



Quelques contributions des invariants projectifs à la vision par ordinateur

Luce Morin

► To cite this version:

Luce Morin. Quelques contributions des invariants projectifs à la vision par ordinateur. Interface homme-machine [cs.HC]. Institut National Polytechnique de Grenoble - INPG, 1993. Français. NNT : . tel-00005137

HAL Id: tel-00005137

<https://theses.hal.science/tel-00005137>

Submitted on 26 Feb 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

présentée par

Luce MORIN

pour obtenir le grade de **DOCTEUR**

de l'**INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE**

(Arrêté ministériel du 30 mars 1992)

Spécialité: Informatique

—

**QUELQUES CONTRIBUTIONS DES INVARIANTS
PROJECTIFS
A LA VISION PAR ORDINATEUR**

—

Date de soutenance: 12 Janvier 1993

Composition du jury:

Président:	Jean-Marc CHASSERY
Rapporteurs:	Serge CASTAN Michel DHOME
Examineurs:	Philippe CINQUIN Roger MOHR

Thèse préparée au sein du laboratoire LIFIA-IRIMAG

sous la direction de **Roger MOHR**

Remerciements

Je dois tout d'abord remercier le GDR Traitement du Signal et Des Images qui m'a accordé sa confiance en finançant la préparation de cette thèse.

Je remercie ensuite Messieurs Jean-Marc Chassery, Phillipe Cinquin, Serge Castan et Michel Dhome qui m'ont fait l'honneur de participer au jury de ma thèse, et plus particulièrement ces deux derniers pour avoir accepté d'être les rapporteurs de mon manuscrit.

Je tiens ensuite à remercier Roger Mohr, qui a encadré et soutenu mon travail de thèse, ainsi que Patrick Gros, Françoise Veillon, Emmanuel Arbogast, Radu Horaud, Long Quan, Danièle Herzog et tous les membres de l'équipe Voir. pour leur aide et leur accueil pendant ces trois années.

Cette thèse doit aussi beaucoup à Enrico Grosso et Kalle Astrom et aux travaux effectués en commun lors de leur visite au Lfia.

Enfin, j'ai apprécié d'avoir préparé cette thèse dans l'environnement chaleureux du Lfia.

Sommaire

Introduction	9
1 Introduction à la géométrie projective	15
1.1 Introduction	15
1.2 Espace projectif	16
1.2.1 Définition	16
1.2.2 Base projective	16
1.2.3 Lien entre $\mathbb{P}^n$ et $\mathbb{R}^n$	17
1.3 Transformations projectives	17
1.3.1 Définition	17
1.3.2 Nombre de points déterminant une transformation projective	18
1.3.3 Transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$ et projections .	18
1.4 Invariants projectifs	19
1.4.1 Qu'est-ce qu'un invariant?	19
1.4.2 Le birapport	21
1.4.3 Invariants projectifs des transformations de $\mathbb{P}^3$ dans $\mathbb{P}^2$. .	23
1.4.4 Coordonnées projectives	24
1.5 Propriétés utiles	27
1.5.1 Conique de Chasles	27
1.5.2 Principe de dualité	30
1.5.3 Géométries affine et euclidienne	31

2	Application au positionnement relatif	35
2.1	Introduction	35
2.1.1	Modèles de projection	36
2.1.2	Rétroprojection	37
2.1.3	Reconstruction avec étalonnage	37
2.1.4	Positionnement relatif	39
2.2	Reconstruction en utilisant des points coplanaires	41
2.2.1	Comment rétroprojeter un point image?	41
2.2.2	Où est la caméra?	42
2.2.3	Où est le point image?	42
2.2.4	Comment localiser un point dans la scène?	43
2.2.5	Comment déterminer les épipolaires?	43
2.3	Reconstruction avec 6 points non coplanaires	44
2.3.1	Où est la caméra?	45
2.3.2	Comment rétroprojeter un point image?	46
2.3.3	Où est le point image?	46
2.3.4	Positionnement relatif symbolique	47
2.4	Expérimentations	49
2.4.1	Reconstruction dans le plan	49
2.4.2	Reconstruction 3D	50
2.5	Conclusion	52
3	Précision du positionnement relatif	55
3.1	Introduction	55
3.2	Analyse différentielle de la reconstruction par le birapport	56
3.3	Calcul des erreurs dans un cas particulier	58
3.3.1	Reconstruction de P	59
3.3.2	Calculs d'erreur	60
3.3.3	Présentation des courbes d'erreur	62
3.4	Conclusion	65
4	Stabilité des invariants projectifs	67
4.1	Introduction	67
4.2	Origine des erreurs	70
4.3	Etude théorique de stabilité du birapport	72
4.3.1	Birapport de quatre points alignés	72
4.3.2	Birapport de cinq points coplanaires	85

4.4	Les invariants de cinq points coplanaires	91
4.4.1	Invariants de cinq points ordonnés	91
4.4.2	Invariants de cinq points non ordonnés	91
4.4.3	Invariants de cinq points : cas de la conservation de l'enveloppe convexe	94
4.5	Simulations	96
4.5.1	Données utilisées pour les test	96
4.5.2	Comparaison de la stabilité	97
4.5.3	Comparaison des performances	104
4.6	Conclusion	106
Conclusion		109
Annexes		112
A Transformations projectives et projections : démonstration		113
A.1	Matrice de projection perspective	114
A.2	Matrice de projection orthographique	116
A.3	Matrice de transformation projective 3D-2D	119
A.3.1	A de rang 3	119
A.3.2	A de rang 2	120
A.3.3	A de rang 1	122
A.4	Démonstration des lemmes	122
B Intersection de deux coniques		125
Bibliographie de l'auteur		127
Bibliographie		129

Introduction

La vision par ordinateur est le cadre général de ce travail. Pour la définir en une phrase, cette discipline s'attache à extraire de manière automatique des informations à partir d'images, en vue d'effectuer des tâches spécifiques.

Par *image* on entend un tableau bidimensionnel de valeurs dont chaque élément (*pixel*) correspond à la mesure d'une grandeur physique. De manière générale, de par son caractère bidimensionnel d'une part, par la restriction de la mesure à une seule grandeur physique d'autre part et enfin par le phénomène de discrétisation, une image représente une perte d'information importante par rapport à la scène tridimensionnelle initiale. Il est donc important, pour pouvoir analyser une image, de connaître le mécanisme de formation de cette image. Dans notre cas, les images utilisées sont issues de caméras CCD noir et blanc. La grandeur mesurée est donc l'intensité lumineuse dans le spectre de la lumière visible.

Les traitements effectués sur ce type d'image en vision par ordinateur sont généralement classés en trois niveaux [Mar 78].

Un premier niveau de traitement, appelé souvent “vision bas-niveau”, est la segmentation, ou regroupement des pixels en ensembles cohérents. On peut s'attacher à différents types de cohérence : on cherche à regrouper les pixels correspondant au même objet, à la même ligne de discontinuité, à la même couleur ou la même intensité, au même mouvement etc. . .

Un deuxième niveau concerne la restitution des caractéristiques spatiales, géométriques et cinématiques de la scène observée : ce sont les domaines de la reconstruction, du positionnement, de l'analyse du mouvement, et aussi de la reconnaissance d'objet, lorsque celle-ci est basée sur la description géométrique des objets.

Un troisième niveau s'attache au contenu sémantique de la scène observée : il s'agit alors de détecter la présence et d'identifier les objets observés comme

appartenant à une catégorie prédéfinie.

Notre travail se situe au niveau intermédiaire de la restitution des caractéristiques géométriques de la scène observée, soit dans le but de déterminer effectivement la forme et la position des objets tridimensionnels, soit dans le but de détecter dans l'image la présence d'objets connus a priori. Les applications sont nombreuses, tant dans le domaine de la robotique pour l'aide à la préhension ou au déplacement autonome dans un environnement éventuellement connu, que dans le domaine médical, pour l'aide au diagnostic ou l'aide au geste chirurgical. Un autre domaine d'application important est celui de la télédétection.

Le premier point auquel nous nous intéressons est la perception tridimensionnelle. Pour restituer la forme tridimensionnelle des objets à partir d'images, l'approche traditionnellement utilisée est la suivante [Bal 82] [Aya 87] [Sko 88].

- La projection effectuée par la caméra est approximée. Les modèles les plus couramment utilisés sont ceux de la projection perspective et de la projection parallèle (ou projection orthographique). Cette dernière conduit à une formulation linéaire des équations de projection. Des modèles plus précis doivent parfois être utilisés, comme une projection perspective, distordue par un difféomorphisme du plan.
- Les valeurs des paramètres intervenant dans ce modèle sont évaluées, à l'aide d'objets dont la position est parfaitement connue (mire ou balises). C'est l'étalonnage de la caméra [Tsa 87].
- Pour chaque point de l'image, on détermine alors le(s) point(s) de la scène susceptible(s) de se projeter en ce point : c'est le problème de la rétroprojection.
- Si on dispose de plusieurs images, les points de chaque image sont mis en correspondance afin de regrouper les points qui sont les projetés d'un même point de la scène. La position de ce point est ensuite déterminée par triangulation.

Cette méthode se heurte à plusieurs difficultés. L'étalonnage est une étape délicate et très sensible aux erreurs de mesure. Par ailleurs, son utilisation suppose que les paramètres de la caméra ne soient pas modifiés au cours du temps. Ceci est pourtant le cas si la caméra est déplacée entre deux prises de vues ou si on effectue une mise au point. De plus, l'étalonnage n'est pas fiable au cours du temps : une caméra immobile, à focale fixe, peut voir ses paramètres biaisés, par un changement des conditions thermiques par exemple. Ceci restreint beaucoup le champ des applications pour lesquelles cette méthode peut être utilisée. C'est pourquoi des alternatives ont été recherchées, et les approches permettant d'éviter entièrement ou partiellement l'étape d'étalonnage suscitent un intérêt croissant.

Certaines techniques de positionnement ne supposent connus qu'une partie des paramètres des caméras [Hor 89], [Yua 89], [Liu 90]. Il est cependant possible d'uti-

liser pour le positionnement des caméras pour lesquelles on ne dispose d'aucun étalonnage, même partiel. En effet les images provenant d'une caméra non étalonnée fournissent des informations sur la position relative des objets observés. C'est en partie ce qui nous permet de reconstituer la scène que nous voyons sur une photographie, même si nous ne connaissons ni la position de l'appareil photo au moment de la prise de vue, ni les caractéristiques optiques de l'objectif (focale, angle de vue ...). Cette information sur la position relative des objets peut être atteinte en utilisant les propriétés géométriques du modèle de projection employé. Si on adopte le modèle de projection parallèle, les propriétés de la géométrie affine sont disponibles et il est possible de calculer une description affine de la scène observée. Ceci est mis en œuvre dans plusieurs travaux : on peut citer ceux de Koenderink et Van Doorn [Koe 89], Sparr et Nielsen [Spa 90] et de Tomasi et Kanade [Tom 91]. Si le modèle de la projection perspective est utilisé, la transformation entre la scène et l'image est une transformation projective. Les propriétés de la géométrie projective peuvent alors être appliquées.

Notre travail s'inscrit dans ce cadre : nous utilisons le modèle de la projection perspective, que nous supposons valide, et nous cherchons alors à déterminer les informations que nous procurent les propriétés de la géométrie projective. Notre première contribution est une méthode de positionnement relatif à l'aide des invariants projectifs. Nous considérons deux images d'une même scène, provenant de caméras non étalonnées, et dont les points sont supposés appariés. Nous montrons alors que si nous connaissons la position dans l'espace de six points dits *de référence* (constitués de deux groupes de quatre points coplanaires ou de six points en position quelconque), alors nous pouvons déterminer la position tridimensionnelle de tous les autres points observés. Cette méthode ne nécessite aucune connaissance sur les caméras, ni leur position, ni leurs caractéristiques optiques. La seule hypothèse concernant les caméras est la validité du modèle de projection perspective. Si la position des points de référence n'est pas connue, ou est seulement connue approximativement, notre méthode permet d'établir la position relative des points observés par rapport aux points de référence.

Dans le cas où on cherche à déterminer les caractéristiques tridimensionnelles des objets, non pas pour les positionner, mais pour les reconnaître, les propriétés de la géométrie projective sont aussi utilisables. En effet, il existe des invariants projectifs, c'est-à-dire des quantités géométriques se conservant au cours des transformations projectives : ils ont donc une valeur identique dans la scène et dans l'image, ou dans deux images d'une même scène. Le modèle de projection parallèle, lié à la géométrie affine, fournit déjà les invariants associés aux affinités [Hut 91] [Gro 91]. Cependant, l'écart entre le modèle de la projection parallèle et la projection effectuée par une caméra réelle ne permet d'utiliser ces invariants que dans des conditions très restreintes, par exemple pour des points de vue très proches. Par contre les invariants projectifs sont conservés entre deux images d'une même vue, quelque soit la position respective des points de vue de chaque image. Ils se présentent donc comme de bons candidats pour les processus de reconnaissance d'objets

et plusieurs travaux concernent des systèmes de reconnaissance basés sur les invariants projectifs [For 90a] [For 90b] [For 91a] [Rot 91] [For 91b] [For 92] [Bat 91]. Coelho et al. [Coe 91] ont étudié la variation de plusieurs invariants en fonction du point de vue, et Lenz et Meer [Len 92] ont soulevé le problème de la symétrisation des invariants projectifs. Cependant peu d'études ont été menées pour évaluer la sensibilité de ces derniers au bruit de mesure ainsi que pour déterminer leur pouvoir de discrimination.

Notre seconde contribution est donc une étude de la stabilité des invariants projectifs. Nous présentons des résultats théoriques sur les invariants projectifs (en particulier l'expression de la fonction de répartition du birapport), et nous proposons des critères de sélection des invariants projectifs dans le cadre d'un processus de reconnaissance.

Le plan suivi dans ce mémoire est le suivant.

Dans un premier chapitre sont introduites les notions fondamentales de la géométrie projective : espace projectif, base projective, transformations projectives et invariants projectifs. Nous présentons aussi les propriétés que nous utilisons dans la suite : le théorème de Chasles, le principe de dualité, la relation entre géométrie projective et géométries affine et euclidienne. En plus des notions générales, nous avons développé dans ce chapitre quelques aspects plus pointus. En particulier, nous explicitons le lien entre les transformations projectives de l'espace projectif dans le plan projectif et les modèles de projection affine et perspective. Nous montrons aussi comment on peut caractériser un point par des coordonnées construites à partir d'invariants projectifs et nous donnons la relation entre ces coordonnées projectives et les coordonnées homogènes habituellement utilisées.

Un deuxième chapitre présente en détails la méthode de positionnement relatif à partir des invariants projectifs. Outre le positionnement des points observés, nous montrons qu'à partir des mêmes hypothèses de départ, nous pouvons aussi déterminer la position de la caméra par rapport aux points de référence, prédire la position d'un point dans une nouvelle image où on connaît les projetés des points de référence, et déterminer la géométrie épipolaire dans le cas de la stéréovision. Les résultats obtenus avec cette méthode sur des images réelles contenant des objets polyédriques sont présentés. Ils nous permettent de valider la méthode et d'évaluer de manière quantitative la précision obtenue.

Une étude de précision du positionnement relatif dans le cas de caméras mobiles constitue le troisième chapitre. Nous cherchons d'abord à évaluer l'incidence des erreurs affectant les points de l'image sur le point reconstruit à l'aide de birapports. Une étude différentielle nous permet de constater qu'en dehors des configurations dégénérées, la précision obtenue est identique à celle obtenue avec un caméra parfaitement étalonnée. Nous cherchons ensuite à comparer de manière quantitative les précisions obtenues en présence d'incertitudes avec une méthode utilisant l'étalonnage et avec une méthode utilisant les invariants projectifs. Pour cela nous considérons une configuration stéréo plane (composée de deux caméras différentes

ou d'une caméra mobile), pour laquelle il existe une incertitude sur la position des caméras. C'est le cas typique d'une caméra fixée sur un robot mobile et dont la position est donnée par l'intermédiaire de capteurs odométriques. Nous dérivons les calculs de reconstruction, soit en supposant les caméras étalonnées, soit en utilisant trois points de référence alignés et la méthode géométrique. Nous constatons que certaines configurations sont dégénérées pour la méthode géométrique : pour certains points de vue, cette méthode ne peut être appliquée. Par contre, en dehors de ces configurations, la méthode géométrique est moins affectée par les erreurs sur la position des caméras que le positionnement effectué par l'intermédiaire de l'étalonnage pour lesquelles elle dégénère. Par contre, les deux méthodes utilisent la position des caméras et nous montrons que la méthode géométrique est moins sensible à l'erreur sur ces positions, dès lors qu'on s'éloigne des configurations dégénérées.

Le dernier chapitre présente notre étude sur la stabilité des invariants projectifs. Nous cherchons à caractériser la sensibilité des invariants projectifs aux erreurs de mesure dans les images, ainsi qu'à déterminer leur pouvoir de discrimination dans le cadre d'un processus de reconnaissance. Nous commençons par une étude de la stabilité du birapport de quatre points alignés. Nous dérivons l'expression analytique de la fonction de répartition du birapport, pour une distribution uniforme des abscisses des quatre points. Ce résultat nous donne une mesure du pouvoir discriminant d'un birapport en fonction de sa valeur. Il nous permet aussi de définir un critère de similarité entre deux valeurs de birapports. Une étude différentielle met en évidence les configurations dégénérées et aboutit à une évaluation a priori de l'incertitude sur le birapport de quatre points connaissant les coordonnées des points et les incertitudes sur ces coordonnées. Nous nous intéressons ensuite aux invariants de cinq points coplanaires. Ce sont ceux qui sont généralement utilisés pour la reconnaissance, car cinq points coplanaires constituent la configuration générale la plus simple admettant des invariants projectifs dans une projection perspective 3D-2D. Une étude théorique du birapport issu de cinq points coplanaires mène à des conclusions analogues à celles obtenus avec le birapport de quatre points alignés. Nous proposons ensuite différents invariants projectifs de cinq points coplanaires, et nous comparons leur stabilité et leurs performances pour la reconnaissance sur des simulations bruitées.

Le chapitre de conclusion résume les différents résultats et présente les perspectives de recherche envisagées à la suite de ce travail.

Chapitre 1

Introduction à la géométrie projective

1.1 Introduction

Ce sont les mathématiciens de la Grèce antique qui les premiers ont introduit la géométrie projective. Ils ont énoncé ses propriétés géométriques et découvert son invariant fondamental : le birapport. La géométrie projective a ensuite été utilisée par les peintres de la Renaissance qui cherchaient à rendre l'effet de profondeur dans leurs toiles, c'est à dire à simuler l'effet de la vision humaine. C'est au XIX^{ème} siècle que la géométrie projective a été étudiée de manière approfondie et qu'ont été énoncés tous les théorèmes que nous connaissons aujourd'hui. On peut citer les travaux de Poncelet, qui dans son *Traité des propriétés projectives des figures* (1822) a étudié les propriétés invariantes par une série de projections perspectives d'un plan dans un autre. Klein, dans son discours *Erlangen Programme* (1872) a généralisé la notion d'invariant associé à un groupe de transformations. Les photogrammètres ont depuis plusieurs décennies utilisé ces propriétés, mais c'est récemment que la géométrie projective et ses invariants ont fait l'objet d'un intérêt croissant dans la communauté de la vision artificielle.

Les concepts fondamentaux de la géométrie projective sont introduits dans ce chapitre. Il présente les notions, le vocabulaire et les notations qui seront utilisés

dans les chapitres suivants.

1.2 Espace projectif

1.2.1 Définition

On définit l'espace projectif $\mathbb{P}^n$ comme l'espace quotient de $\mathcal{C}^{n+1} \setminus \{(0, \dots, 0)\}$ par la relation d'équivalence :

$$(x_1, \dots, x_{n+1}) \sim (x_1', \dots, x_{n+1}') \\ - \exists \lambda \neq 0 \quad / \quad (x_1', \dots, x_{n+1}') = \lambda(x_1, \dots, x_{n+1}).$$

Ainsi les deux $(n+1)$ -uplets $(x_1, \dots, x_{n+1})$ et $\lambda(x_1, \dots, x_{n+1})$ représentent le même point de l'espace projectif $\mathbb{P}^n$.

Un point de $\mathbb{P}^n$ est donc représenté par ses *coordonnées homogènes* $(x_1, \dots, x_{n+1})$ qui sont définies à un facteur de proportionnalité près.

1.2.2 Base projective

Une *base projective* de $\mathbb{P}^n$ est un ensemble de $n + 2$ points dont tous les sous-ensembles de $n + 1$ points sont linéairement indépendants.

Définissons d'abord la notion de points linéairement dépendants. Des points sont dits linéairement dépendants s'il existe une relation linéaire liant leurs coordonnées homogènes.

La base projective canonique est formée des $n + 2$ points suivants : $(1, 0, \dots, 0)$, $(0, 1, \dots, 0)$, $\dots$ $(0, \dots, 1)$ et $(1, 1, \dots, 1)$. Les $n + 1$ premiers points de la base canonique sont appelés *points de référence* et le dernier point est appelé *point unité*, selon la terminologie employée par Semple et Kneebone.

