat{m}_i, \widehat{m}'_i, F} \text{ t.q. } \widehat{m}_i F \widehat{m}'_i = 0 \sum_i d(\widehat{m}_i - m_i)^2 + d(\widehat{m}'_i - m'_i)^2$$

c'est-à-dire qu'on estime simultanément  $F$  et les positions réelles  $(\widehat{m}_i, \widehat{m}'_i)$  des points à partir des observations bruitées de ces dernières. On note  $d(., .)$  la distance dans l'espace euclidien de deux points en coordonnées homogènes. D'un point de vue probabiliste, cette estimation atteint le maximum de vraisemblance au vu des hypothèses sur les erreurs.

L'ingénieur objectera qu'il n'y a pas de solution analytique à ce problème et que sa résolution itérative est beaucoup plus coûteuse en temps de calcul pour un gain parfois minime en précision. . .

#### 4.4 Le cas des erreurs de mise en correspondance

Un couple de points issu d'une erreur de mise en correspondance est inexploitable. Il faut donc s'en débarrasser dans l'estimation de  $F$ . Supposons connu un échantillon minimal de 9 couples de points sans erreur à partir duquel on peut estimer  $F$ . Cette première estimation permet de sélectionner parmi les couples restants ceux qui ne sont pas grossièrement faux.

Un processus peut alors être le suivant :

1. trouver une première estimation de  $F$  à partir de 9 couples non contaminés par des erreurs ;
2. utiliser cette estimation pour rejeter les mises en correspondances grossières. Pour le couple  $(m, m')$ , la correspondance est considérée grossière si  $m'$  est distante de plus d'un certain seuil de la ligne épipolaire sur  $I'$  correspondant à  $m$  ;
3. réestimer  $F$  à partir des couples retenus.

Le problème est de trouver les 9 couples non contaminés. On pourrait essayer toutes les géométries épipolaires déterminées par un ensemble de 9 couples. On retiendrait l'échantillon le plus « consensuel », c'est-à-dire celui pour lequel le nombre de mises en correspondances grossières est le plus faible. Cela conduirait à une combinatoire déraisonnable ( $C_n^9$  échantillons pour  $n$  de l'ordre de la centaine).

On préfère procéder à un sous-échantillonnage aléatoire de cet ensemble. La probabilité  $q$  d'avoir un échantillon de 9 ne comprenant pas de correspondance grossièrement erronée paraît faible ; elle est par exemple de  $q = (1/2)^9 = 1/512$  si 50% des correspondances sont erronées. Cependant, après  $k$  tirages aléatoires, la probabilité de n'en avoir aucun se réduit à  $(1 - q)^k$ . En prenant  $k = 2500$  (rapide sur les ordinateurs actuels), cette probabilité devient inférieure à 1%. Ce type d'algorithme est appelé RANSAC (RANdom SAMple Consensus, consensus d'échantillons aléatoires) et est d'une utilité très générale pour tout problème d'estimation avec des données dites grossières ou aberrantes. Il en existe de nombreuses variantes [4].

## 5 Reconstruction 3D à partir d'images non calibrées

Dans cette section, nous complétons notre étude du problème esquissé dans l'introduction de cet article. Grâce aux sections précédentes, nous savons maintenant comment calculer la matrice fondamentale pour une paire d'images d'une même scène, comment mettre en correspondance des

points particuliers et comment calculer la position des points 3D originaux. Cette dernière tâche, décrite au §3.2, nécessite la connaissance des matrices de projection, or généralement elles ne sont connues a priori que dans des cas de figure particuliers comme expliqué à la fin de la section 3.

Dans la suite, il est montré comment remonter aux matrices de projection, à partir des seules informations données par des points homologues dans les images. Ceci est décrit pour le cas général de deux images ou plus.

Soit une scène fixe, vue par deux caméras au moins. Supposons que  $p$  points 3D  $M_i, i = 1, \dots, p$  sont observés par  $c$  caméras de matrices de projection  $P_j, j = 1, \dots, c$ . Les coordonnées  $M_i$  et les matrices  $P_j$  sont inconnues, mais nous disposons, grâce à la mise en correspondance, pour chaque image  $j$ , des coordonnées  $m_{ij}$  de la projection de chaque point  $m_i$ .