Une base projective définit de manière unique un système de coordonnées pour $\mathbb{P}^n$ dans lequel elle est la base projective canonique : pour définir un système de coordonnées de $\mathbb{P}^n$, il suffit donc de choisir $n + 2$ points formant une base projective, et de leur attribuer les coordonnées $(1, 0, \dots, 0)$, $(0, 1, \dots, 0)$, $\dots$ $(0, \dots, 1)$ et $(1, 1, \dots, 1)$.

Un changement de base consiste en une transformation linéaire : les coordonnées homogènes Y d'un point dans la nouvelle base s'expriment en fonction des coordonnées X du point dans l'ancienne base et d'une matrice M de taille $(n + 1) \times (n + 1)$ de changement de base :

$$Y^t = M X^t$$

La matrice M est elle aussi définie à facteur de proportionnalité près et elle est complètement déterminée par le choix des points de la nouvelle base projective.

1.2.3 Lien entre $\mathbb{P}^n$ et $\mathbb{R}^n$

On établit généralement une correspondance entre les points de $\mathbb{R}^n$ et $\mathbb{P}^n$ par l'injection suivante :

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{P}^n \\ (x_1, \dots, x_n) &\longmapsto (x_1, \dots, x_n, 1) \end{aligned}$$

Certains points de $\mathbb{P}^n$ ne sont pas atteints par cette correspondance : ce sont les points de l'hyperplan de $\mathbb{P}^n$ d'équation $x_{n+1} = 0$.

Ces points ne sont associés à aucun point usuel de $\mathbb{R}^n$. Cependant, on peut faire correspondre à chaque point de $\mathbb{P}^n$ tel que $x_{n+1} = 0$, le n-uplet suivant :

$$\lim_{x_{n+1} \rightarrow 0} \left(\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}} \right)$$

Ce n-uplet peut être interprété soit comme le point à l'infini dans la direction $(x_1, \dots, x_n)$, soit comme la direction vectorielle $(x_1, \dots, x_n)$. Nous utiliseront généralement cette première interprétation et l'hyperplan $x_{n+1} = 0$ sera appelé *hyperplan de l'infini*.

Le choix de $x_{n+1} = 0$ comme hyperplan contenant les points de l'infini est totalement arbitraire : tout hyperplan de $\mathbb{P}^n$ peut être choisi comme hyperplan de l'infini, pour établir une correspondance entre $\mathbb{P}^n$ et $\mathbb{R}^n$. Pour la géométrie projective, il n'y a aucune distinction entre des points à l'infini et les autres : tous les points de $\mathbb{P}^n$ sont équivalents. On peut donc voir $\mathbb{P}^n$ comme une extension de $\mathbb{R}^n$, dont le formalisme permet de manipuler les points à l'infini de manière identique aux autres points de $\mathbb{R}^n$.

1.3 Transformations projectives

1.3.1 Définition

On appelle *transformation projective* toute application linéaire des coordonnées homogènes.

Une transformation projective de $\mathbb{P}^n$ dans $\mathbb{P}^m$ s'exprime donc sous forme matricielle : l'image Y d'un point X de $\mathbb{P}^n$ est définie par :

$$Y^t = AX^t \quad \text{où } A \text{ est une matrice de taille } (m+1) \times (n+1)$$

On peut remarquer que A et λA définissent la même application projective. La matrice associée à une projection perspective est donc définie à un facteur multiplicatif près : elle n'a que $(n+1)(m+1) - 1$ degrés de liberté.

1.3.2 Nombre de points déterminant une transformation projective

Cas des homographies

Les transformations de $\mathbb{P}^n$ dans lui-même, ou *homographies* forment un groupe. Une homographie équivaut à un changement de base. Elle est donc définie de manière unique par l'image d'une base projective de $\mathbb{P}^n$. Une démonstration complète de cette propriété peut être trouvée dans [Sem 52]. Nous ne ferons ici qu'une brève considération de degrés de liberté.

Déterminer une homographie est équivalent à déterminer à un facteur multiplicatif près la matrice qui lui est associée. Cette matrice contient $(n+1)^2 - 1 = n^2 + 2n$ paramètres indépendants.

Chaque correspondance entre un point de la base et son image fournit $(n+1)$ équations linéaires en les inconnues. n équations seulement sont indépendantes, puisque les points sont exprimés en coordonnées homogènes. On dispose donc de $n(n+2) = n^2 + 2n$ équations contenant les $n^2 + 2n$ inconnues. Dans le cas où les points utilisés forment une base projective, ces équations sont indépendantes et il existe une solution unique. Si les points ne forment pas une base projective, le nombre d'équations indépendantes est insuffisant pour déterminer tous les paramètres de la transformation projective : il existe alors plusieurs solutions. Une homographie de $\mathbb{P}^n$ est donc définie par les images de $n+2$ points formant une base projective.

Cas général

Dans le cas d'une transformation projective quelconque de $\mathbb{P}^n$ dans $\mathbb{P}^m$, l'image d'une base de $\mathbb{P}^n$ n'est pas nécessairement suffisante pour déterminer la transformation.

Envisageons le cas de la projection de $\mathbb{P}^3$ dans $\mathbb{P}^2$ qui nous intéresse tout particulièrement. Une telle transformation est déterminée par une matrice 3×4 à $3 \times 4 - 1 = 11$ paramètres indépendants. Chaque correspondance entre un point de $\mathbb{P}^3$ et son image dans $\mathbb{P}^2$ fournit $m = 2$ équations indépendantes. Une base de $\mathbb{P}^3$, constituée de 5 points ne fournit donc que 10 équations, et ne suffit pas pour déterminer la transformation.

Une transformation projective 3D-2D est donc déterminée par les images de 6 points de $\mathbb{P}^3$. Il faut remarquer cependant que 6 points fournissent 12 équations : la dernière est redondante.

1.3.3 Transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$ et projections

Si nous pouvons appliquer les propriétés de la géométrie projective au modèle de la projection perspective, c'est parce que le modèle de projection perspective que nous utilisons est une transformation projective de $\mathbb{P}^3$ dans $\mathbb{P}^2$. Inversement,

on peut se demander à quelles projections géométriques correspondent les transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$. En fait, nous avons démontré la propriété suivante.

Théorème 1 *Soit une transformation projective de $\mathbb{P}^3$ dans $\mathbb{P}^2$, représentée par une matrice P de taille 3×4 en coordonnées homogènes. Considérons dans $\mathbb{P}^3$ le plan $x_4 = 0$, auquel nous donnons un sens particulier : c'est le plan de l'infini. Nous définissons aussi un repère euclidien de l'espace, compatible avec le choix du plan de l'infini. La matrice P peut alors avoir différentes interprétations géométriques. Soit la matrice A de taille 3×3 formée des trois premières colonnes de la matrice P .*

Si A est de rang 3, alors P représente une projection perspective, c'est à dire une projection centrale, dont le centre n'appartient pas au plan de l'infini, et dont le plan de projection est différent du plan de l'infini et ne contient pas le centre de projection.

Si A est de rang 2, alors P représente la composition d'une projection orthogonale sur un plan et d'une homographie de ce plan. Dans le cas particulier où la dernière ligne de la matrice A est nulle, alors l'homographie est une transformation affine du plan. La matrice P représente alors une transformation parallèle de l'espace euclidien dans le plan de projection.

Si A est de rang 1, alors P est une projection de l'espace dans le plan mais tous les points se projettent sur une droite du plan.

La démonstration est donnée en annexe. La première partie de ce théorème (concernant le cas de la matrice A de rang 3) a aussi été démontrée par Faugeras dans son livre *3D Computer Vision*, à paraître [Fau 92b].

1.4 Invariants projectifs

Les invariants projectifs sont les invariants géométriques associés aux transformations projectives.

1.4.1 Qu'est-ce qu'un invariant?

Commençons par donner une définition très générale d'un invariant. On considère deux ensembles E et F et un ensemble T de transformations de E dans F . Soient e un élément quelconque de E et I une fonction des éléments de F . I est un invariant si elle prend la même valeur pour toutes les images de e par n'importe quel élément de T :

$$\forall e \in E \quad \forall t, t' \in T \quad I(t'(e)) = I(t(e))$$

Si $E = F$, on a plus simplement :

$$\forall e \in E \quad \forall t \in T \quad I(t(e)) = I(e)$$

On définit l'*orbite* O_e d'un point e de E comme l'ensemble des points images de e par les transformations de T .

$$O_e = \{t(e) | t \in T\}$$

Il découle des précédentes définitions que si I est un invariant pour les transformations T de E dans F , alors I est constant en tous les points d'une même orbite. Un invariant est dit *complet* si il prend des valeurs différentes sur chaque orbite, c'est à dire si et seulement si :

$$I(f_1) = I(f_2) \iff \exists e \in E \quad \text{tel que } f_1 \in O_e \text{ et } f_2 \in O_e$$

Cas des groupes Dans le cas où E est une variété et où T est un groupe continu, localement paramétrable, agissant sur E , on peut appliquer la théorie des invariants algébriques associés aux groupes de Lie et on dispose de propriétés supplémentaires sur les invariants. En particulier, dans ce cas, les orbites sont disjointes et forment une partition de l'ensemble vectoriel E .

On dispose aussi d'un théorème permettant de déterminer le nombre d'invariants indépendants. Précisons d'abord la notion d'*invariants indépendants*. Soit I un invariant d'une configuration x et soit f une fonction quelconque, alors $f(I(x))$ est aussi un invariant de x . On voit donc qu'à partir d'un invariant I , on peut générer une infinité d'invariants $f(I)$, mais liés par une dépendance fonctionnelle : ils ne peuvent pas varier indépendamment. Des invariants indépendants sont caractérisés par des dérivées partielles linéairement indépendantes. Le théorème suivant donne le nombre d'invariants indépendants associés au groupe de transformations G opérant sur un espace vectoriel E :

$$\text{Nombre d'invariants} = \dim E - (\dim G - \min_{x \in E} (\dim G_x))$$

où G_x est le *groupe d'isotropie* ou *groupe stabilisateur* de x . Il est défini par :

$$G_x = \{g \in G | g(x) = x\}$$

Par exemple, on peut calculer le nombre d'invariants indépendants associés au groupe G des homographies de $\mathbb{P}^2$ agissant sur l'espace E des quintuplets de $\mathbb{P}^2$. La dimension de G est 8 (nombre de coefficients indépendants de la matrice associée). La dimension de E est 10 (2 degrés de liberté par point). Si x est un quintuplet contenant une base projective, le groupe d'isotropie est réduit à l'identité. $\min_{x \in E} (\dim G_x)$ est donc nul. Le nombre d'invariants indépendants est donc $10 - (8 - 0) = 2$. Cinq points du plans admettent donc 2 invariants indépendants pour les homographies planes.

On dispose aussi de différentes méthodes (différentielle et symbolique) pour calculer analytiquement les invariants. Nous ne développerons pas ce point, et le lecteur est renvoyé à l'article de Forsyth [For 90a] et à celui de P. Gros [Gro 92].

1.4.2 Le birapport

Le birapport est l'invariant projectif fondamental : la plupart des invariants projectifs plus complexes peuvent s'exprimer en termes de birapports. Par exemple, la fonction algébrique utilisée par Zisserman et al. pour caractériser deux coniques coplanaires est basée sur les matrices symétriques associées à chacune des coniques. On peut montrer qu'une paire de coniques est caractérisée par des invariants s'exprimant en fonction de birapports. Leur calcul ne nécessite alors que des constructions géométriques simples sur les coniques [Qua 92].

Birapport de quatre points alignés

Soient A, B, C, D quatre points colinéaires, nous définissons leur *birapport*, noté $[A, B, C, D]$, comme :

$$[A, B, C, D] = \frac{\overline{AC} \times \overline{BD}}{\overline{AD} \times \overline{BC}} \quad (1.1)$$

où $\overline{AB}$ est la mesure algébrique de AB . Cette formulation du birapport est valable pour les points situés à l'infini à condition d'utiliser les conventions suivantes :

$$\frac{\infty}{\infty} = 1, \quad \frac{a}{\infty} = 0, \quad \frac{\infty}{a} = \infty \quad \text{avec } a \in \mathbb{R}$$

Théorème fondamental :

Théorème 2 *Toute homographie conserve le birapport.*

La conservation du birapport est illustrée par la figure 1.1, et peut s'exprimer par :

$$[A, B, C, D] = [A', B', C', D'].$$

Ce théorème peut se démontrer en utilisant le théorème de Thalès. Semple et Kneebone donnent une autre démonstration utilisant les secteurs angulaires [Sem 52].

Birapport d'un faisceau de quatre droites

Le birapport d'un faisceau de quatre droites l_1, l_2, l_3, l_4 concourantes en O est défini par le birapport $[A, B, C, D]$ des points d'intersection du faisceau avec une droite quelconque l , qui ne contient pas O .

On le note aussi $[O; A, B, C, D]$ (birapport des droites OA, OB, OC, OD).

Le birapport de quatre droites peut s'exprimer sous différentes formes. On peut le calculer en fonction des coordonnées homogènes des points O, A, B, C, D sous la forme proposée par Möbius [Möb 85], lorsque les points A, B, C, D ne sont pas alignés :

$$k = \frac{|OAC| |OBD|}{|OAD| |OBC|} \quad (1.2)$$

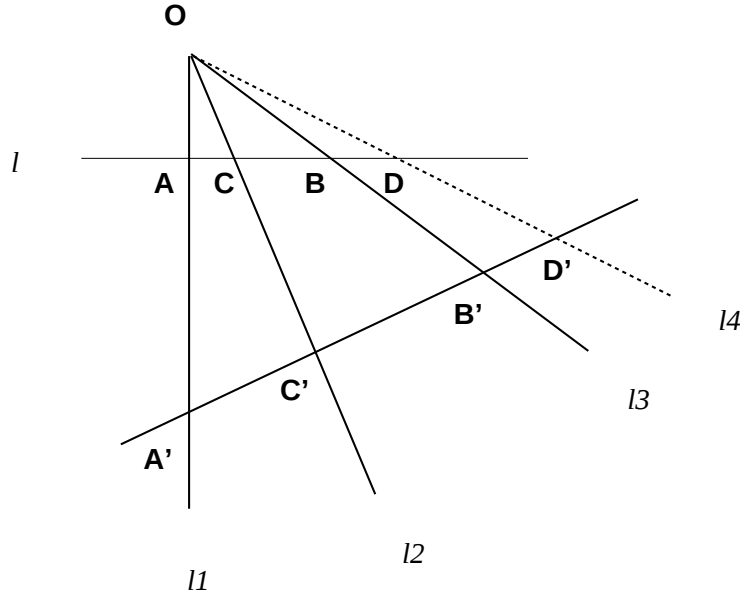


FIG. 1.1 - Birapport de quatre droites

où

$$|P_1 P_2 P_3| = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$$

et (x_i, y_i, z_i) sont les coordonnées homogènes de P_i .

De cette expression découle immédiatement l'expression du birapport en fonction des angles entre les droites du faisceau :

$$k = \frac{\sin(\vec{OA}, \vec{OC}) \sin(\vec{OB}, \vec{OD})}{\sin(\vec{OA}, \vec{OD}) \sin(\vec{OB}, \vec{OC})} \quad (1.3)$$

Birapport d'un faisceau de quatre plans

Le birapport d'un faisceau de quatre plans p_1, p_2, p_3, p_4 ayant une droite commune est défini comme le birapport $[l_1, l_2, l_3, l_4]$ de leurs quatre droites d'intersection avec un plan quelconque p . Ceci est bien sûr indépendant du choix de p . (voir figure 1.2).

La notion de birapport, ainsi que la formule de Möbius se généralisent dans $\mathbb{P}^n$ pour tout faisceau de quatre hyperplans concourants en un espace de dimension $n - 1$ [Bri 83].

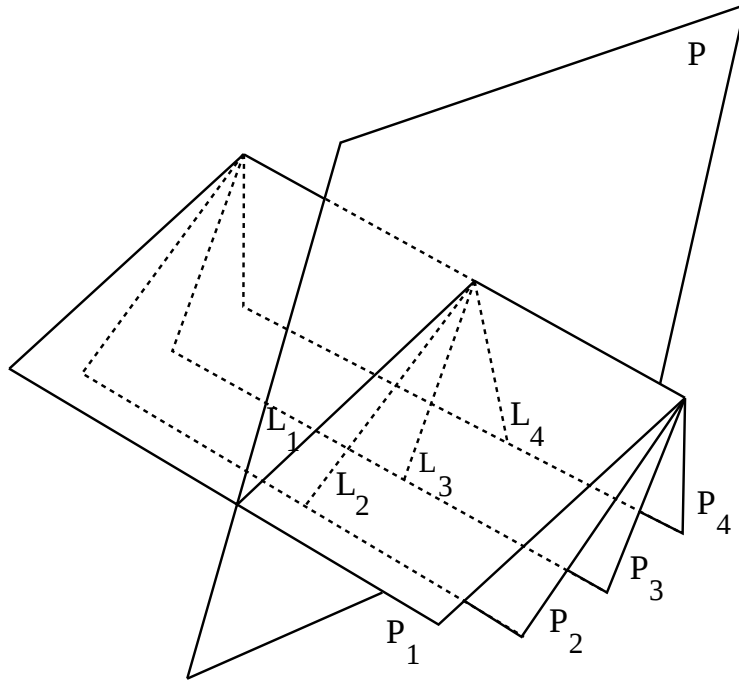


FIG. 1.2 - Birapport de quatre plans

1.4.3 Invariants projectifs des transformations de $\mathbb{P}^3$ dans $\mathbb{P}^2$

Les transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$ nous intéressent plus particulièrement puisque ce sont celles que nous utiliserons pour modéliser la projection effectuée par une caméra. On peut alors se demander quels sont les invariants projectifs associés à ces transformations. Burns [Bur 90] a démontré le résultat suivant :

Théorème 3 *Il n'existe aucun invariant général des transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$.*

Plus précisément, soient $P = (p_1, \dots, p_n)$ et $Q = (q_1, \dots, q_n)$ deux ensembles de n points de $\mathbb{P}^3$. Alors on peut construire une suite de k n -uplets $M_1 = (m_{11}, \dots, m_{1n})$, $\dots$, $M_k = (m_{k1}, \dots, m_{kn})$ telle que :

- Les orbites de P et M_1 s'intersectent.
- Pour tout i , les orbites de M_i et M_{i+1} s'intersectent.
- Les orbites de M_k et Q s'intersectent.

Dire que les orbites de deux n -uplets de $\mathbb{P}^3$ s'intersectent signifie qu'il existe deux transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$ projetant respectivement ces deux n -uplets sur le même n -uplet de $\mathbb{P}^2$.

Un invariant, s'il existe, doit être constant sur chacune des orbites ; il prend donc la même valeur sur des orbites qui s'intersectent. Ceci implique qu'un invariant des transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$, applicable aux n -uplets de $\mathbb{P}^3$ est nécessairement constant.

En revanche, si on se restreint à certains n -uplets de $\mathbb{P}^3$, les orbites associées à ces n -uplets peuvent être disjointes et il devient alors possible de trouver des invariants pour ces n -uplets. Par exemple, les n -uplets de points coplanaires présentent des invariants pour les transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$: en effet, il subissent une homographie de $\mathbb{P}^2$ et on peut utiliser les invariants vus précédemment, en particulier les birapports de faisceaux plans.

En conclusion, les invariants projectifs utilisables dans le cadre des transformations de $\mathbb{P}^3$ dans $\mathbb{P}^2$ se limitent à des ensembles de points qui ne sont pas en position générale, en particulier les ensembles de points alignés ou coplanaires. Une autre solution, est, comme le propose Hartley [Har 92b] de construire des invariants à partir de plusieurs images d'un même ensemble de points.

1.4.4 Coordonnées projectives

Etant donné P un point de $\mathbb{P}^n$, et B une base projective $\mathbb{P}^n$, il est possible de caractériser P par rapport à B par des quantités définies géométriquement à partir d'invariants projectifs : c'est ce que nous appellerons les *coordonnées projectives* de P par rapport à la base projective B . Il découle du paragraphe précédent que ces "coordonnées" sont invariantes par toute homographie.

On donnera une définition des coordonnées projectives uniquement dans les cas de la droite projective $\mathbb{P}^1$ et du plan projectif $\mathbb{P}^2$, mais cette notion est facilement généralisable dans $\mathbb{P}^n$.

Coordonnées projectives sur la droite

Soit une droite l et 3 points de cette droite A, B et C formant une base projective de l . Il suffit pour cela que A, B et C soient distincts. Soit P un point de l . P et A, B, C définissent un birapport, éventuellement infini. $\lambda = [A, B, C, P]$. Inversement, étant donné un scalaire λ , prenant ses valeurs dans $\mathbb{R} \cup \{\infty\}$, il existe un unique point P tel que le birapport $[A, B, C, P]$ soit égal à λ .

Tout couple de réels (x_1, x_2) tel que $x_1/x_2 = \lambda$ est appelé *coordonnées projectives* du point P dans la base projective définie par (A, B, C) .

En fait, le birapport λ suffit à caractériser P . Il permet de déterminer la position d'un point P dans la base projective (A, B, C) en utilisant la relation suivante :

$$\overline{AP} = \frac{\overline{AC} \times \overline{BA}}{\lambda \overline{BC} - \overline{AC}}$$

Dans le cas où $\lambda = \infty$, P est confondu avec A

Coordonnées projectives dans le plan

Dans le plan projectif $\mathbb{P}^2$, quatre points quelconques A, B, C, D non alignés trois à trois définissent une base projective. (voir figure 1.3). Etant donné un point P de $\mathbb{P}^2$, trois réels (x_1, x_2, x_3) tels que

$$x_1 = x_3[A; B, C, D, P] = x_3[AB, AC, AD, AP]$$

$$x_2 = x_3[B; A, C, D, P] = x_3[BA, BC, BD, BP]$$

sont les coordonnées projectives de P dans la base (A, B, C, D) .

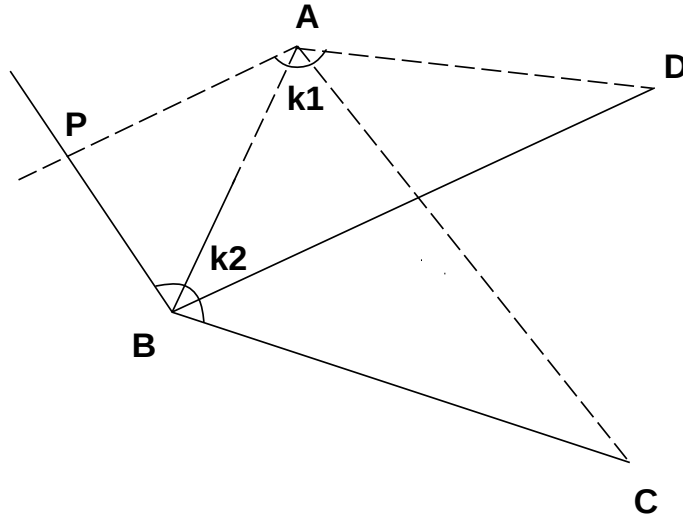


FIG. 1.3 - Coordonnées projectives dans un plan

Tout point P de $\mathbb{P}^2$ peut être référencé de manière unique par ses coordonnées projectives, mis à part les points de la droite (AB) qui ont tous les mêmes coordonnées projectives. Les coordonnées projectives sont définies à un facteur multiplicatif près et ne dépendent pas du système de référence des coordonnées homogènes dans lequel sont décrits les points.

En fait, les birapports $k_1 = x_1/x_3$ et $k_2 = x_2/x_3$ suffisent pour déterminer un point quelconque en dehors de la droite (AB) et ils seront donc utilisés par la suite à la place des coordonnées projectives.

Lien entre coordonnées projectives et coordonnées homogènes

Nous l'avons vu dans le paragraphe précédent, les coordonnées projectives se définissent par rapport à une base projective, et à un facteur de proportionnalité près. Les coordonnées homogènes présentent les mêmes propriétés. Quel est le lien entre ces deux caractérisations?

On se donne un point P du plan projectif, défini par ses coordonnées projectives (x_1, x_2, x_3) dans une base projective (A, B, C, D) . Soient (x, y, z) les coordonnées

homogènes de P dans la représentation du plan projectif où A, B, C, D ont les coordonnées suivantes :

$$\begin{aligned} A &= [0, 1, 0] \\ B &= [1, 0, 0] \\ C &= [0, 0, 1] \\ D &= [1, 1, 1] \end{aligned}$$

On cherche à exprimer les coordonnées projectives en fonction des coordonnées homogènes. Les coordonnées (x_1, x_2, x_3) sont déterminées par la valeur des deux birapports :

$$k_1 = [A; B, C, D, P] = \frac{|ABD| |ACP|}{|ABP| |ACD|}$$

$$k_2 = [B; A, C, D, P] = \frac{|BAD| |BCP|}{|BAP| |BCD|}$$

et

$$\frac{x_1}{x_3} = k_1$$

$$\frac{x_2}{x_3} = k_2$$

Les coordonnées projectives sont donc, à un facteur de proportionnalité près :

$$x_1 = \frac{|ACP|}{|ACD|} = \frac{\begin{vmatrix} 0 & 0 & x \\ 1 & 0 & y \\ 0 & 1 & z \end{vmatrix}}{\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix}} = x$$

$$x_2 = \frac{|BCP|}{|BCD|} = \frac{\begin{vmatrix} 1 & 0 & x \\ 0 & 0 & y \\ 0 & 1 & z \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix}} = y$$

$$x_3 = \frac{|ABP|}{|ABD|} = \frac{\begin{vmatrix} 0 & 1 & x \\ 1 & 0 & y \\ 0 & 0 & z \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{vmatrix}} = z$$

Ceci signifie que les coordonnées projectives d'un point P dans la base projective (A, B, C, D) sont exactement les coordonnées homogènes de ce point dans la représentation où A, B, C et D ont pour coordonnées respectives : $(0, 1, 0)$, $(1, 0, 0)$, $(0, 0, 1)$ et $(1, 1, 1)$.

Les coordonnées projectives des points de la base projective ne sont pas définies, si on utilise la construction géométrique. Mais à la suite de la remarque précédente, on peut définir par extension les coordonnées projectives des points de la base projectives, comme étant $(0, 1, 0)$, $(1, 0, 0)$, $(0, 0, 1)$ et $(1, 1, 1)$.

Cette formulation fournit aussi une méthode pratique pour calculer les coordonnées projectives associées à un point, étant données les coordonnées homogènes des cinq points A, B, C, D, P dans une représentation quelconque : les coordonnées projectives de P seront alors :

$$x_1 = \frac{|ACP|}{|ACD|} \quad x_2 = \frac{|BCP|}{|BCD|} \quad x_3 = \frac{|ABP|}{|ABD|}$$

Inversement, comment calculer les coordonnées homogènes d'un point connaissant ses coordonnées projectives ? En reprenant les expressions ci-dessus, on obtient trois équations linéaires en les coordonnées homogènes de P . La résolution de ce système donne les coordonnées homogènes de P en fonction de ses coordonnées projectives et des coordonnées homogènes de A, B, C, D .

1.5 Propriétés utiles

Dans ce paragraphe, nous énonçons quelques unes des propriétés fondamentales de la géométrie projective qui nous seront utiles par la suite : nous verrons le théorème de Chasles, le principe de dualité et le lien entre géométrie projective et géométries affine et euclidienne.

1.5.1 Conique de Chasles

Les coniques en géométrie projective

Les coniques sont des objets projectifs : *un changement de repère projectif transforme une conique en une conique.*

démonstration

L'équation d'une conique est une équation homogène de degré 2 en fonction des coordonnées homogènes (x_1, x_2, x_3) d'un point P :

$$\sum_{i=1}^3 \sum_{j=1}^3 a_{ij} x_i x_j = 0$$

On peut supposer sans faire de restriction que $a_{ij} = a_{ji}$. L'équation d'une conique peut alors s'écrire sous une forme matricielle :

$$X^t A X = 0 \quad \text{où } X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{et } A \text{ est une matrice } 3 \times 3 \text{ symétrique}$$

Soit une conique $\mathcal{C}$ associée à une matrice symétrique A , d'équation :

$$X^t A X = 0$$

On effectue un changement de repère associé à une matrice P inversible. Les nouvelles coordonnées X' et anciennes coordonnées X d'un point sont liées par la relation :

$$X = P X'$$

d'où

$$X^t = X'^t P^t$$

En remplaçant X et X^t dans l'équation de $\mathcal{C}$, on obtient l'équation de $\mathcal{C}$ dans le nouveau repère :

$$X'^t P^t A P X' = 0$$

On pose $A' = P^t A P$. Le lieu des points vérifiant l'équation $X'^t A' X' = 0$ est une conique. $\square$

Une conique est transformée en une autre conique par un changement de base projective. Cependant, en géométrie projective, il n'existe aucune distinction entre les différents types de coniques : ellipse, parabole et hyperbole n'ont de sens qu'en géométrie affine, c'est à dire si on a défini la droite de l'infini. Dans ce cas, un changement de repère projectif peut alors transformer une ellipse en une hyperbole et inversement. Il suffit pour cela que le changement de base transforme une droite coupant l'ellipse en la droite de l'infini.

Toute homographie transforme donc une conique en une autre conique, éventuellement dégénérée (réduite à une droite ou un segment).

Nombre de degrés de liberté d'une conique

Une conique est entièrement déterminée par la matrice A qui lui est associée. La matrice A est une matrice 3×3 symétrique, définie à un facteur de proportionnalité près. Son degré de liberté est donc 5. Un point de la conique fournit une équation. Cinq points du plan non quatre à quatre alignés définissent donc de manière unique une conique passant par ces points.

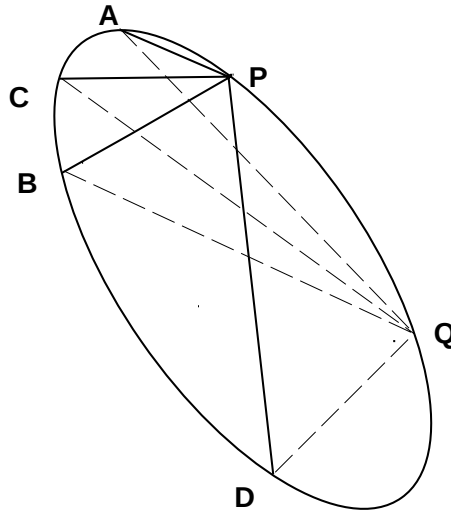


FIG. 1.4 - Conique de Chasles : $[P; A, B, C, D] = [Q; A, B, C, D]$

Théorème de Chasles

Théorème 4 (Chasles) Soient quatre points A, B, C, D formant une base du plan projectif. Soit k un réel. Le lieu des points sommets d'un faisceau de droites passant par A, B, C, D et de birapport k est une conique.

Inversement, si P appartient à une conique contenant A, B, C, D , alors le birapport du faisceau de droites de sommet P et passant par A, B, C, D est indépendant de P . (voir figure 1.4)

L'équation de la conique $\mathcal{C}$ définie par quatre points A, B, C, D et un birapport k s'écrit facilement :

$$\begin{aligned}
 P \in \mathcal{C} &\iff [P; A, B, C, D] = k \\
 &\iff \frac{|PAC| |PBD|}{|PAD| |PBC|} = k \\
 &\iff |PAC| |PBD| - k |PAD| |PBC| = 0
 \end{aligned}$$

Chaque déterminant $|PXY|$ est un polynôme homogène du premier degré en (x, y, z) . Le membre gauche de l'équation est donc un polynôme homogène de degré 2 en (x, y, z) . C'est donc bien l'équation d'une conique. On vérifie facilement que les points A, B, C et D annulent cette équation et appartiennent donc à la conique. Ceci démontre la partie directe du théorème.

La réciproque de ce théorème se démontre très élégamment : il suffit d'observer qu'elle est immédiate dans le cas du cercle : les angles du faisceau issu de P se conservent lorsque P varie sur le cercle. Le birapport du faisceau est donc constant

(voir eq.1.3). Cette propriété ne faisant intervenir que des notions projectives (incidence, birapport), elle se généralise à toute conique, puisque cercle et conique sont équivalents en géométrie projective.

1.5.2 Principe de dualité

A tout hyperplan de $\mathbb{P}^n$ est associé le $n+1$ -uplet $(a_1, \dots, a_{n+1})$ tel que

$$\sum_{i=1}^{n+1} a_i x_i = 0$$

est l'équation de l'hyperplan. Cette équation caractérise l'ensemble des points de coordonnées $(x_1, \dots, x_{n+1})$ appartenant à l'hyperplan. Inversement si on fixe un point $(x_1, \dots, x_{n+1})$, cette équation représente l'ensemble des hyperplans $(a_1, \dots, a_{n+1})$ contenant ce point.

L'équation $\sum_{i=1}^{n+1} a_i x_i = 0$ caractérise en fait la relation d'incidence entre un point et un hyperplan, et ce de manière symétrique pour les points et les hyperplans. L'ensemble des points $(x_1, \dots, x_{n+1})$ de $\mathbb{P}^n$ et l'ensemble des hyperplans $(a_1, \dots, a_{n+1})$ de $\mathbb{P}^n$ sont équivalents. C'est le principe de dualité : toute propriété des points est transposable en une propriété des hyperplans.

Le principe de dualité n'apporte rien sur la droite projective $\mathbb{P}^1$, puisque les points eux-mêmes sont les hyperplans de $\mathbb{P}^1$.

Par contre, les points et les droites sont duaux dans le plan projectif $\mathbb{P}^2$. Cette dualité s'exprime essentiellement par l'équivalence entre la notion d'alignement pour les points et la notion de concourance pour les droites : deux points définissent une droite et deux droites définissent un point. Un ensemble de points alignés est dual d'un faisceau de droites concourantes.

On peut donner un exemple de propriété immédiatement déduite du principe de dualité. On sait que 5 points P_1, P_2, P_3, P_4, P_5 , définissent un birapport, celui du faisceau de droites de sommet P_1 et passant par les points P_2, P_3, P_4, P_5 . On peut donc en déduire que 5 droites l_1, l_2, l_3, l_4, l_5 , définissent un birapport, celui des points appartenant à la droite l_1 , et intersections de l_1 avec l_2, l_3, l_4, l_5 . (voir figure 1.5)

Toujours grâce au principe de dualité, on en déduit directement que le birapport ainsi défini a pour valeur :

$$k = \frac{|l_1 l_2 l_4| |l_1 l_3 l_5|}{|l_1 l_2 l_5| |l_1 l_3 l_4|}$$

où $|l_i l_j l_k|$ est le déterminant de la matrice 3×3 contenant les coordonnées homogènes des droites l_i, l_j, l_k .

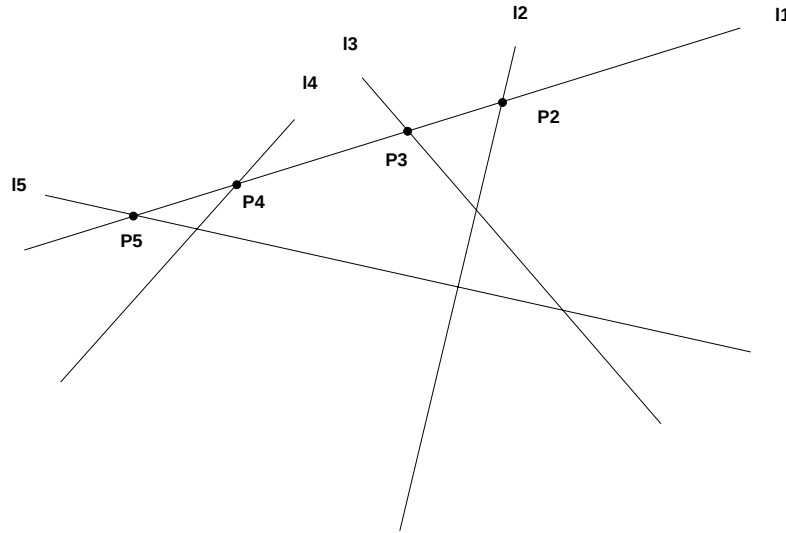


FIG. 1.5 - Birapport associé à cinq droites : $[l_1; l_2, l_3, l_4, l_5] = [P_2, P_3, P_4, P_5]$

1.5.3 Géométries affine et euclidienne

La géométrie projective englobe les géométries affine et euclidienne : ces dernières sont des restrictions de la géométrie projective. De ce fait, les transformations affines et les transformations euclidiennes sont des sous-groupes des homographies. Inversement, les invariants projectifs sont aussi des invariants affines et euclidiens.

Géométrie affine

La géométrie affine se déduit de la géométrie projective en fixant l'hyperplan de l'infini.

Transformations affines

Les transformations affines sont donc les transformations projectives qui conservent globalement l'hyperplan de l'infini. On peut dire de manière équivalente qu'une transformation est affine si et seulement si elle conserve le parallélisme. En effet des droites sont parallèles si et seulement si elles s'intersectent en un point du plan de l'infini. Dans le cas de $\mathbb{P}^n$, l'hyperplan de l'infini peut donc être déterminé par la donnée de n paires indépendantes de droites parallèles.

Si on fixe comme hyperplan de l'infini le plan $x_{n+1} = 0$, alors la matrice associée à toute transformation affine a la forme suivante :

$$\begin{bmatrix} & t_1 \\ A & \vdots \\ & t_n \\ 0 \dots 0 & 1 \end{bmatrix}$$

A est une matrice $n \times n$ correspondant à la partie linéaire de la transformation affine et $(t_1, \dots, t_n)$ est le vecteur de translation.

Invariants affines

Le parallélisme, qui n'est pas une notion projective, est un invariant affine fondamental. A partir du parallélisme de droites, on peut définir la notion de plans parallèles (sécants en une droite située dans le plan de l'infini), et droites parallèles (sécantes en un point du plan de l'infini).

L'invariant algébrique fondamental en géométrie affine est le rapport des longueurs de 3 points alignés. On peut noter que cet invariant est en fait un birapport faisant intervenir un point à l'infini. Si A, B et C sont 3 points d'une droite l , et si D est le point à l'infini sur l , alors, en utilisant les conventions vues au paragraphe 1.4.2, on a

$$[A, B, C, D] = \frac{\overline{AC} \times \overline{BD}}{\overline{AD} \times \overline{BC}} = \frac{\overline{AC}}{\overline{BC}}$$

Géométrie Euclidienne

La géométrie euclidienne est une restriction supplémentaire de la géométrie affine : l'hyperplan de l'infini est globalement fixé, ainsi que certains points de cet hyperplan.

Transformations euclidiennes du plan

Les transformations euclidiennes du plan sont les transformations affines qui conservent une paire de points de la droite de l'infini. Ces deux points I, J sont appelés *points absolus* et ils peuvent être choisis arbitrairement sur la droite de l'infini. Une transformation euclidienne qui conserve le couple (I, J) , est une transformation directe. Inversement, si le couple (I, J) a pour image le couple (J, I) , la transformation est dite indirecte.

Etant données deux droites l_1, l_2 et A_1, A_2 leurs intersections respectives avec la droite de l'infini (voir figure 1.6), le birapport

$$k_{12} = [A_1, A_2, I, J]$$

est invariant par toute transformation euclidienne directe, et il est transformé en son inverse par toute transformation euclidienne indirecte.

Si on a choisi $x_3 = 0$ comme droite de l'infini et si on choisit comme points absolus les *points cycliques* $(1, i, 0)$ et $(1, -i, 0)$, les transformations euclidiennes sont les similitudes planes : rotations translations, homothéties et symétries orthogonales. On définit alors les cercles comme les ellipses intersectant la droite de l'infini en les points cycliques. L'angle entre deux droites l_1 et l_2 est défini par la formule de Laguerre :

$$\alpha = (\vec{l}_1, \vec{l}_2) = \frac{1}{2i} \log[A_1, A_2, I, J]$$

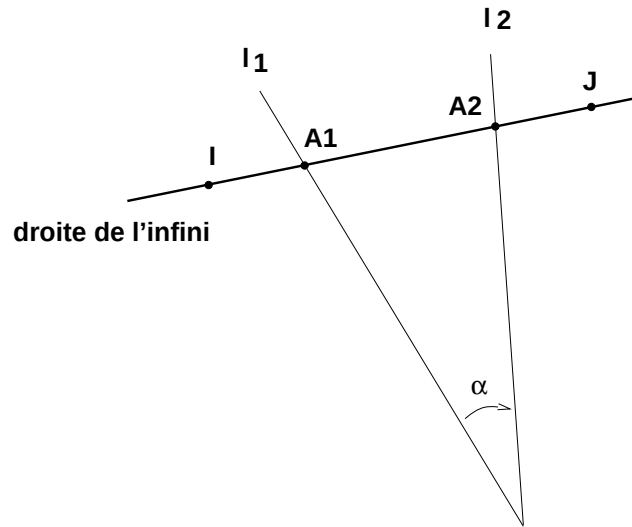


FIG. 1.6 - Les points cycliques et le lien avec la notion d'angle en géométrie euclidienne.

Dans cette même représentation, la matrice associée à une transformation euclidienne directe prend la forme suivante [Sem 52] :

$$\begin{bmatrix} a \cos \alpha & a \sin \alpha & t_1 \\ -a \sin \alpha & a \cos \alpha & t_2 \\ 0 & 0 & 1 \end{bmatrix}$$

et une transformation euclidienne indirecte a la forme :

$$\begin{bmatrix} a \cos \alpha & a \sin \alpha & t_1 \\ -a \sin \alpha & -a \cos \alpha & t_2 \\ 0 & 0 & 1 \end{bmatrix}$$

On retrouve les matrices associées aux similitudes du plan.

Transformations euclidiennes de l'espace

On obtient la géométrie euclidienne de l'espace à partir de la géométrie affine de l'espace en fixant une conique du plan de l'infini. Cette conique ne contient que des points imaginaires et elle est appelée *conique absolue* ; on la note Ω . Les transformations euclidiennes de l'espace sont les transformations affines qui laissent globalement invariante la conique absolue. Les intersections d'un plan avec Ω définissent les points absolus associés à ce plan.