## 5.1 Reconstruction projective

Un simple décompte des paramètres montre que nous avons  $2nm$  mesures et  $11m + 3n$  inconnues (11 pour chaque matrices de projection et 3 par point 3D), donc le problème devrait être soluble s'il y a assez de points et d'images. Cependant, rien ne fixe le système de coordonnées 3D, donc la solution n'est définie qu'à un changement de système de coordonnées près. Plus généralement, toute transformation linéaire en coordonnées homogènes 3D de la solution est aussi une solution. En effet,

$$m_{ij} \sim P_j M_i \quad i = 1, \dots, p, \quad j = 1, \dots, c \quad (5)$$

On peut donc transformer  $P_j \rightarrow P_j H^{-1}$  et  $M_i \rightarrow H M_i$ , avec  $H \in \mathbb{R}^{4 \times 4}$ .

Ainsi, sans plus de contraintes, la reconstruction n'est possible qu'à une transformation projective 3D près. Modulo cette ambiguïté fondamentale, la solution est unique en général.

Il y a différentes manières pour procéder à la reconstruction 3D. Souvent, on tente d'abord de calculer les matrices de projection, puis de les utiliser afin de reconstruire les positions des points 3D, en utilisant une version généralisée pour plusieurs images, de l'approche du §3.2. Quant aux matrices de projection, elles peuvent être calculées pour des paires d'images à partir de la matrice fondamentale (nous omettons les détails de ce calcul). Une manière plus simple de procéder est de prendre une base projective liée aux points 3D [2]. Pour cela, il suffit de choisir cinq points sans qu'il y en ait quatre coplanaires (cf. §2.6).

**Exercice 5.1** : *Etant donnés les projetés de quatre points 3D dans deux images et la géométrie épipolaire de cette paire d'images, montrez comment on peut décider si ces points 3D sont coplanaires ou non. Conseil : considérer, dans les images, les intersections de paires de droites dont chacune passe par deux des quatre points image.*

Une approche alternative est de sélectionner comme origine et point unité les centres de projection des deux premières caméras, et de compléter la base avec trois points visibles. Là encore, aucun quadruplet de points ne doit être coplanaire.

**Exercice 5.2** : *Concevez un test pour décider si, à partir de leurs projetés, trois points sont coplanaires avec le centre de projection. Concevez un test pour vérifier si deux points sont coplanaires avec la droite joignant les deux centres de projections d'une paire stéréo dont la géométrie épipolaire est donnée. Déduisez-en un test pour vérifier si cinq points forment une base projective valide.*

Soient  $m_1, m_2, m_3$  et  $m'_1, m'_2, m'_3$  les projetés dans les images 1 et 2 des points 3D  $M_1, M_2, M_3$ . Appliquons une transformation projective à chacune des images telle que ces trois points et l'épipole deviennent une base projective standard :

$$m_1 = m'_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad m_2 = m'_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad m_3 = m'_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad e = e' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Fixons également les coordonnées 3D de  $M_1, M_2, M_3$  :

$$M_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Avec ces choix, les matrices de projection sont données par :

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad P' = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

**Exercice 5.3** : *Prouvez que les matrices de projection ont la forme donnée ci-dessus.*

Cette approche peut être généralisée afin de calculer les matrices de projection de toutes les images.

Cette section ne montre que de simples exemples pédagogiques, donnés sans trop de détails. Remarquons qu'en pratique, des résultats précis requièrent une optimisation non linéaire aux moindres carrés en partant de cette solution algébrique initiale.

## 5.2 Reconstruction euclidienne

Dans la section précédente il est expliqué qu'a priori, la reconstruction 3D ne peut être obtenue qu'à une transformation projective inconnue près. On peut s'interroger sur l'utilité de tels modèles 3D... Il s'avère pourtant que passer d'une reconstruction dite projective à une reconstruction dite euclidienne, beaucoup plus utile, n'est pas trop difficile.