Si on choisit comme plan de l'infini le plan $x_4 = 0$ et si on choisit dans ce plan la conique d'équation $x_1^2 + x_2^2 + x_3^2 = 0$ comme conique absolue, alors on retrouve ainsi qu'en géométrie euclidienne du plan, la notion d'angle. L'angle entre deux

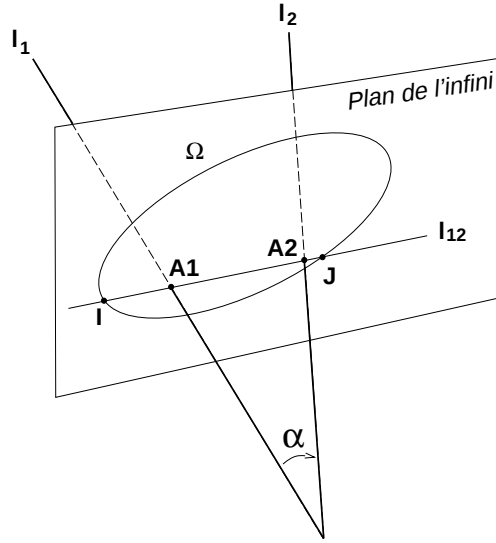


FIG. 1.7 - La conique absolue et la notion d'angle dans l'espace euclidien

droites sécantes l_1 et l_2 est défini par :

$$\alpha = (\vec{l}_1, \vec{l}_2) = \left| \frac{1}{2i} \log[A_1, A_2, I, J] \right|$$

où I et J sont les intersections du plan contenant l_1 et l_2 avec Ω (points cycliques associés à ce plan) et A_1 et A_2 sont les intersections de l_1 et l_2 avec le plan de l'infini. Ces quatre points sont alignés car ils appartiennent à la droite l_{12} , intersection du plan contenant l_1 et l_2 et du plan de l'infini. Ils définissent donc bien un birapport (voir figure 1.7).

Dans la pratique, la position de Ω peut être déterminée en utilisant des connaissances a priori sur les propriétés euclidiennes de l'espace [Bou 93]. Une première solution consiste à utiliser le fait que tous les cercles de l'espace intersectent Ω . Si on connaît trois cercles non deux à deux coplanaires, chacun intersecte le plan de l'infini en deux points qui appartiennent à Ω . Ces six points déterminent l'équation de Ω . En effectuant le changement de repère affine qui ramène l'équation de Ω à $x_1 + x_2 + x_3 = 0$, on obtient une représentation euclidienne de l'espace.

Une deuxième solution utilise l'invariant euclidien : l'angle. Si on connaît l'angle entre deux droites concourantes, et en reprenant les notations ci-dessus, on peut écrire : $I = A + tB$ et $J = A + t'B$ et $[A, B, C, D] = \frac{t}{t'} = 2i\alpha$. En exprimant que I et J appartiennent à Ω , on obtient des contraintes polynômiales sur les coefficients de l'équation de Ω .

La conique absolue est un objet important en vision : en effet connaître la projection de la conique absolue sur le plan image est équivalent à l'étalonnage des paramètres intrinsèques de la caméra [May 92].

Chapitre 2

Application au positionnement relatif

2.1 Introduction

Nous nous plaçons ici dans la problématique de la perception tridimensionnelle en vision par ordinateur : quelle information tridimensionnelle peut être extraite d'une ou plusieurs images ? En particulier, on cherche à déterminer la position ou la forme des objets observés, ainsi que la ou les positions de la caméra dans l'espace.

Plusieurs questions se posent alors :

- Quel modèle de projection faut-il utiliser ?
- Quel type d'information est recherché : en particulier, dans quel repère veut-on exprimer les points calculés ?
- Quelles informations a priori sont alors nécessaires ?

Nous allons dans la suite présenter les différentes solutions généralement apportées à ce problème et situer notre approche par rapport à ces solutions.

2.1.1 Modèles de projection

Plusieurs modèles permettent de représenter la projection effectuée par une caméra. Le plus courant est le modèle sténopé, dans lequel une caméra est définie par son centre de projection et son plan de projection. A chaque point P de l'espace 3D est associé la droite passant par P et le centre de projection, appelée *ligne de vue*. L'image p de P est alors l'intersection de la ligne de vue et du plan de projection (voir figure 2.1).

Ce modèle ne tient pas compte de la nécessité d'une mise au point : en toute rigueur, seuls les points du plan de netteté sont projetés en un point du plan image, et le phénomène de flou intervient pour tous les autres points. Le modèle des lentilles minces doit être utilisé pour tenir compte de cet aspect.

De plus, les objectifs des caméras réelles sont généralement constitués de plusieurs lentilles, et les rayons lumineux qui les traversent ne forment pas un faisceau de droites convergentes. Dans les cas où la distance focale doit être connue avec une grande précision, il est nécessaire de considérer le modèle des systèmes épais, pour lequel une caméra est caractérisée par ses deux plans principaux et ses deux points principaux. C'est le modèle employé dans les travaux de Lavest et al. [Lav 91] qui utilisent un objectif à focale variable (zoom) pour le positionnement et l'étalonnage.

Les photogramètres introduisent d'autres raffinements destinés à modéliser la non linéarité des systèmes optiques, en particulier les distorsions radiales.

Dans le domaine de la vision par ordinateur, la précision du modèle sténopé est généralement suffisante. De plus, la transformation effectuée est alors une projection perspective pure, qui appartient aux transformations projectives de $\mathbb{P}^3$ dans $\mathbb{P}^2$. Cette transformation s'exprime donc sous la forme très simple d'une transformation linéaire en coordonnées homogènes. Cependant, Peuchot [Peu 92b] [Peu 92c] [Peu 92a] a observé que le modèle sténopé introduit un biais de l'ordre du pixel et pouvant aller jusqu'à 4 ou 5 pixels sur les bords de l'image. Il propose une technique permettant de transformer l'image issue d'une caméra réelle en une projection perspective pure, et ceci sans connaître les caractéristiques optiques de la caméra. Elle a été reprise dans notre équipe par P. Brand [Bra 92], qui a obtenu une précision du dixième de pixel sur l'image corrigée.

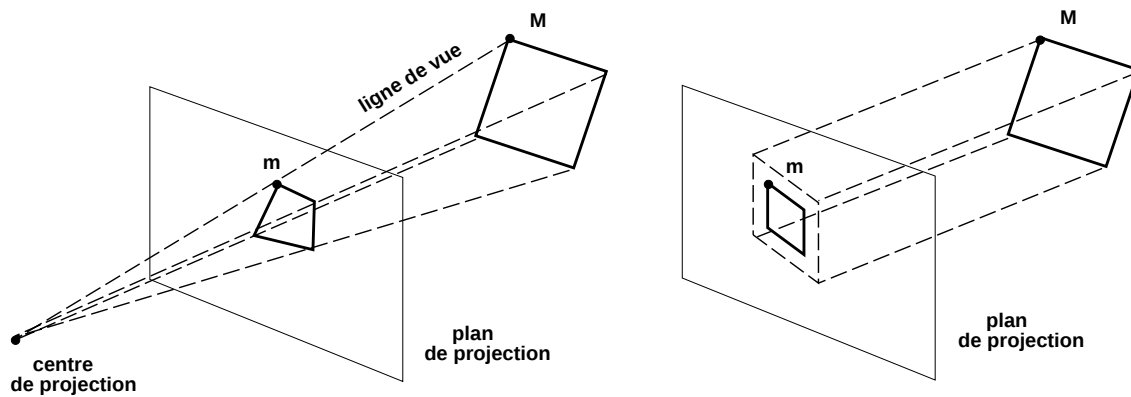


FIG. 2.1 - Les modèles de projection couramment utilisés : projection perspective pure et projection perspective faible

Un autre modèle souvent utilisé est celui de la projection perspective faible, résultat d'une projection orthographique composée avec une homothétie (voir la figure 2.1). Ce modèle est valide pour les objets dont la profondeur est réduite par rapport à la distance entre l'objet et l'image : Thompson et Mundy donnent un rapport de 1/10 entre profondeur et distance comme limite de validité [Tho 87]. Ce modèle est donc essentiellement utilisable dans le cas d'objets plans positionnés parallèlement au plan image.

2.1.2 Rétroprojection

Etant donné un modèle de projection, la rétroprojection consiste à déterminer l'ensemble des points de l'espace susceptibles de se projeter en un point donné de l'image. Pour des modèles de projection pure (perspective ou orthographique), cet ensemble est une droite. Dans le cas de la projection perspective, c'est la ligne de vue.

Pour un point du plan image donné, il y a donc une infinité de points se projetant en ce point. C'est pourquoi on doit disposer d'au moins deux images du même point pour pouvoir déterminer sa position tridimensionnelle : c'est le principe de la *stéréovision*. Une fois déterminées les lignes de vues associées à un même point, le point est reconstruit par triangulation, c'est-à-dire par intersection des lignes de vue. Il se pose donc le problème de la mise en correspondance des points entre les images (en anglais "matching") : c'est un préalable indispensable à la reconstruction et un des problèmes épineux en vision.

2.1.3 Reconstruction avec étalonnage

La solution classique au problème de la reconstruction est la suivante : les points de la scène sont reconstruits dans un repère orthonormé fixé a priori et généralement

appelé *repère scène*. Le modèle utilisé est celui de la projection perspective, et la rétroprojection est réalisée grâce à l'étalonnage préalable des caméras.

L'étalonnage d'une caméra consiste à déterminer la position de son centre de projection et de son plan de projection dans le repère scène, à l'aide d'une mire dont la position est parfaitement connue dans ce repère [Tsa 87]. On définit pour cela un repère orthonormé lié à la caméra dont l'origine est située au centre optique, et dont l'axe $\vec{Z}$ (axe optique) est orthogonal au plan image.

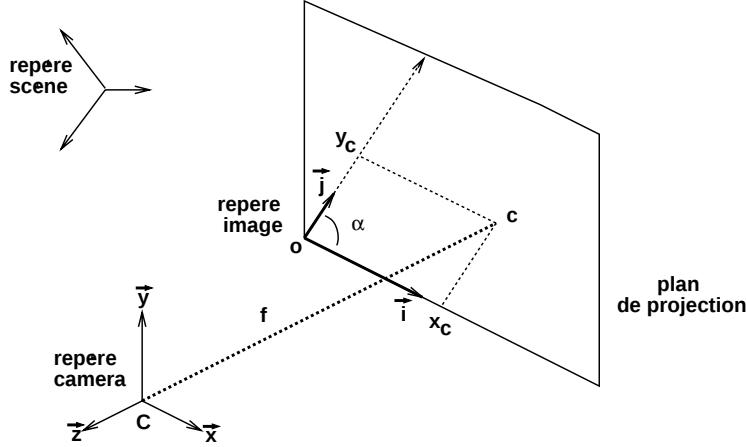


FIG. 2.2 - Les paramètres de l'étalonnage

L'étalonnage correspond à la détermination de 11 paramètres.

5 paramètres appelés *paramètres intrinsèques* ou *paramètres intérieurs* définissent la position du repère image par rapport au repère de la caméra. Le repère image $(o, \vec{i}, \vec{j})$ est le repère dans lequel sont effectuées les mesures, c'est à dire le repère dans lequel on exprime la position d'un point de l'image en pixels. Ce repère dépend du capteur et il n'est généralement pas orthonormé : les pixels peuvent être des rectangles ou mêmes des parallélogrammes. Le repère caméra est un repère orthonormé direct d'origine le centre de projection de la caméra, dont l'axe $\vec{x}$ est parallèle à l'axe $\vec{i}$ du repère image et dont l'axe $\vec{z}$ (axe principal parfois aussi appelé axe optique) est orthogonal au plan de projection de la caméra. Les paramètres nécessaires pour caractériser la position relative de ces deux repères sont les suivants (voir figure 2.2) :

- les coordonnées x_c et y_c de c , intersection de l'axe optique avec le plan de projection. x_c et y_c sont exprimés dans le repère image,
- les deux unités i et j sur les axes du plan image (pour une distance focale f arbitrairement fixée),
- et l'angle α entre ces axes (s'ils ne sont pas orthogonaux).

6 paramètres appelés *paramètres extrinsèques* ou *paramètres extérieurs* selon

la terminologie des photogramètres, définissent la position du repère associé à la caméra dans un repère scène :

- 3 paramètres déterminent la rotation entre les axes des deux repères
- 3 paramètres déterminent la translation entre les origines des deux repères.

L'étalonnage d'une caméra détermine ces paramètres et il permet aussi le calcul de la position dans le repère scène de la ligne de vue associée à tout point du plan image. L'étape de mise en correspondance est ensuite réalisée entre les images (elle est facilitée par l'étalonnage qui engendre des contraintes géométriques) et les points sont enfin reconstruits par triangulation.

Cependant l'étalonnage s'est révélé une étape délicate et numériquement instable (voir [Fau 87] [Aya 87], [Bal 82]) : il fait partie de la classe des problèmes mal conditionnés. De plus, les paramètres d'étalonnage sont très sensibles aux modifications de l'environnement tels qu'un changement de la température ou un choc mécanique. Plus simplement, une mise au point correspond à un changement de focale et rend donc nécessaire un nouvel étalonnage des paramètres intrinsèques.

L'étalonnage est aussi parfois inutilisable, comme dans le cas de la vision active : la caméra se déplace autour de l'objet observé et la géométrie de cet objet est obtenue par la fusion des informations extraites des différents points de vue. Dans ce contexte, l'utilisation de l'étalonnage pour la détermination des lignes de vue est impossible : les paramètres intrinsèques peuvent être déterminés a priori, mais ils subissent une dérive temporelle et ils sont entièrement modifiés par chaque mise au point. Les paramètres extrinsèques, qui déterminent la position du repère caméra dans un repère absolu, sont différents pour chaque point de vue. Les informations de position issues des capteurs odométriques comportent une trop grande imprécision pour pouvoir être utilisés pour la détermination des lignes de vue. Il faudrait alors un étalonnage préalable à chaque point de vue, ce qui est impossible à réaliser.

2.1.4 Positionnement relatif

Les difficultés liées à l'étalonnage ont conduit à rechercher des méthodes qui permettent d'éviter un étalonnage explicite des caméras. Des solutions ont été proposées qui ne supposent connus que les paramètres intrinsèques de la caméra : le problème de la reconstruction se simplifie alors beaucoup (voir par exemple [Hor 89], [Yua 89], [Liu 90] pour des contributions récentes).

La notion d'*autocalibrage* a ensuite été introduite : l'idée est d'utiliser les points observés eux-mêmes pour constituer le repère dans lequel on positionne la scène. L'étalonnage est alors effectué simultanément à la prise d'image, et on évite ainsi les problèmes liés à la dérive des paramètres.

Le deuxième principe novateur est l'utilisation de propriétés géométriques liées au modèle de projection employé. L'utilisation de la géométrie permet d'éviter un étalonnage explicite et les problèmes numériques qui y sont liés.

Des méthodes ont été proposées qui utilisent le modèle de la projection orthographique ([Koe 89] [Spa 90] [Lee 90]). Dans ce cas, on dispose des propriétés issues de la géométrie affine. Sparr et Nielsen en particulier introduisent grâce à une légère généralisation de la notion de coordonnées barycentriques, la notion de “forme” affine, qui est conservée par toute transformation affine. Tomasi et Kanade [Tom 91], proposent une méthode de reconstruction à partir d’une séquence d’images dont les points sont mis en correspondance.

De façon similaire, le modèle de projection perspective permet d’utiliser les propriétés de la géométrie projective. Comme le mentionnent Duda et Hart [Dud 73], le lien entre la géométrie projective, et les problèmes d’étalonnage, de rétroprojection et de reconstruction était jusqu’à présent mal compris et donc peu utilisé. Dans ce cadre, le pionnier est Longuet-Higgins [LH 81] qui le premier a formulé les contraintes géométriques liées à la reconstruction avec un modèle de projection perspective. Cependant, sa technique suppose connus certains paramètres de la caméra (la focale et l’intersection de l’axe optique avec le plan de l’image). Larabi [Lar 91b] [Lar 91a] utilise les propriétés de la géométrie projective et un système de vision stéréo partiellement étalonné pour réaliser l’appariement entre les deux images et déterminer la profondeur des points de l’image.

L’approche que nous proposons s’inscrit dans cette démarche générale et s’appuie sur les choix suivants :

- Le modèle de projection employé est celui de la projection perspective : il est relativement fidèle à la réalité et nous permet d’utiliser les propriétés de la géométrie projective. En particulier, l’originalité de notre méthode réside dans le fait qu’elle n’utilise que des constructions géométriques simples issues des propriétés de la géométrie projective.
- Nous cherchons à reconstruire les points dans un repère lié à des points de référence appartenant à l’image : ainsi, si la position des points utilisés comme références est connue, on obtient une reconstruction similaire à celle réalisée avec un étalonnage explicite utilisant ces points. Par contre, si la position des points de référence n’est pas connue, les propriétés géométriques nous permettent néanmoins de reconstruire les autres points *relativement* aux points de référence. Ce positionnement relatif est suffisant pour de nombreuses tâches en vision : par exemple, pour la manipulation, l’information recherchée est la position de la pince du robot par rapport à l’objet à saisir, et non leurs positions dans un repère absolu.
- Enfin, aucune connaissance préalable sur les paramètres des caméras n’est utilisée. Nous ne nous donnons comme connaissance a priori que la mise en correspondance des points entre les différentes images et éventuellement la position tridimensionnelle des points de référence.

Nous allons montrer dans la suite de ce chapitre comment nous pouvons alors résoudre les différents problèmes liés à la reconstruction : la rétroprojection, la

localisation du centre optique de la caméra, la localisation d'un point dans la scène et enfin la détermination de la géométrie épipolaire.

2.2 Reconstruction en utilisant des points coplanaires

Ce paragraphe propose une résolution des problèmes mentionnés dans l'introduction dans le cas plus simple où au moins quatre des points de référence sont coplanaires. Nous résolvons d'abord le problème de la perspective inverse et la technique utilisée est ensuite appliquée à la résolution des autres problèmes.

Cette méthode utilise comme points de référence deux groupes de quatre points coplanaires $\{A, B, C, D\}$ et $\{E, F, G, H\}$ et leurs projections dans le plan image $\{a, b, c, d\}$ et $\{e, f, g, h\}$.

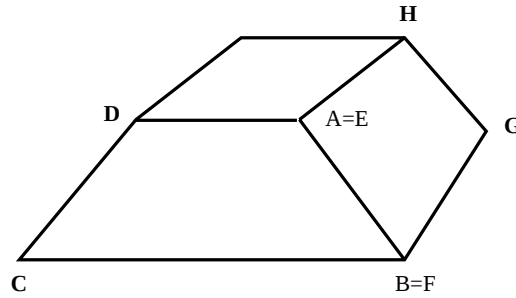


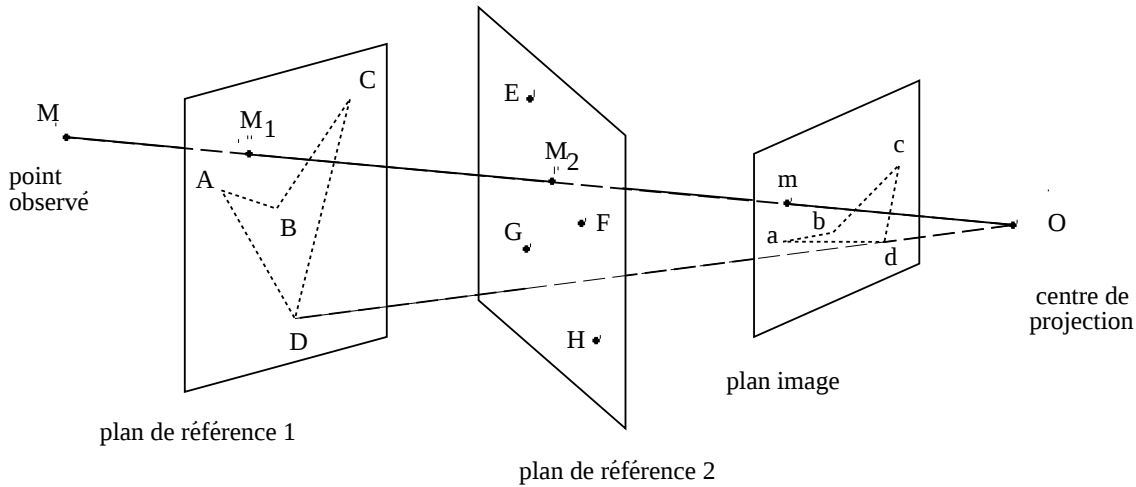
FIG. 2.3 - 6 points sur deux faces adjacentes fournissent 2 groupes de 4 points coplanaires

On peut remarquer que cette configuration peut être obtenue avec $a = e$ et $b = f$; c'est le cas lorsque nous choisissons comme points de référence 6 sommets d'un parallélépipède appartenant à deux faces adjacentes comme c'est le cas sur la figure 2.3.

2.2.1 Comment rétroprojeter un point image?

Etant donné un point observé dans le plan image m , il est possible de déterminer la position dans la scène de la ligne de vue Om passant par m , sans aucun étalonnage.

Démonstration : Comme on l'a vu au paragraphe 1.2.2, les quatre projetés des points de référence, a, b, c et d forment une base projective du plan image. On peut donc déterminer les coordonnées projectives (x_1, x_2, x_3) de m par rapport à ce repère, à partir de mesures dans l'image. Soit M_1 , l'intersection de la ligne de vue Om avec le plan $ABCD$ (voir la figure 2.4). M_1 se projette en m dans le plan image. Comme les coordonnées projectives sont invariantes par projection (voir paragraphe 1.4.4), les coordonnées projectives de M_1 par rapport à (A, B, C, D)

FIG. 2.4 - La rétroprojection du point image m

sont (x_1, x_2, x_3) . Ces coordonnées déterminent la position de M_1 dans le plan de référence $ABCD$.

On peut de même déterminer M_2 , l'intersection de la ligne de vue Om avec le plan $EFGH$. On dispose ainsi de deux points de la ligne de vue M_1 et M_2 qui déterminent sa position par rapport à la scène, et ceci *sans connaître la position de la caméra dans la scène*.

2.2.2 Où est la caméra ?

Nous venons de montrer dans le paragraphe 2.2.1 comment reconstruire à partir d'une seule vue la ligne de vue associée à un point de l'image. Il est aussi possible de reconstruire la position du centre optique, O à partir d'une seule vue, en utilisant la même méthode. Si nous disposons de deux points m et p dans le plan image, il est possible de reconstruire les lignes de vue Om et Op . Or ces deux droites s'intersectent dans l'espace en O , ce qui permet de calculer la position de O .

En pratique, la présence de bruit entraîne généralement que les deux lignes de vues ne s'intersectent pas : la position du centre optique est alors déterminée aux moindres carrés. De plus, en utilisant plus de deux points, l'intersection aux moindres carrés des lignes de vues associées donne une estimation du centre optique O qui est beaucoup plus résistante aux bruits de mesure.

2.2.3 Où est le point image ?

Nous supposons ici que la position du centre de la caméra O a été calculée en utilisant la méthode précédente. On peut alors facilement construire le point m , image d'un point M connu de la scène : il faut d'abord calculer le point d'intersection de OM avec un plan connu contenant quatre points de référence ; on

détermine ensuite les coordonnées projectives du point d'intersection ainsi qu'elles ont été définies au paragraphe 1.4.4. Enfin, on peut reconstruire le point m dans le plan image : c'est le point défini par ces mêmes coordonnées projectives dans la base projective formée par les projetés des quatre points de référence.

2.2.4 Comment localiser un point dans la scène?

La méthode précédente peut être appliquée pour localiser un point dans une scène contenant des points de référence. Pour cela deux prises de vue sont nécessaires. Un point donné M de l'espace 3D est projeté respectivement en m_1 et m_2 dans chacune de ces deux images.

Supposons que la correspondance entre m_1 et m_2 a été établie. Pour chaque image, nous savons reconstruire la ligne de vue passant par M et nous pouvons donc déterminer M comme intersection des deux lignes de vue m_1M et m_2M . Cette méthode dégénère uniquement lorsque les deux positions du centre optique sont alignées avec la ligne de vue.

2.2.5 Comment déterminer les épipolaires?

En stéréovision, tout point M de l'espace 3D est soumis à la contrainte épipolaire : si m et m' sont respectivement les projetés de M dans la première et la deuxième image, le plan épipolaire $OO'M$ intersecte le plan image selon les droites om et $o'm'$. La droite om (resp. $o'm'$) est appelée la droite épipolaire associée au point m' (resp. m). Ainsi, étant donné un point m dans la première image, son correspondant m' dans la deuxième image appartient nécessairement à la droite épipolaire associée à m .

Une procédure identique à celle décrite au paragraphe 2.2.1 permet de déterminer les droites épipolaires. On suppose que l'on observe un point M depuis 2 points de vue O et O' . M se projette respectivement en m et m' dans les 2 images I et I' . On se donne 2 ensembles de 4 points de référence coplanaires $\{A, B, C, D\}$ et $\{E, F, G, H\}$ se projetant respectivement en $\{a, b, c, d\}$ et $\{e, f, g, h\}$ dans l'image I et en $\{a', b', c', d'\}$ et $\{e', f', g', h'\}$ dans l'image I' . On cherche la droite épipolaire associée à m' dans l'image I , c'est-à-dire la projection de la ligne de vue $O'M'$ dans la première image I . Pour cela, il suffit de construire les projetés p_1 et p_2 des 2 points P_1 et P_2 , intersections de $O'M'$ avec les plans de référence. Par conservation des coordonnées projectives, les coordonnées projectives de p_1 par rapport à a, b, c, d sont les mêmes que celles de P_1 par rapport à A, B, C, D et donc les mêmes que celles de m' par rapport à a', b', c', d' , ces dernières étant mesurables dans l'image I' . De même, les coordonnées projectives de p_2 par rapport à e, f, g, h sont celles de m' par rapport à e', f', g', h' . On peut donc reconstruire les points p_1 et p_2 dans l'image I , déterminant ainsi la droite épipolaire associée à m' .

On peut noter que cette construction est uniquement basée sur des mesures

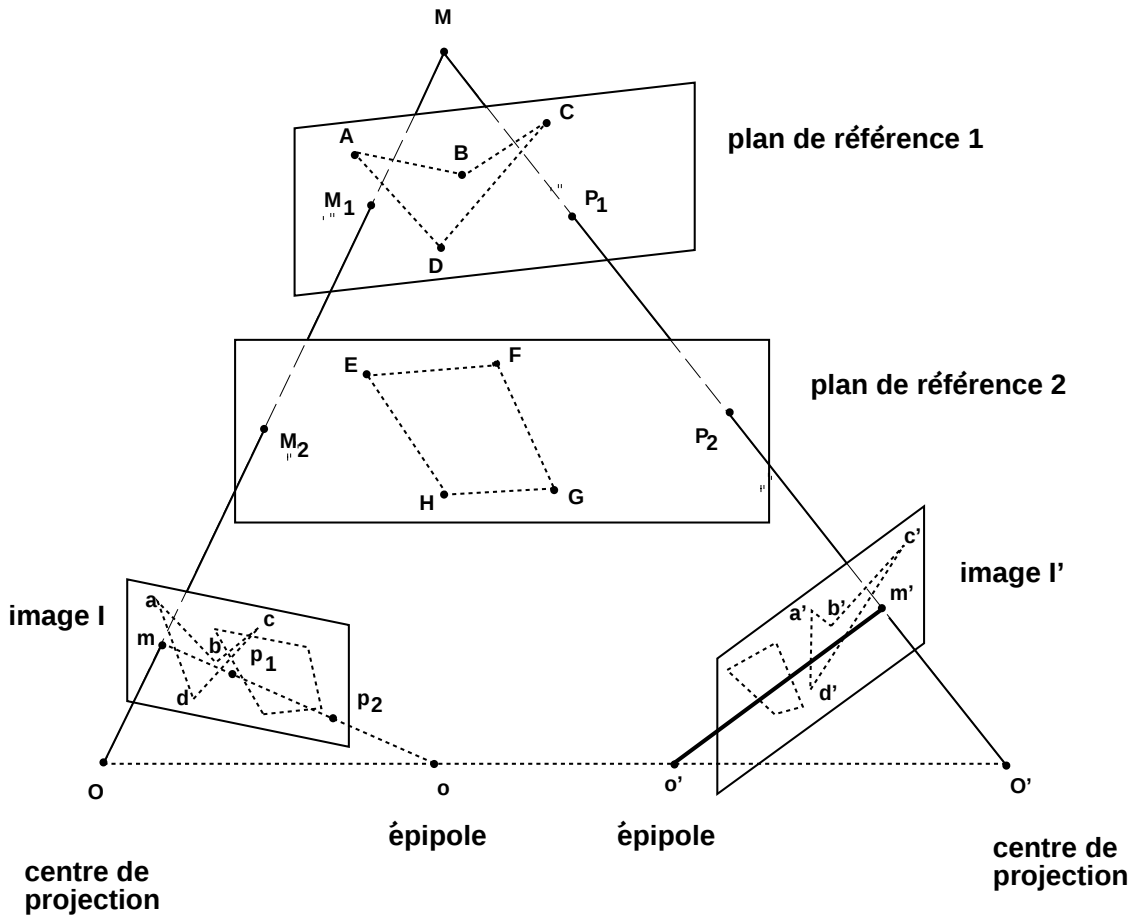


FIG. 2.5 - Reconstruction des droites épipolaires

dans les images : la position tridimensionnelle des points de référence n'est pas utilisée, seule la position de leurs images intervient. La détermination de la géométrie épipolaire est donc possible simplement à partir de deux images où les points de référence ont été identifiés et mis en correspondance.

2.3 Reconstruction avec 6 points non coplanaires

Nous avons montré dans la partie précédente comment résoudre les problèmes présentés en introduction, à l'aide de deux groupes de référence de quatre points coplanaires. La même technique est en fait applicable avec six points de référence non quatre à quatre coplanaires, c'est-à-dire en position générique.

Pour cela on étend au cas tridimensionnel la technique développée par Trip [Tri 87] et Stewart [Ste 90] dans le cas plan. Cette technique est basée sur le théorème de Chasles (voir paragraphe 1.5.1) et la conservation du birapport, non plus pour des droites mais pour des plans contenant les points de référence.

Tous les problèmes précédemment traités sont résolus, sauf celui de la géométrie épipolaire: en effet, si il n'existe aucune contrainte sur les points de référence, sept points sont nécessaires pour calculer la géométrie épipolaire et huit points permettent une résolution linéaire [Fau 92a].

2.3.1 Où est la caméra?

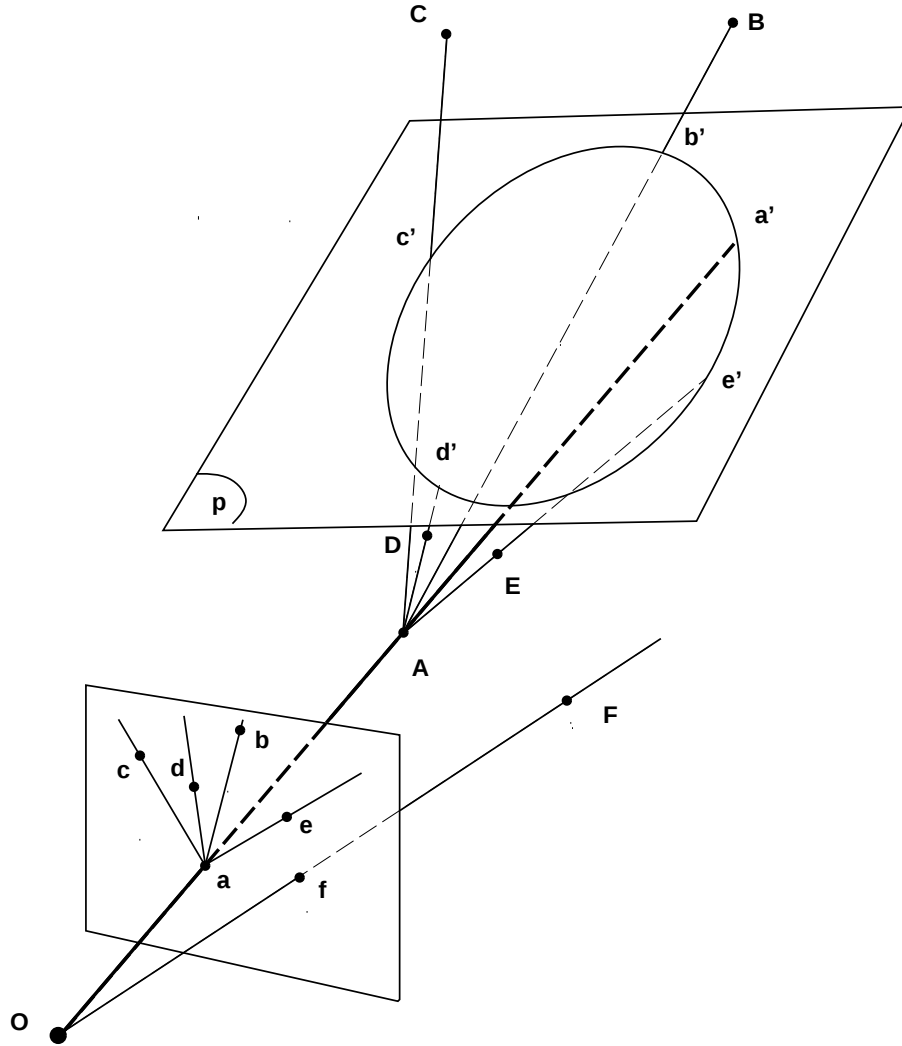


FIG. 2.6 - Détermination de la ligne de vue associée à un point connu A

Considérons un ensemble de six points A, B, C, D, E, F de l'espace 3D et leurs projections a, b, c, d, e, f sur le plan image. Soit O le centre optique que l'on cherche à déterminer. Considérons le faisceau des plans passant par la droite OA et contenant respectivement les points B, C, D, E, F (voir la figure 2.6).

Les intersections de ces plans avec le plan image sont les droites $ab, ac, ad, ae,$

af). Nous pouvons donc mesurer le birapport de tout faisceau formé de quatre de ces plans.

Tout plan connu $\mathcal{P}$ intersecte ce faisceau en des droites passant par a' , intersection de OA avec $\mathcal{P}$. La droite AB (resp. AC, AD, AE, AF) intersecte le plan $\mathcal{P}$ au point connu b' (resp. $c', \dots, f'$).

On connaît le birapport $[a'; b', c', d', e']$. D'après le théorème de Chasles, le point a' appartient à la conique de $\mathcal{P}$ déterminée par b', c', d', e' et ce birapport. La même construction peut être faite avec le point f' au lieu de e' . a' appartient donc à la conique de $\mathcal{P}'$ déterminée par b', c', d', f' et le birapport $[a'; b', c', d', f']$.

On en déduit finalement que a' appartient simultanément à deux coniques ayant déjà les trois points b', c', d' en commun. a' est donc la quatrième intersection des deux coniques. La solution explicite du calcul de l'intersection de deux coniques données par trois de leur points d'intersection et deux birapports est donnée en annexe.

Nous avons ainsi reconstruit la ligne de vue Aa' . La même méthode est utilisée pour les autres points, et le centre optique O est l'intersection de ces différentes lignes de vue.

2.3.2 Comment rétroprojeter un point image?

Considérons à présent la reconstruction de la ligne de vue associée à un point m quelconque, observé dans le plan image. m correspond à la projection d'un point M de la scène dont on cherche à déterminer la position.

Nous gardons les mêmes notations que dans le paragraphe précédent. En utilisant la même procédure de reconstruction du point a' , nous pouvons déterminer le faisceau des plans passant par Aa' . Ainsi le birapport $[a; b, c, d, m]$ mesuré dans l'image détermine le plan OMA dans la scène. De la même manière, en utilisant le point de référence B au lieu de A , nous pouvons déterminer le plan OMB . L'intersection de ces deux plans est la droite OM , ligne de vue du point m .

2.3.3 Où est le point image?

Inversement, étant donné un point M dans la scène, comment le localiser dans l'image?

O, A, B, C, M étant connus, soit l_a la droite contenant a , intersection du plan OAM et du plan image. l_a est déterminée par le birapport de droites $[ab, ac, ad, l_a]$ égal au birapport de plans $[OAB, OAC, OAD, OAM]$, qui est directement calculable dans l'image.

On peut déterminer de la même façon une autre droite du plan image, l_b par exemple. Le point image m est alors l'intersection des droites l_a et l_b .

Cette construction n'est pas valable si le point M appartient aux plans intro-

duits pendant la construction. Ceci peut toujours être évité en intervertissant les points de référence.

2.3.4 Positionnement relatif symbolique

Un autre aspect fondamental de notre approche est qu'elle permet d'induire des informations relatives importantes avec peu de données structurées sur les références (pas suffisamment pour un étalonnage habituel). Il est possible par exemple de localiser des objets relativement les uns aux autres. Pour illustrer ce point, nous allons considérer le problème de localiser la ligne de vue associée à un point, par rapport à un parallélépipède de dimensions et angles non connus. En choisissant cet objet, on simplifie le problème en réduisant le nombre de ses paramètres : dans la figure 2.7 les six points de référence sont entièrement définis par le sommet A et les vecteurs $\vec{AB}$, $\vec{AC}$, $\vec{AE}$; par exemple le point D est défini par la relation : $\vec{AD} = \vec{AB} + \vec{AC}$.

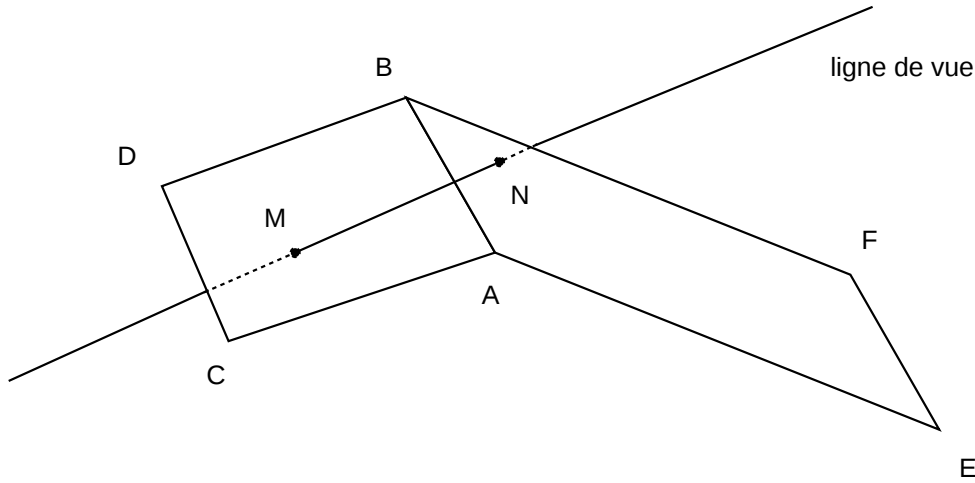


FIG. 2.7 - Localisation d'une ligne de vue par rapport à un parallélépipède

Nous devons d'abord être capables de localiser un point à partir de ses coordonnées projectives, en utilisant un parallélogramme comme référence. Par exemple, nous calculons la position du point M , intersection d'une ligne de vue avec le parallélogramme $ABDC$ (voir figure 2.7).

Soit :

$$k_1 = [AB, AC, AD, AM]$$

$$k_2 = [BA, BD, BC, BM]$$

où k_1 et k_2 ont été calculés à partir des points images.

On peut montrer que dans le cas où $ABDC$ est un parallélogramme,

- le vecteur $\vec{AM}$ est colinéaire à $\vec{AC} + k_1 \vec{AB}$

– le vecteur $B\vec{M}$ est colinéaire à $\vec{AC} - k_2\vec{AB}$

Ceci détermine les deux droites AM et BM et on peut alors calculer leur intersection, M . La position de M peut être exprimée en fonction du système de référence du parallélogramme, $(A, \vec{AB}, \vec{AC})$ et des birapports k_1 et k_2 :

$$\vec{AM} = \frac{1}{k_1 + k_2}(\vec{AC} + k_1\vec{AB})$$

De même, un autre point appartenant à la ligne de vue, N peut être exprimé dans le parallélogramme $ABFE$ en fonction des birapports :

$$k_3 = [AB, AE, AF, AN]$$

$$k_4 = [BA, BF, BE, BN]$$

On peut alors donner une équation paramétrique de la ligne de vue NM dans le système de référence du parallélépipède, $(A, \vec{AC}, \vec{AE}, \vec{AB})$:

$$P \in NM \quad - \quad \exists \lambda \quad /$$

$$\vec{AP} = \frac{1 - \lambda}{k_1 + k_2}\vec{AC} + \frac{\lambda}{k_3 + k_4}\vec{AE} - \left[\frac{(1 - \lambda)k_1}{k_1 + k_2} + \frac{\lambda k_3}{k_3 + k_4} \right] \vec{AB}$$

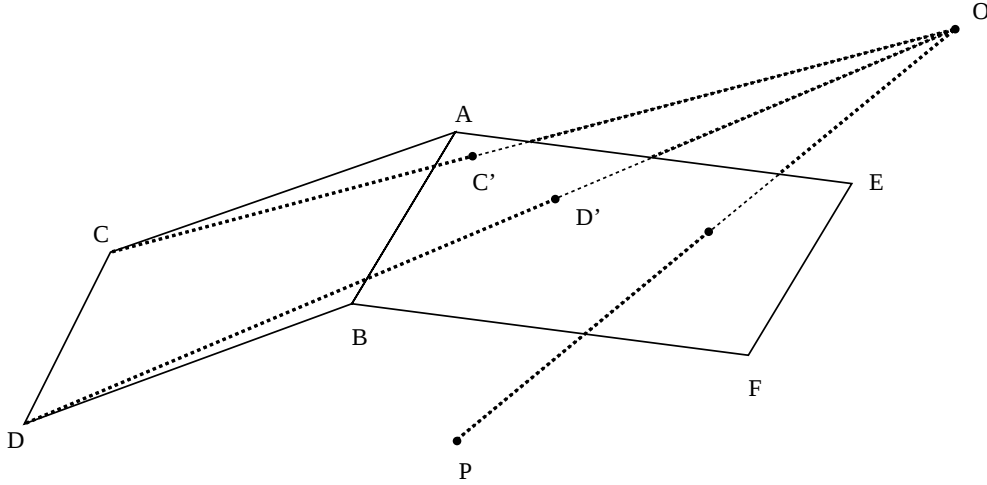


FIG. 2.8 - La contrainte que toutes les lignes de vue doivent contenir O aboutit à une relation liant les paramètres de l'équation de la ligne de vue.