Une méthode relativement brutale pour passer d'une reconstruction projective à une reconstruction euclidienne est d'injecter sous forme d'équations les contraintes euclidiennes connues sur le monde : orthogonalité de droites, connaissances a priori de certaines coordonnées, etc. Pour cela il faut trouver une transformation projective  $H$  de l'équation (5) qui envoie la reconstruction projective dans un espace qui satisfera un ensemble de contraintes euclidiennes qu'il faut souhaiter redondantes.

On doit ainsi résoudre un ensemble d'équations hautement non linéaires pour lequel on minimisera simultanément l'erreur de mesure et le respect des contraintes. Ces équations étant hautement non linéaires, il est souhaitable d'avoir une première solution approximative depuis laquelle des processus itératifs vont chercher à approcher la solution réelle. Dans nos expérimentations, ces solutions approchées ont été obtenues par une simplification des hypothèses : on suppose que les projections perspectives sont des projections affines. Sous cette hypothèse, une reconstruction affine peut être obtenue de manière efficace et robuste en utilisant la décomposition en valeurs singulières vue en section 4.2 ([7]).

Les figures 9 et 10 montrent un exemple de ce processus.

## 6 Pour conclure et mettre la main à la pâte

Nous espérons avoir amené le lecteur à ce que nous avons promis dans l'introduction : une promenade agréable à travers une modélisation utilisant de vieux outils géométriques. Pour le problème de la reconstruction 3D, les parties précédentes ont illustré le besoin de concilier une modélisation mathématique avec des outils de résolution numériques. La prise en compte des erreurs dans les mesures, démarche classique de la physique appliquée, conduit à des développements supplémentaires.



FIG. 9 – Trois images d’une maison. Sont montrés par des croix les points image utilisés pour la reconstruction 3D. La base de la reconstruction projective initiale est attachée aux cinq points marqués par des disques blancs dans l’image 3.

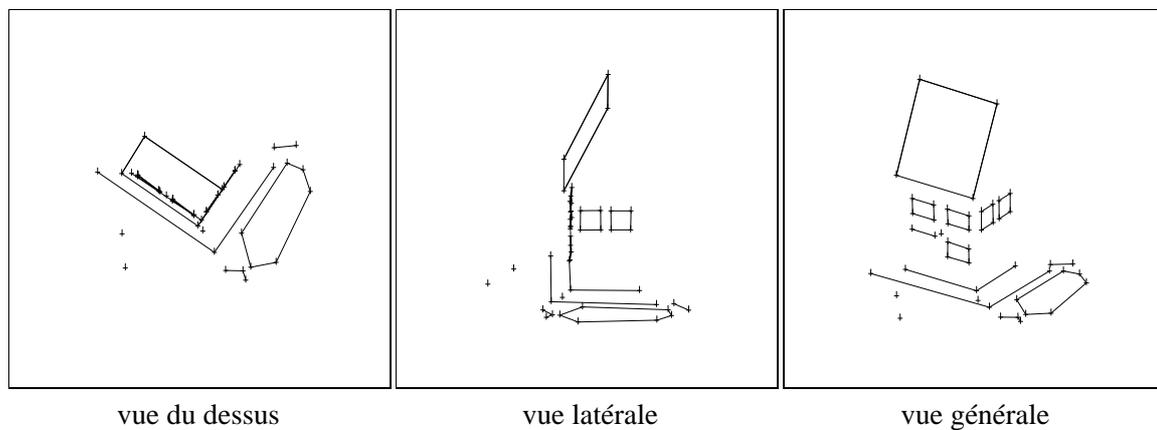


FIG. 10 – Reconstruction euclidienne de la maison, obtenue en utilisant la connaissance des positions réelles des cinq points. Afin de faciliter l’interprétation du résultat, les points reconstruits sont liés par des segments de droite.

Il n’en reste pas moins qu’il reste des points dans l’ombre, et seul un investissement plus fort, avec une dimension expérimentale pour valider les idées, permettrait d’aller plus au fond des choses.