On peut remarquer que la ligne de vue est entièrement définie par les quatre paramètres k_1, k_2, k_3 et k_4 . Or une droite de l'espace 3D passant en un point fixe O a deux degrés de liberté : ces quatre paramètres sont donc liés par deux relations. En exprimant que le centre de projection O est l'intersection des lignes de vue associées à C et D et que toute ligne de vue contient O , alors nous obtenant 2 équations liant respectivement k_1 et k_3, k_2 et k_4 :

$$k_1 = \frac{l_{3c} - k_3}{l_{3d} - l_{3c}}$$

$$k_2 = 1 - \frac{l_{4c} - k_4}{l_{4d} - l_{4c}}$$

où

$$l_{3c} = [AB, AE, AF, AC']$$

$$l_{4c} = [BA, BF, BE, BC']$$

$$l_{3d} = [AB, AE, AF, AD']$$

$$l_{4d} = [BA, BF, BE, BD']$$

et où C' et D' sont les intersections de OC et OD avec le plan $ABFE$.

En conclusion, des constructions de cette nature autorisent à positionner à partir d'une seule image un objet relativement à un autre. Il faut cependant une information de profondeur ; celle-ci peut être obtenue par exemple en connaissant à priori la taille relative des objets à positionner.

2.4 Expérimentations

Nous présentons ici quelques résultats destinés à tester la robustesse de notre méthode. Nous avons utilisé une scène polyédrique simple, présentant un cube et un polyèdre au premier plan et à l'arrière-plan deux feuilles de papier contenant des rectangles peints en noir.

Les positions des points de référence ont été mesurées avec une règle standard, et donc avec une précision de l'ordre du millimètre. D'autre part, les deux plans verticaux contenant les points de référence ont été supposés orthogonaux alors qu'ils forment en réalité un angle de 92 degrés.

La figure 2.9 montre une image de la scène. De cette image ont été extraits les contours à l'aide de l'extracteur de contours de Deriche ([Der 87]) puis une approximation polygonale aux moindres carrés a été effectuée sur les contours, pour déterminer les arêtes des objets de l'image. Les points utilisés pour la reconstruction sont les jonctions (obtenues par intersections des arêtes).

2.4.1 Reconstruction dans le plan

Une première expérimentation consiste à reconstruire des points appartenant à un plan connu.

Quatre points coplanaires sont choisis comme points de référence. Ils forment une base projective du plan qui les contient. Pour reconstruire un point appartenant

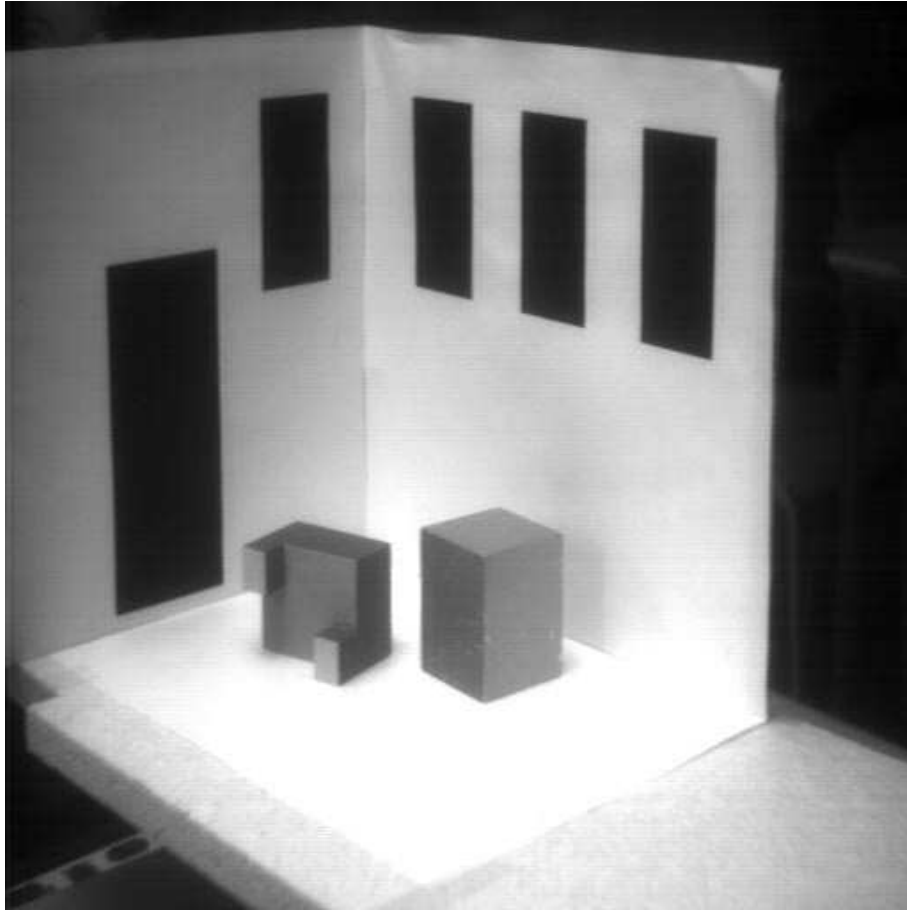


FIG. 2.9 - Image de la scène utilisée

à ce même plan, il suffit de déterminer l'intersection du plan et de la ligne de vue associée au point. Une seule image suffit pour déterminer la ligne de vue en utilisant la méthode décrite au paragraphe 2.2.1. La position ainsi calculée est comparée aux mesures faites à la règle directement dans la scène. Le tableau 2.1 montre les résultats obtenus en prenant comme points de références les points 8, 9, 18, 19.

Le choix d'autres points comme points de référence peut conduire à des résultats très différents : par exemple, si les points 8, 9, 10, 11 sont utilisés comme références, les points 16, 17, 18, 19 sont calculés avec des erreurs atteignant $6mm$. Ceci est dû au fait que ces points sont éloignés des points de référence : les erreurs sur la position des points de références sont amplifiées dans le calcul de reconstruction.

2.4.2 Reconstruction 3D

On reconstruit des points de la scène en stéréo en utilisant deux images entre lesquelles la scène a subi une rotation d'environ 40 degrés. La correspondance des

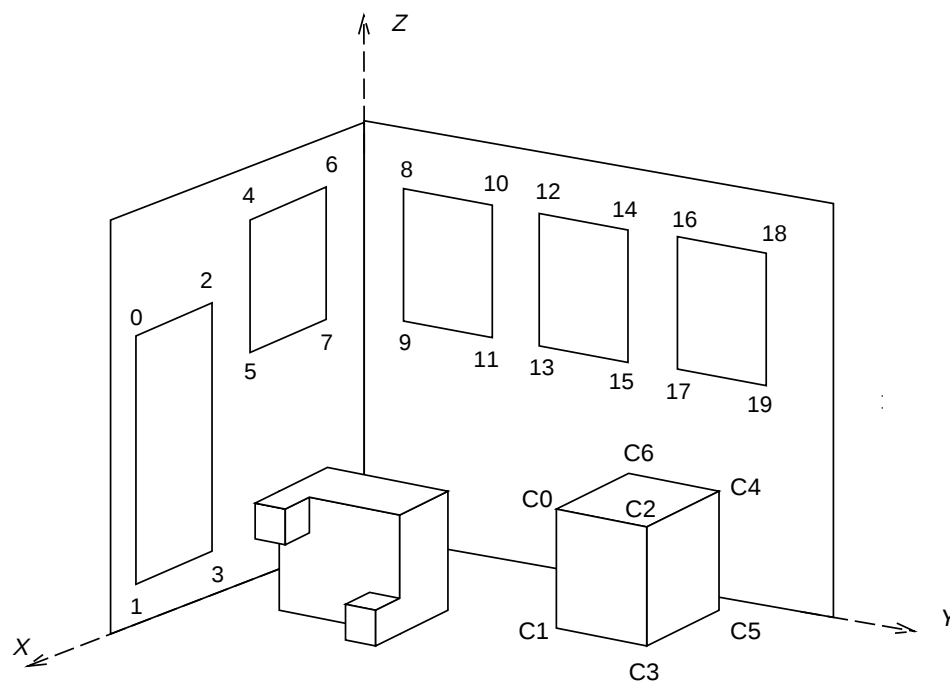


FIG. 2.10 - Les points utilisés pour le test

<i>points</i>	<i>valeurs mesurées</i>		<i>valeurs calculées</i>		<i>différence</i>	
	<i>y</i>	<i>z</i>	<i>y</i>	<i>z</i>	<i>dy</i>	<i>dz</i>
10	90	185	92.5	185	2.5	±1
11	90	113	91	113	1	±1
12	128	185	129	185	1	±1
13	128	113	128	113	±1	±1
14	174	185	175	185	1	±1
15	174	113	174	113	±1	±1
16	212	185	213	185	1	±1
17	212	113	210	113	2	±1

TAB. 2.1 - Résultats expérimentaux pour des points d'un plan connu. Les valeurs sont données en millimètres

points a été faite à la main, puisque le problème de la correspondance stéréo n'est pas ici l'objet de notre travail.

Les 6 points de référence utilisés sont 0, 1, 6, 7, 18, 19 (voir figure 2.10). Le tableau 2.2 présente les résultats de la reconstruction des sommets du cube. La taille exacte du cube est de 50mm. Les résultats sont donc exacts à 4% près.

Il faut noter que ces dernières mesures sont obtenues par différences de mesures dans l'espace, elles-mêmes bruitées.

<i>points</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>côtés</i>	<i>longueur</i>
C0	78.9	140	48.5	C0–C1	50.5
C1	79.1	141	-2	C0–C2	49.1
C2	81.3	189	47.5	C0–C6	48.9
C3	82.0	188	-1.5	C1–C3	47.1
C4	33.2	195	48.5	C2–C3	49.5
C5	34.4	194	-1.5	C2–C4	48.9
C6	30.3	145	49.0	C3–C5	48.0
				C4–C5	49.8
				C4–C6	50.1

TAB. 2.2 - Résultats expérimentaux de la reconstruction 3D du cube

2.5 Conclusion

Le positionnement relatif a déjà été exploré par Koenderink et van Doorn [Koe 89], mais uniquement en géométrie affine. Ils montrent comment la position affine d'un point peut être dérivée simplement à partir de deux projections orthogonales d'une scène contenant quatre points de référence. Cette approche est étendue ici à la géométrie projective. Une telle extension est nécessaire si on veut éviter une modélisation trop approximative du processus de formation de l'image. Nous avons montré que des problèmes fondamentaux de la vision artificielle tels que

- comment rétroprojeter un point image?
- comment projeter un point de l'espace?
- comment déterminer la géométrie épipolaire?

peuvent être résolus par des constructions géométriques directes en utilisant uniquement des propriétés de la géométrie projective.

Cette méthode de positionnement relatif a plusieurs avantages :

- Un étalonnage numérique explicite, souvent coûteux, est ici évité ; seuls des points de référence sont nécessaires.
- La méthode n'implique que des calculs simples, presque toujours le calcul arithmétique de birapports.
- La méthode fournit les mêmes possibilités que l'étalonnage traditionnel, et parfois plus, par exemple des inférences géométriques à partir des données de référence partielles.
- On utilise uniquement des propriétés des invariants projectifs purs.

- La méthode permet une appréciation intuitive du processus d'étalonnage ; les phénomènes d'instabilité peuvent être facilement isolés, nous donnant ainsi des clés pour l'élaboration de méthodes plus robustes et plus stables.

A la suite de ce travail, une méthode plus générale a été élaborée et implémentée dans notre équipe [Moh 92]. Elle permet de calculer la position relative des points d'une scène, à partir d'une séquence d'images réalisées avec une caméra non étalonnée. Les points observés sont mis en correspondance par une méthode de corrélation et la solution du positionnement est calculée par un processus d'optimisation non linéaire. Les résultats obtenus sur des images réelles ont confirmé la robustesse de cette méthode en présence de bruit.

Des méthodes similaires mais fondées sur une résolution linéaire non itérative et semble-t-il moins robustes ont été développées indépendamment par Faugeras et Maybank [Fau 92c] [May 92] et Hartley [Har 92b] [Har 92a]. Faugeras a aussi présenté les principes et les implications de la vision avec des caméras non étalonnées [Fau 92a]. Plus récemment ce type d'approche a été repris outre-atlantique par Shashua [Sha 92] et Barrett [Bar 91] [Bar 92].

Chapitre 3

Précision du positionnement relatif

3.1 Introduction

Ce court chapitre présente une étude comparative des méthodes de positionnement avec caméra étalonnée et à l'aide des invariants projectifs. L'objectif est d'obtenir une évaluation quantitative de la sensibilité aux erreurs de ces deux méthodes.

Lorsque l'on compare l'approche utilisant l'étalonnage traditionnel et l'approche géométrique, on doit prendre en compte le fait que les deux méthodes sont intrinsèquement différentes. La précision des résultats obtenus avec une caméra étalonnée dépend des erreurs sur les paramètres intrinsèques et extrinsèques de la caméra, des erreurs de discrétisation (dus à la numérisation de l'image) et des erreurs de distorsion (dus aux non-linéarités de l'optique).

Pour la méthode géométrique, le calcul des birapports et la détermination des positions spatiales qui en découle dépendent exclusivement de mesures dans l'image. En conséquence, seules les erreurs de discrétisation et de distorsion peuvent affecter le calcul. On peut remarquer cependant que la précision sur le point reconstruit dépend aussi de sa position par rapport aux points de référence.

Il est donc délicat de comparer des méthodes qui font intervenir des paramètres différents. Nous avons considéré un cas particulier simple, en choisissant des paramètres “standards”, et nous avons calculé l’erreur obtenue sur le point reconstruit en fonction des incertitudes sur les paramètres d’entrée.

Le plan de ce chapitre est le suivant : tout d’abord, nous considérons la méthode de rétroprojection à l’aide du birapport et nous cherchons à évaluer comment les incertitudes sur les points de l’image affectent l’incertitude sur le point ainsi reconstruit. Nous montrons qu’en dehors des cas dégénérés, la précision obtenue est comparable à celle obtenue avec une caméra étalonnée.

Nous considérons ensuite un cas simple de système stéréoscopique plan pour lequel la position des deux caméras est connue approximativement. Nous dérivons les expressions de la reconstruction associées à un point observé par ce système, d’une part en supposant les caméras étalonnées et d’autre part en utilisant une droite de référence et les propriétés de la géométrie projective. De ces expressions sont tirés les calculs d’incertitudes sur le point reconstruit, pour les deux méthodes.

Des courbes d’erreur sont ensuite présentées, qui permettent d’apprécier l’influence des différents paramètres sur la précision de la reconstruction. Ces courbes permettent aussi de comparer les performances des deux méthodes.

Une discussion des avantages et limitations respectives des deux méthodes conclue ce chapitre.

3.2 Analyse différentielle de la reconstruction par le birapport

Nous considérons ici les erreurs de discrétisation et de distorsion qui ont pour effet un déplacement du point dans l’image, de l’ordre du pixel dans le cas de la discrétisation et pouvant atteindre plusieurs pixels dans le cas de la distorsion. Nous montrons comment cette déviation affecte le calcul de la position d’un point appartenant à une droite munie d’une base projective. Soient A, B, C , trois points colinéaires formant le système de référence sur la droite L et soient a, b, c leurs projections sur une droite l par rapport au centre optique O . Soit d la projection du quatrième point de position inconnue, D (voir figure 3.1).

On note A l’abscisse du point A sur la droite L , par rapport à un repère affine de L arbitrairement choisi et de même pour tous les autres points. Ainsi $(B - A)$ est la longueur algébrique $\overline{AB}$. Avec :

$$k = \frac{(c - a)(d - b)}{(d - a)(c - b)} = \frac{(C - A)(D - B)}{(D - A)(C - B)} \quad (3.4)$$

on obtient

$$D = \frac{A(c - a)(d - b)(C - B) + B(d - a)(c - b)(A - C)}{(c - a)(d - b)(C - B) + (d - a)(c - b)(A - C)} \quad (3.5)$$

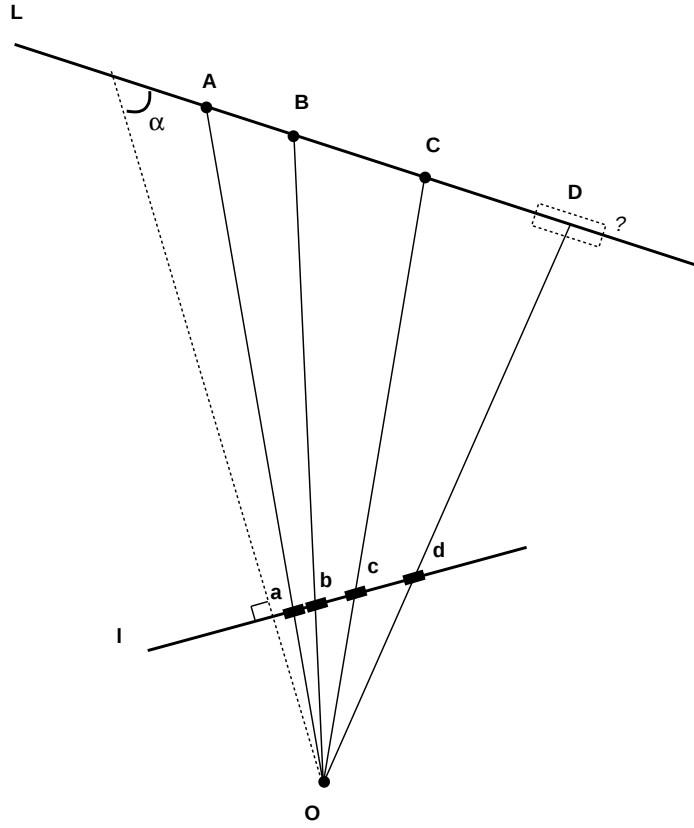


FIG. 3.1 - Erreur sur la position d'un point d'une droite reconstruit à partir du birapport

Pour obtenir l'effet de la variation de a, b, c, d sur la détermination de D , nous calculons les dérivées partielles de (3.5) par rapport à ces paramètres.

$$\frac{\delta D}{\delta a} = H(c - b)(b - d)(d - c)$$

$$\frac{\delta D}{\delta b} = H(d - a)(a - c)(c - d)$$

(3.6)

$$\frac{\delta D}{\delta c} = H(a - d)(d - b)(b - a)$$

$$\frac{\delta D}{\delta d} = H(b - c)(c - a)(a - b)$$

où

$$H = \frac{(A - B)(B - C)(C - A)}{[(c - a)(d - b)(C - B) + (d - a)(c - b)(A - C)]^2}$$

L'incertitude totale sur D est donnée par :

$$\delta D = \left| \frac{\delta D}{\delta a} \right| \delta a + \left| \frac{\delta D}{\delta b} \right| \delta b + \left| \frac{\delta D}{\delta c} \right| \delta c + \left| \frac{\delta D}{\delta d} \right| \delta d \quad (3.7)$$

Ce résultat indique que l'erreur de position sur D ne dépend que des positions relatives des points dans l'image (une erreur systématique est entièrement compensée). De plus les distances relatives interviennent avec un facteur globalement négatif. Ceci implique de façon immédiate que lorsqu'on éloigne le centre optique O (la caméra) du système de référence, la mesure de D est moins précise. De même une rotation du centre optique (et du plan image) par rapport au système de référence peut modifier de manière significative les distances relatives dans l'image et ainsi la précision obtenue sur la position de D (voir figure 3.2).

Ainsi que le montre aussi la figure 3.1, la relation entre deux segments en correspondance (par exemple $[a, b]$ dans l'image et $[A, B]$ sur la droite de référence) peut être approximée au moyen de l'équation suivante :

$$(A - B) \simeq \frac{D}{f} \cdot \frac{(a - b)}{\sin \alpha} = \frac{K}{\sin \alpha} \cdot (a - b) \quad (3.8)$$

Le facteur K dépend de la distance des points image au système de référence alors que le facteur $\sin \alpha$ prend en compte la rotation relative entre l'image et la droite de référence. L'approximation (3.8) est valide pour la plupart des configurations, excepté le cas d'un grand angle de vue.

Enfin, en considérant les équations (3.6), (3.7) et (3.8) il est facile de déduire que l'erreur sur D , δd , est aussi proportionnelle à K :

$$\delta D \simeq \frac{K}{4 \sin \alpha} (\delta a + \delta b + \delta c + \delta d) \simeq K \frac{\varepsilon_{(distorsion+discretization)}}{\sin \alpha} \quad (3.9)$$

Avec un étalonnage parfait, les erreurs de discrétisation génèrent sur la position d'un point P une incertitude proportionnelle à la distance de P à la caméra. On peut alors écrire, pour une caméra parfaitement étalonnée,

$$\varepsilon_{position} = K \varepsilon_{(distorsion+discretization)}$$

On voit donc qu'en dehors des configurations extrêmes, si une droite quelconque de la scène contient au moins trois points de référence connus, on peut déterminer la position d'un quatrième point D de la droite avec une précision comparable à celle obtenue avec une caméra parfaitement étalonnée.

3.3 Calcul des erreurs dans un cas particulier

Pour plus de simplicité, nous nous sommes placés dans une configuration bidimensionnelle : un point P est observé depuis deux caméras linéaires C_1 et C_2 (ou

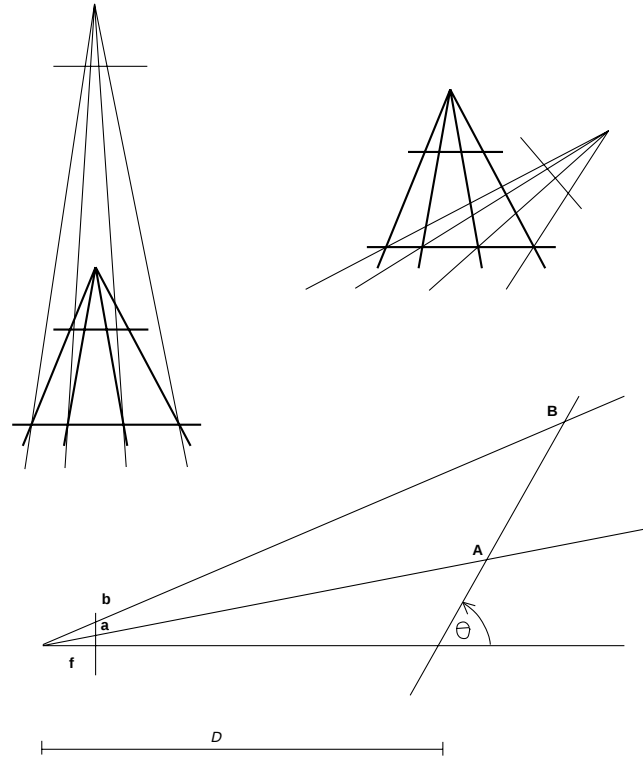


FIG. 3.2 - Si on éloigne la caméra de la droite de référence, ou si on incline l'axe optique, les points projetés sont plus rapprochés dans l'image.

deux positions éloignées d'une même caméra en mouvement). Les images obtenues sont ici linéaires. Les hypothèses sont les suivantes : on suppose que les deux caméras sont parfaitement étalonnées, (paramètres intrinsèques connus parfaitement) et que leurs positions et orientations sont connues avec une incertitude fixée. On se fixe aussi une incertitude sur les points de l'image, prenant en compte la discrétisation et la distorsion. On suppose de plus que le champ de vue contient une droite de référence L munie d'un repère projectif. La position de L dans le plan est supposée connue parfaitement. La figure 3.3 montre une des configurations étudiées, ainsi que les valeurs des différents paramètres, exprimées en *pixels*.

3.3.1 Reconstruction de P

En utilisant la méthode traditionnelle, il est possible de reconstruire la position de P à partir des deux images : pour chaque image, l'étalonnage permet de déterminer l'orientation θ de la ligne de vue passant par le centre optique C et le point image p . P est alors reconstruit par triangulation. La position (x, y) de P est donnée en fonction de $C_1(X_1, Y_1)$, $C_2(X_2, Y_2)$, θ_1 et θ_2 :

$$x = \frac{X_2 \tan \theta_2 - X_1 \tan \theta_1 - (Y_2 - Y_1)}{\tan \theta_2 - \tan \theta_1}$$

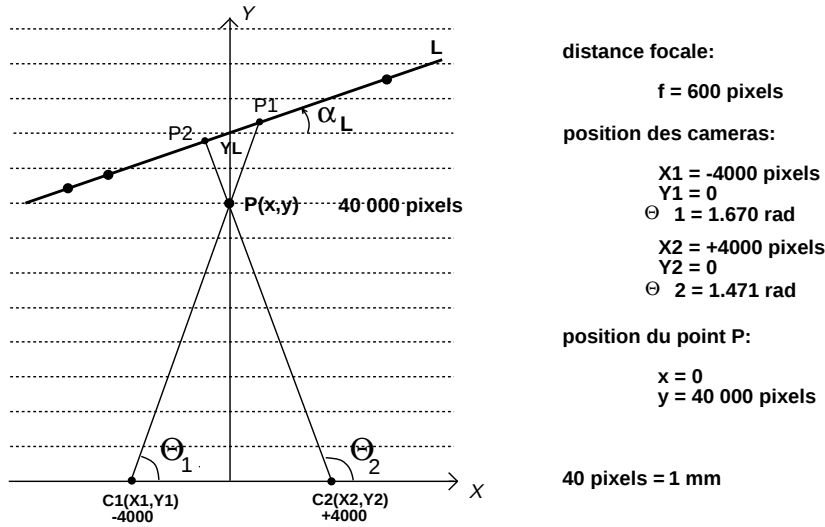


FIG. 3.3 - Un cas particulier : reconstruction stéréo planaire.

$$y = \frac{\tan \theta_1 \tan \theta_2 (X_2 - X_1) - Y_2 \tan \theta_1 + Y_1 \tan \theta_2}{\tan \theta_2 - \tan \theta_1} \quad (3.10)$$

La position de P peut aussi être calculée en utilisant partiellement la méthode du birapport. Considérons l'image issue de C_1 : le point P_1 de la droite de référence L peut être reconstruit par rétroprojection à l'aide du birapport. La ligne de vue issue de C_1 est ainsi définie par P_1 et le centre optique C_1 . De façon analogue, on détermine la seconde ligne de vue (C_2, P_2). La position de P est alors calculée comme l'intersection des deux lignes de vue (C_1, P_1) et (C_2, P_2). Cette procédure donne $P(x, y)$ en fonction de $C_1(X_1, Y_1)$, $C_2(X_2, Y_2)$, $P_1(x_1, y_1)$ et $P_2(x_2, y_2)$:

$$x = \frac{(X_1 y_1 - Y_1 x_1)(X_2 - x_2) - (X_2 y_2 - Y_2 x_2)(X_1 - x_1)}{(X_1 - x_1)(Y_2 - y_2) - (X_2 - x_2)(Y_1 - y_1)}$$

$$y = \frac{(X_1 y_1 - Y_1 x_1)(Y_2 - y_2) - (X_2 y_2 - Y_2 x_2)(Y_1 - y_1)}{(X_1 - x_1)(Y_2 - y_2) - (X_2 - x_2)(Y_1 - y_1)} \quad (3.11)$$

3.3.2 Calculs d'erreur

Dans le cas de la reconstruction traditionnelle, il suffit de calculer la matrice Jacobienne associée aux équations (3.10). On peut alors déterminer l'erreur sur les coordonnées x et y de P en fonction des différents paramètres d'entrée.

Les calculs d'erreurs pour la méthode géométrique sont plus délicats. On détermine d'abord la matrice Jacobienne associée aux expressions (3.11) de x et y et on en déduit l'incertitude sur x et y en fonction de $X_1, Y_1, X_2, Y_2, x_1, y_1, x_2, y_2$ et des incertitudes sur ces paramètres. Cependant, les coordonnées (x_1, y_1) et (x_2, y_2) de

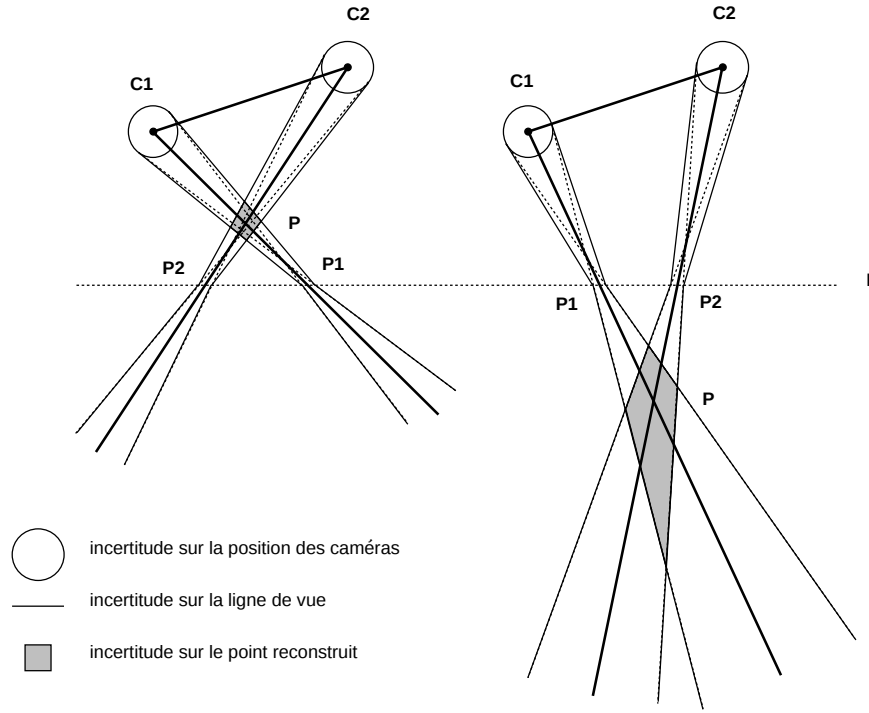


FIG. 3.4 - Incertitude sur la ligne de vue et le point reconstruit avec la méthode géométrique pour deux positions différentes de P , proche ou éloigné de la droite de référence L .

P_1 et P_2 sont calculées à partir de birapports faisant intervenir les points de référence ; il est donc impossible de les déterminer, ainsi que les incertitudes associées, indépendamment de la position des trois points de référence (voir eq. (3.6)).

Les points P_1 et P_2 peuvent être calculés comme intersections de la droite de référence L (définie par Y_L et α_L , voir figure 3.3) et des deux lignes de vue issues de C_1 et C_2 , ce qui donne :

$$\begin{aligned}
 x_1 &= \frac{Y_1 - Y_L - \tan \theta_1 X_1}{\tan \alpha_L - \tan \theta_1} \\
 y_1 &= \frac{\tan \alpha_L (Y_1 - \tan \theta_1 X_1) - \tan \theta_1 Y_L}{\tan \alpha_L - \tan \theta_1} \\
 x_2 &= \frac{Y_2 - Y_L - \tan \theta_2 X_2}{\tan \alpha_L - \tan \theta_2} \\
 y_2 &= \frac{\tan \alpha_L (Y_2 - \tan \theta_2 X_2) - \tan \theta_2 Y_L}{\tan \alpha_L - \tan \theta_2}
 \end{aligned}$$

Ceci permet de calculer les valeurs de (x_1, y_1) et (x_2, y_2) intervenant dans l'expression de l'incertitude sur x et y .

L'incertitude sur (x_1, y_1) et (x_2, y_2) est dérivée de l'équation (3.9), où $\varepsilon_{(distorsion+discretisation)}$ est fixé à 1 *pixel* et où K est exprimé en fonction de α_L , Y_L .

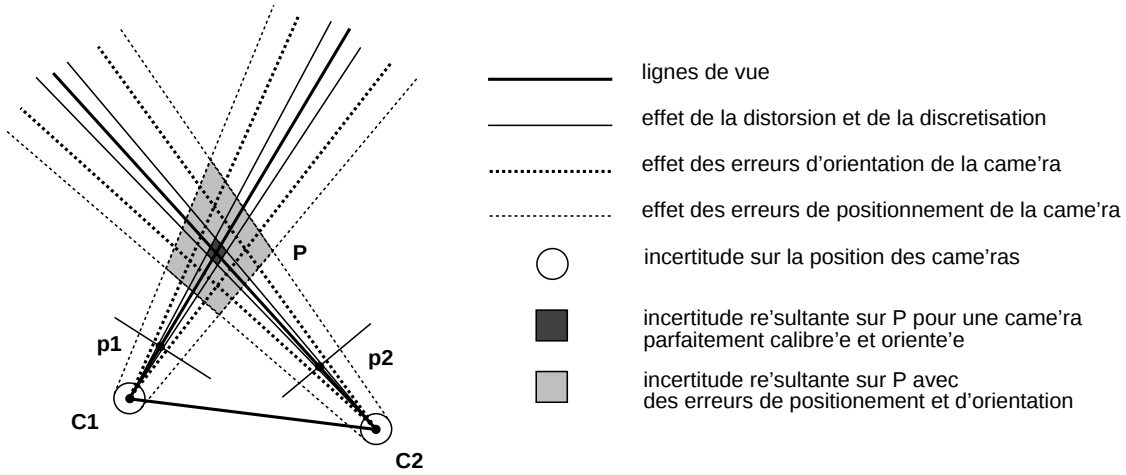


FIG. 3.5 - Incertitude sur la ligne de vue et le point reconstruit avec la méthode traditionnelle.

3.3.3 Présentation des courbes d'erreur

Pour permettre des considérations quantitatives, nous produisons des courbes représentant l'erreur sur le point reconstruit en fonction des incertitudes sur les données. Pour chacune des courbes, on étudie les variations de l'erreur sur P en fonction d'un paramètre alors que les autres paramètres sont fixés à des valeurs "standards".

Seuls les résultats pour la coordonnée y sont présentés : en effet on observe que les résultats obtenus sont analogues pour x et y , avec un facteur d'échelle qui dépend de la configuration : pour la configuration présentée, l'erreur δx est inférieure d'un facteur 10 à l'erreur δy .

Nous avons utilisé la même échelle sur l'axe des ordonnées pour toutes les courbes, ceci pour permettre une comparaison immédiate des différents résultats.

Considérons tout d'abord la figure 3.6 qui permet de comparer directement la sensibilité des deux méthodes aux erreurs sur la position des caméras. C'est le seul paramètre qui intervient à la fois dans les deux méthodes. Les erreurs sur les autres paramètres, telles que les erreurs de discrétisation et l'erreur sur l'orientation des caméras, sont supposées nulles. L'orientation de la droite de référence est fixée à sa valeur optimale, c'est-à-dire $\alpha_L = 0$. Ainsi, la seule erreur prise en compte est l'incertitude sur la position des caméras, δX .

La figure 3.6 représente sur le même graphe l'erreur obtenue avec la méthode traditionnelle et avec la méthode géométrique pour différentes valeurs de l'erreur de positionnement des caméras $\delta X = 100, 200, 300 \text{ pixels}$. Sur l'axe horizontal apparaît Y_L , la distance entre la droite de référence et les caméras.

La méthode avec étalonnage ne dépend bien sûr pas de la position de la droite de référence. C'est pourquoi l'erreur obtenue avec la méthode traditionnelle ap-

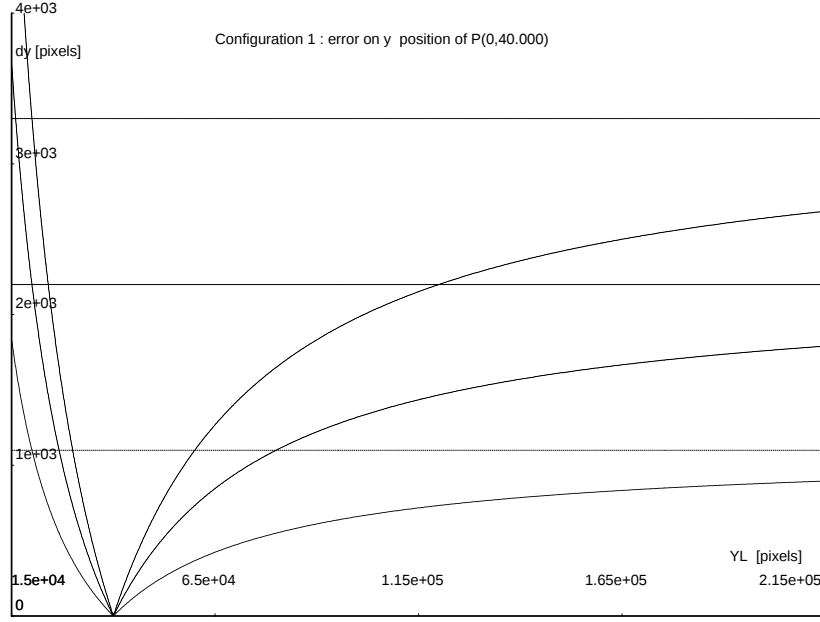


FIG. 3.6 - Configuration 1. Erreur sur les coordonnées x et y de P en fonction de Y_L . Chaque courbe correspond à un différent $\delta = \delta X_1 = \delta Y_1 = \delta X_2 = \delta Y_2$: de bas-en-haut $\delta = 100 \text{ pixels}$, $\delta = 200 \text{ pixels}$, $\delta = 300 \text{ pixels}$.

paraît comme une droite horizontale pour chacune des trois différentes valeurs de δX . Les courbes correspondent à la méthode géométrique et présentent comme attendu un minimum pour $Y_L = 40.000 \text{ pixels}$, lorsque le point P appartient à la droite de référence; il y a par contre dégénérescence lorsque la droite de référence devient proche des caméras ($Y_L \rightarrow 0$): dans ce cas la méthode géométrique n'est plus applicable. Cependant on peut observer que chaque droite horizontale est une asymptote pour la courbe correspondante. Ceci signifie que si la droite de référence est suffisamment éloignée des caméras, l'erreur obtenue par la méthode géométrique est toujours inférieure à celle obtenue avec la méthode traditionnelle. Dans notre cas cette distance minimale correspond à $1/4$ de la distance entre les caméras et l'objet observé.

Les figures 3.7 et 3.8 montrent l'incertitude sur le point reconstruit par la méthode géométrique en fonction de la position de la droite de référence. Elles représentent respectivement l'erreur en fonction de la distance Y_L entre la droite de référence et les caméras, son orientation α_L étant fixée et l'erreur en fonction de l'orientation α_L de la droite de référence, sa distance aux caméras Y_L étant fixée. Pour ces courbes nous considérons une erreur de 100 pixels sur les coordonnées des caméras. Par ailleurs, l'erreur sur P_1 et P_2 est calculée pour une erreur de discrétisation de 1 pixel .

La figure 3.7 présente des caractéristiques analogues à celles de la figure 3.6, en particulier la présence d'un minimum d'erreur lorsque le point est situé sur la

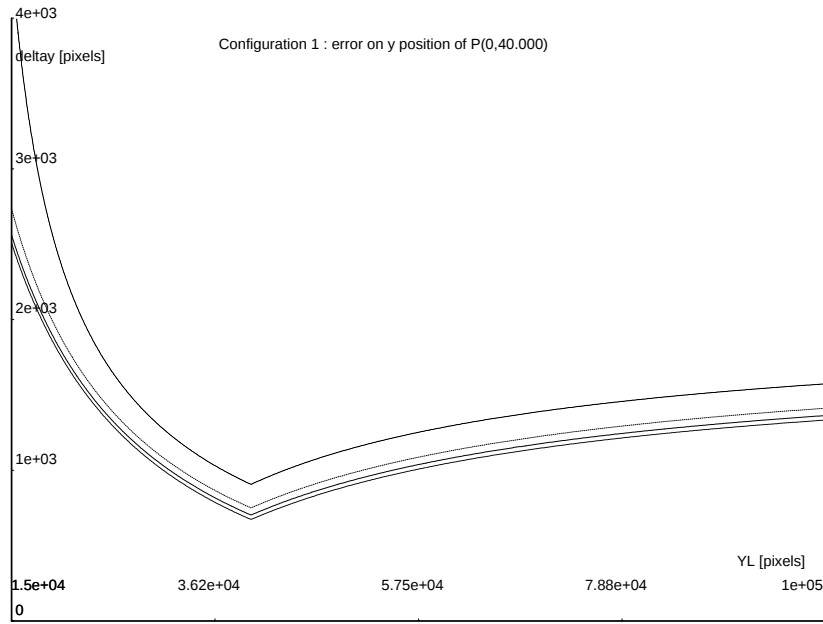


FIG. 3.7 - Configuration 1. Erreur sur la coordonnée y de P en fonction de Y_L . Chaque courbe correspond à un α_L différent : de bas-en-haut $\alpha_L = 0 \text{ rad}$, $\alpha_L = 0.4 \text{ rad}$, $\alpha_L = 0.8 \text{ rad}$, $\alpha_L = 1.2 \text{ rad}$.