Voici quelques questions qui n’ont pas été posées, ou qui n’ont eu que des réponses partielles :

- en 4.4 la sélection des points se fait de manière entièrement aléatoire. Pour l’estimation d’une première solution qui ne soit pas trop affectée par l’erreur sur les 9 points, leurs positions sont-elles indifférentes ? En déduire une amélioration dans le processus de sélection aléatoire des points.
- La matrice fondamentale étant estimée, comment corriger les positions des points mis en correspondance pour qu’ils satisfassent exactement la contrainte épipolaire (équation (2)) ?
- Comment reconstruire non seulement un ensemble de points 3D distincts, mais une surface ?

Un résultat récent, issu de nos travaux [1] est montré sur la figure 11.

Pour mener à bien des expérimentations, on pourra utiliser le site <http://pascal.inrialpes.fr/corresp/ImgSelectionEng.php> qui permet, à partir d’images d’une scène prises sous deux angles de vues différents, d’obtenir des mises en correspondance de manière automatique.

Les bibliothèques de calcul pour les différentes opérations sont disponibles sur <http://lear.inrialpes.fr/software>.

**Remerciements.** Nous remercions Pierre Bouchard pour une relecture détaillée de cet article et des remarques pertinentes qui ont permis d’en améliorer la forme et le contenu.

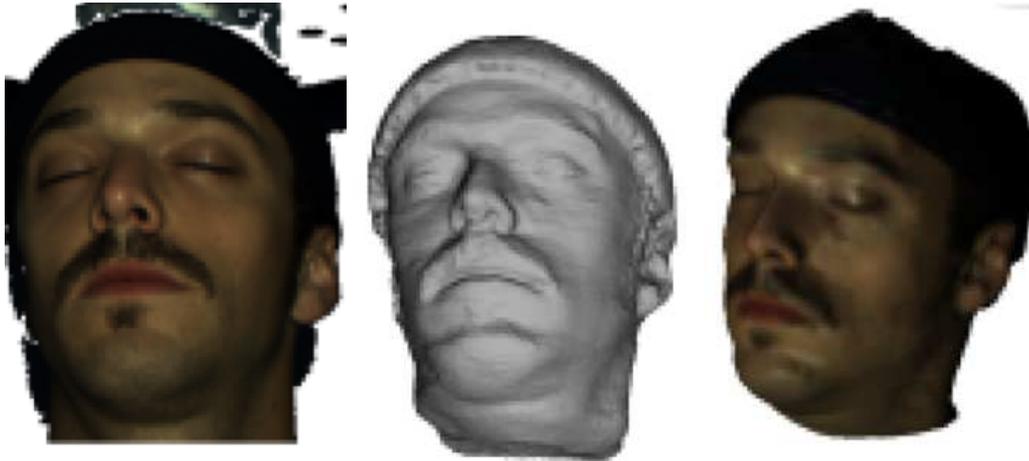


FIG. 11 – Gauche : une image originale. Milieu : modèle 3D reconstruit à partir de quelques dizaines d’images. Droite : le même modèle 3D, enrichi par des informations de couleur extraites des images originales.

## Références

- [1] N. BIRKBECK, D. COBZAS, P. STURM et M. JAGERSAND : Variational shape and reflectance estimation under changing light and viewpoints. Actes du *European Conference on Computer Vision*, Graz, Autriche, pp. 536–549, 2006.
- [2] O. FAUGERAS : What can be seen in three dimensions with an uncalibrated stereo rig ? Actes du *European Conference on Computer Vision*, Santa Margherita Ligure, Italie, pp. 563–578, 1992.
- [3] R.I. HARTLEY et A. ZISSERMAN : *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [4] P.J. ROUSSEEUW : Introduction to positive-breakdown methods. *Handbook of Statistics*, 15:101–121, 1997.
- [5] J.G. SEMPLE et G.T. KNEEBONE : *Algebraic Projective Geometry*. Oxford Science Publication, 1952.
- [6] P. STURM : *Vision 3D non calibrée : contributions à la reconstruction projective et étude des mouvements critiques pour l’auto-calibrage*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1997.
- [7] C. TOMASI et T. KANADE : Shape and motion from image streams under orthography : A factorization method. *International Journal of Computer Vision*, 9(2):137–154, 1992.