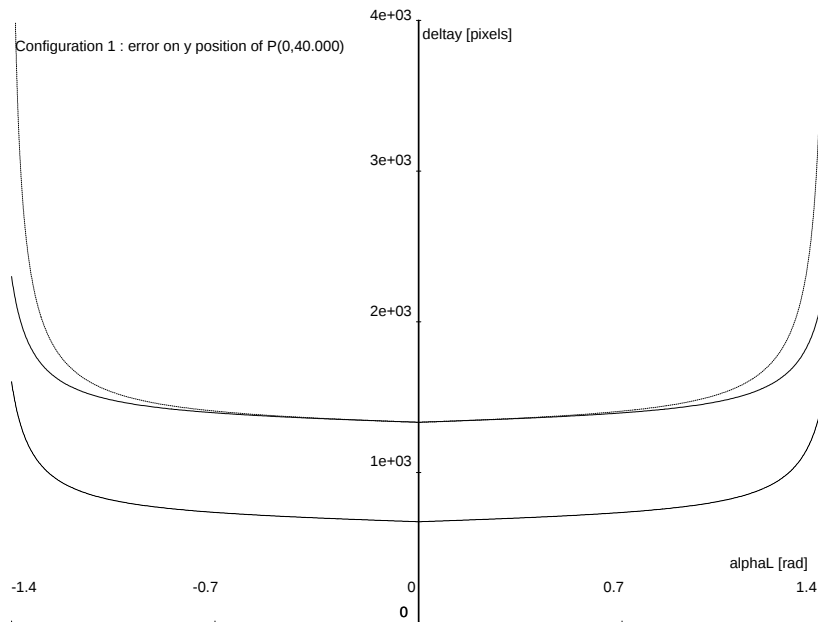


FIG. 3.8 - Configuration 1. Erreur sur la coordonnée y de P en fonction de α_L . Chaque courbe correspond à un Y_L différent : de bas-en-haut $Y_L = 40000 \text{ pixels}$, $Y_L = 100000 \text{ pixels}$, $Y_L = 25000 \text{ pixels}$.

droite de référence.

La figure 3.8 montre une bonne stabilité de l'erreur par rapport à l'orientation de la droite de référence. En fait, l'erreur est stable sur un intervalle d'environ $\frac{\pi}{2}$ autour d'une orientation préférentielle, augmentant seulement au voisinage des cas dégénérés (droite de référence contenant le centre de projection et donc confondue avec une ligne de vue!).

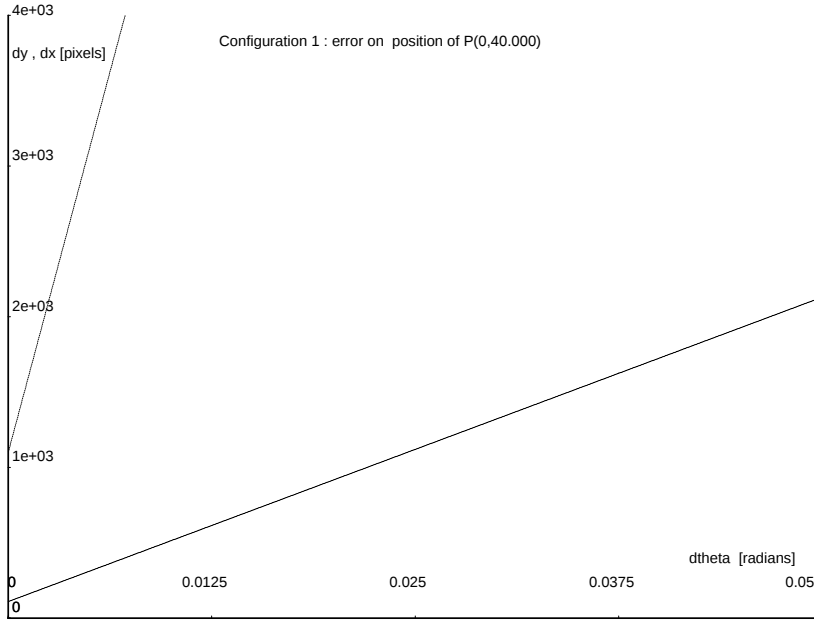


FIG. 3.9 - Configuration 1. Erreur sur les coordonnées de P en fonction de $\delta\theta$, l'incertitude sur l'orientation des lignes de vue θ .

De manière plus quantitative, on observe que la méthode géométrique produit des erreurs inférieures à 1500 pixels ; dans le cas traditionnel, ceci signifie une précision de 0.004 radians (~ 0.2 deg) sur la mesure des angles θ_1 et θ_2 , qui définissent l'orientation des lignes de vues. Cette évaluation est déduite de la figure 3.9 qui montre pour la méthode traditionnelle, l'erreur sur la position du point P en fonction de l'incertitude sur l'orientation des lignes de vue. (L'erreur sur la position des caméras est fixée à une valeur identique au cas géométrique, c'est-à-dire $\delta X_1 = \delta Y_1 = \delta X_2 = \delta Y_2 = 100$ pixels).

Ces observations se généralisent pour les autres configurations étudiées mais non présentées ici. (voir [Moh 91])

3.4 Conclusion

Les conclusions de cette étude sont les suivantes : la méthode géométrique, si elle ne demande pas d'étalonnage des caméras, requiert en revanche la connaissance de

points de référence. Si ces derniers sont insuffisants pour déterminer complètement la vue, il s'agit d'une configuration "dégénérée" pour la quelle la reconstruction ne peut pas être effectuée avec cette méthode. Cependant, en dehors de ces configurations, la précision obtenue sur le point reconstruit est comparable ou meilleure que la précision obtenue avec une caméra étalonnée :

- La méthode géométrique est alors moins sensible aux erreurs sur la position des caméras.
- Pour obtenir avec une caméra étalonnée une précision comparable à celle donnée par la méthode géométrique, l'étalonnage doit fournir une précision angulaire de 0.2 deg. sur les lignes de vue.

De plus, dans les discussions précédentes, nous avons toujours supposé que les paramètres intrinsèques de la caméra étaient parfaitement étalonnés. Mais ceci n'est jamais vraiment exact et il faut remarquer que les erreurs sur les paramètres intrinsèques affectent la reconstruction avec étalonnage alors qu'elles n'ont aucune incidence sur les résultats obtenus avec la méthode du birapport. Le calcul du birapport dans l'image est indépendant de la distance focale et de la position du projeté du centre optique (axe optique). Ceci est particulièrement intéressant lorsqu'on utilise une caméra mobile ; dans de nombreux cas, l'expérience a montré que l'évaluation de la position de la caméra est imprécise, à cause des conditions mécaniques. On peut aussi remarquer qu'une distance focale constante entre les vues interdit de pouvoir focaliser sur les objets observés. L'approche géométrique élimine ces contraintes et, en général, permet de s'affranchir d'une procédure de relocalisation continue, généralement basée sur des algorithmes très coûteux et sensibles. En fait, l'approche géométrique fournit explicitement une information basée sur l'objet, avec une incertitude proportionnelle à la distance. C'est donc un moyen tout à fait naturel de décrire l'environnement pour un grand nombre de tâches en robotique.

Chapitre 4

Stabilité des invariants projectifs

4.1 Introduction

Les applications des invariants projectifs au domaine de la vision sont nombreuses, comme l'a montré une récente conférence sur ce thème [Mun 92] : une première classe d'applications concerne la restitution des caractéristiques spaciales de la scène observée. Dans cette catégorie entrent le positionnement de la caméra, la reconstruction des points observés et l'étalonnage ou l'autocalibrage à l'aide d'invariants projectifs. Un autre domaine où les invariants projectifs sont naturellement utilisés est celui de la reconnaissance d'objets. En effet l'objectif de la reconnaissance est, à partir d'une base d'objets connus a priori, de détecter la présence d'un ou plusieurs objets de la base dans une image. Les objets sont identifiés grâce à des caractéristiques qui doivent être mesurables dans l'image, si possible indépendantes du point de vue et elles doivent permettre de discriminer les objets de la base. Pour les objets polyédriques, les propriétés topologiques sont largement utilisées. Des systèmes de reconnaissance d'objets polyédriques qui sont basés sur ces caractéristiques ont été développés. On peut citer en particulier les travaux de H. Sossa dans notre équipe [Sos 92c] [Sos 92b] [Sos 92a].

La description des propriétés géométriques des objets est plus ardue : les caractéristiques euclidiennes (distances, angles) et affines (rapports de distance, parallélisme) ne sont pas conservées dans les images et leur valeur dépend du point de

vue. Leur utilisation est donc limitée aux cas où soit le point de vue est connu, soit la transformation entre les modèles et les images peut être modélisée de manière satisfaisante par une transformation euclidienne, une similitude ou une affinité. Dans le cas d'une séquence d'images proches, P. Gros [Gro 91] montre comment les invariants affines permettent un appariement fiable des points des images. Huttenlocher [Hut 91] propose une méthode de reconnaissance fondée sur les invariants affines.

Les invariants projectifs fournissent une caractérisation géométrique invariante par transformation projective. Ils sont donc conservés dans toutes les images d'un même objet, à condition que le modèle de projection perspective soit valide.

Plusieurs travaux proposent des méthodes de reconnaissance fondées sur les invariants projectifs. On peut citer en particulier ceux de Zisserman et al. [For 90a] [For 90b] [For 91a] [Rot 91] [For 91b] [For 92] qui utilisent les invariants projectifs de droites et de coniques pour la reconnaissance d'objets plans. Batatia et al. [Bat 91] ont développé un système de reconnaissance de polyèdres utilisant des caractérisations à la fois topologiques et quantitatives à partir d'invariants projectifs de cinq points coplanaires. Ils proposent en particulier des solutions pour caractériser les configurations dégénérées (points alignés ...). Meer [Mee 92] propose une méthode de reconnaissance fondée sur la mise en correspondance de groupes de cinq points ou cinq droites coplanaires : la comparaison des invariants permet un premier filtrage, puis un système de vote définit la correspondance point à point entre le modèle et l'image.

D'une manière générale, ainsi que nous l'avons nous même constaté à la suite de nos travaux sur le positionnement relatif, les techniques utilisant les invariants projectifs se heurtent à l'absence de résultats ou d'études concernant la stabilité de ces invariants. Ce type de considérations n'entraîne pas dans les préoccupations des théoriciens de la géométrie au siècle dernier. Les questions qui se posent sont les suivantes :

- Comment les invariants projectifs sont ils affectés par la présence de bruit dans les images?
- La stabilité des invariants dépend elle des configurations (configurations de points, de droites ...) sur lesquelles on les calcule? Existe-t-il des configurations dégénérées?
- L'incertitude sur les invariants dépend elle du point de vue?

et dans le cas de la reconnaissance :

- Parmi les invariants à notre disposition, comment choisir les plus appropriés?
- Les invariants projectifs sont ils discriminants?

Des études approfondies ont déjà été menées pour répondre à ces questions dans le cas des invariants affines du plan [Hut 91]. Mais pour les invariants projectifs,

peu de résultats sont encore disponibles. Coelho et al. [Coe 91] ont observé les variations des invariants projectifs selon le point de vue. Ils utilisent des images réelles de figures géométriques planes très contrastées, pour éliminer l'influence des erreurs dues à l'extraction des contours. Différents invariants projectifs sont calculés pour plusieurs points de vues et leurs valeurs sont comparées. Cette étude permet d'évaluer de manière quantitative et sur un cas réel la stabilité des invariants projectifs. Cependant, le problème de l'évaluation a priori de cette incertitude n'est pas abordé et aucune information n'est apportée quand à l'incidence des erreurs sur les performances de la reconnaissance : en effet, les primitives de l'objet observé sont identifiées préalablement au calcul des invariants. Sanfeliu [San 92] propose un critère permettant de filtrer les configurations de cinq points coplanaires générant des invariants stables, et donc constituant des caractéristiques fiables pour la reconnaissance. Ce critère est basé sur une description probabiliste de la position des points et il suppose connus les effets de la transformation projective sur les incertitudes. Lenz et Meer [Len 92] ont recherché les invariants projectifs de cinq points coplanaires indépendants de l'ordre des points. Ils en proposent plusieurs et ils donnent des résultats sur leurs performances pour la reconnaissance avec des simulations bruitées. Meer [Mee 92] propose une mesure de qualité de la reconnaissance effectuée par son système. Cette mesure s'appuie sur une évaluation de la probabilité des mauvais appariements entre les groupes de cinq primitives.

Nous nous intéressons à la robustesse des invariants projectifs en nous plaçant dans le cadre de la reconnaissance.

Pour évaluer les performances des invariants projectifs comme indices de reconnaissance, nous considérons des configurations élémentaires formées de cinq points coplanaires, et caractérisées par deux invariants projectifs. Ces configurations forment notre base de modèles. Nous cherchons ensuite à identifier parmi tous les groupes de cinq points présents dans une image, ceux susceptibles d'être les images d'un groupe modèle, c'est-à-dire ceux caractérisés par les mêmes invariants projectifs.

Le choix des configurations de cinq points coplanaires a plusieurs raisons :

- Le choix d'une configuration plane est a priori restrictif, puisqu'on s'intéresse à des méthodes utilisables dans le cadre de la projection perspective 3D–2D. Cependant, Burns et al. [Bur 91] ont montré qu'il n'existe aucun invariant de la projection perspective 3D–2D pour un groupe de N points 3D quelconques. Comme le remarque Hartley [Har 92a], on est donc contraint de considérer :
 - soit les invariants de configurations 3D quelconques, mais calculables à partir de plusieurs images,
 - soit des configurations 3D non quelconques, c'est-à-dire vérifiant une contrainte géométrique, et qui peuvent alors admettre des invariants pour la projection 3D–2D, calculables à partir d'une seule image.

C'est pourquoi nous avons choisi d'étudier une configuration plane, la contrainte

de coplanarité engendrant l'existence d'invariants projectifs. (On est alors dans le cas d'une homographie de $\mathbb{P}^2$.)

- Si nous avons choisi un nombre de cinq points, c'est qu'il n'existe pas d'invariant projectif général pour quatre points coplanaires ou moins. En effet quatre points formant une base projective de $\mathbb{P}^2$ déterminent de manière unique une homographie de $\mathbb{P}^2$, ce que nous avons vu au paragraphe 1.3.2. Ceci implique que toute fonction de quatre points coplanaires invariante par transformation projective est constante : elle prend la même valeur pour tous les quadruplets de $\mathbb{P}^2$. Il n'existe donc pas d'invariant projectif (non constant) de quatre points coplanaires non alignés.
- On peut discuter le choix de points et non de droites qui sont des primitives a priori préférables lorsqu'on s'attache à la robustesse. D'un point de vue théorique, les invariants déduits des droites ont même forme que ceux déduits des points puisque les points et les droites sont duaux dans l'espace $\mathbb{P}^2$. Cependant, les incertitudes sur une droite ne s'expriment pas de manière analogue aux incertitudes sur un point.

Nous commençons ce chapitre en présentant les causes d'erreurs dans un processus de reconnaissance à l'aide d'invariants projectifs. Nous poursuivons par une étude théorique de l'incertitude sur le birapport de quatre points alignés. Nous dérivons en particulier le calcul de la fonction de répartition associée au birapport de quatre points alignés. Ceci nous permet une première évaluation du pouvoir discriminant d'un birapport. Nous proposons aussi une mesure de similarité entre deux birapports basée sur la fonction de répartition. Une étude différentielle nous permet ensuite d'observer l'effet sur le birapport des erreurs sur les quatre points. L'incertitude sur le birapport peut être évaluée par approximation linéaire ou majorée en considérant la monotonie de la fonction birapport. Cette étude met en évidence l'existence de configurations dégénérées et nous procure des outils pour la reconnaissance.

Nous considérons ensuite les groupes de cinq points coplanaires pour lesquels nous proposons différents invariants projectifs, pouvant être utilisés comme indices de reconnaissance. Nous cherchons d'une part à évaluer le pouvoir discriminant des invariants proposés et d'autre part nous proposons une évaluation a priori de l'incertitude sur un invariant projectif, afin d'obtenir un critère de rejet au moment de la reconnaissance. Nous présentons enfin les résultats obtenus sur des données simulées.

4.2 Origine des erreurs

Considérons un processus de reconnaissance où les configurations élémentaires sont les groupes de cinq points coplanaires. Pour chaque groupe de cinq points de

la base de modèles, on calcule les invariants projectifs k_1 et k_2 qui lui sont associés. On cherche dans l'image le ou les groupes de cinq points caractérisés par les mêmes invariants projectifs, appelés groupes *candidats*.

Pour chaque groupe modèle on trouve en général plusieurs groupes candidats, dont au plus un n'est pas un *faux candidat*. Les faux candidats ont plusieurs origines :

- D'un point de vue strictement théorique, c'est-à-dire en ne considérant aucune incertitude sur les données, on peut voir apparaître de faux candidats. En effet, deux groupes de cinq points différents peuvent avoir les mêmes invariants projectifs : il faut et il suffit qu'ils soient en correspondance homographique. Si il existe une homographie entre cinq points de l'image et un groupe modèle de cinq points, ils auront exactement les mêmes invariants projectifs, et les cinq points de l'image constitueront un faux candidat pour le groupe modèle (voir figure 4.1).

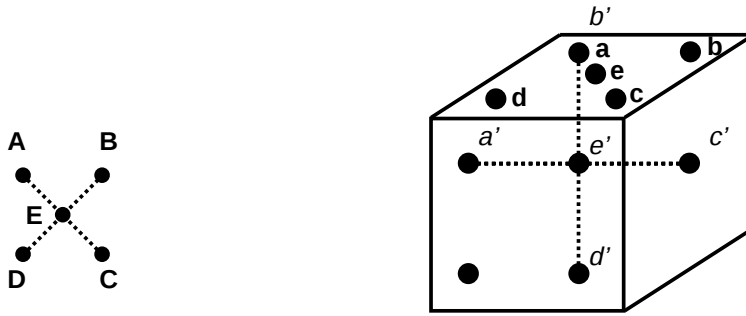


FIG. 4.1 - Premier type de faux candidat : dans une vue représentant des dés, on cherche le motif “5” : le groupe modèle est A, B, C, D, E . Dans une vue perspective parfaite, on a trouvé deux groupes candidats a, b, c, d, e et a', b', c', d', e' . Ce dernier est un faux candidat.

Ce cas de figure a une probabilité très faible de survenir effectivement. On ne l'a jamais observé dans nos simulations non bruitées utilisant des points répartis aléatoirement (voir paragraphe 4.5).

- Dans la pratique, il existe une incertitude sur les coordonnées des points images : elle provient de la discrétisation de l'image, des traitements bas-niveau (segmentation, extraction de contours et de jonctions) et du fait que le système optique ne réalise pas une projection perspective parfaite (distorsion). Si on tient compte des incertitudes sur les coordonnées des points images, on doit considérer un intervalle d'erreur associé à chaque invariant projectif. Dans l'espace des invariants, chaque groupe image sera donc associé à une zone d'incertitude et chaque groupe modèle sera associé à un point. Un groupe image dont la zone d'incertitude contient le point correspondant à un groupe modèle sera retenu comme un candidat potentiel pour ce modèle. C'est l'origine principale des faux candidats.

Il se pose donc le problème d'évaluer l'intervalle de confiance associé à un invariant projectif : si cet intervalle de confiance est sous-évalué, le "vrai" candidat ne sera pas retenu ; s'il est sur-évalué, un nombre superflu de faux candidats seront retenus.

La qualité des invariants projectifs comme indices de reconnaissance est donc directement liée à la proportion de faux candidats retenus par groupe modèle, et par rapport au nombre total de groupes images considérés, ainsi qu'à la possibilité d'évaluer de manière fiable l'intervalle de confiance associé chaque invariant projectif.

4.3 Etude théorique de stabilité du birapport

Dans cette section, nous cherchons à explorer le comportement des invariants projectifs en présence d'incertitudes sur les points qui sont utilisés pour leur calcul. Pour cela nous considérons l'invariant fondamental, le birapport, d'abord pour quatre points alignés, ensuite pour cinq points coplanaires. Une première étude probabiliste, nous conduit au calcul de la fonction de répartition du birapport de quatre points alignés. Dans une deuxième partie, une approche différentielle nous permet d'évaluer l'incertitude sur le birapport en fonction de l'incertitude sur les points.

4.3.1 Birapport de quatre points alignés

Quatre points d'une droite définissent un invariant projectif : leur birapport. On cherche à étudier le pouvoir discriminant de ce birapport, ainsi que sa stabilité par rapport aux erreurs sur les coordonnées des points.

Densité de probabilité du birapport

Le pouvoir discriminant du birapport est limité dans la mesure où deux configurations de quatre points différentes peuvent définir le même birapport. Ceci nous amène à considérer d'abord un point de vue probabiliste. La question qui se pose est la suivante : étant donné un réel k , quelle est la probabilité pour qu'un quadruplet de réels admette k comme birapport ?

La réponse est donnée par la fonction densité de probabilité du birapport, dont nous produisons ici l'expression analytique. Ce résultat a été établi en collaboration avec Kalle Åström, lors de sa visite au LIFIA en Septembre 1992 [Åst 92].

Considérons a, b, c et d , les réalisations de quatre variables aléatoires indépendantes A, B, C et D de distribution uniforme sur l'intervalle $[0, 1]$. Leur birapport est :

$$k(a, b, c, d) = \frac{(c - a)(d - b)}{(d - a)(c - b)}$$

Nous allons calculer la densité de probabilité et la fonction de répartition de la variable aléatoire $K = k(A, B, C, D)$.

Un cas particulier

Supposons pour l'instant que :

$$0 \leq a < b < c < d \leq 1 \quad (4.12)$$

La probabilité de cette configuration est $1/24$. Nous allons calculer la probabilité jointe que

$$P(K > \kappa, 0 \leq A < B < C < D \leq 1) = \iiint_{\Omega_\kappa} f_{(A,B,C,D)}(a, b, c, d) da db dc dd$$

où

$$\Omega_\kappa = \{(a, b, c, d) \mid k(a, b, c, d) \geq \kappa, 0 \leq a < b < c < d \leq 1\}$$

et $f_{(A,B,C,D)}(a, b, c, d)$ est la densité de probabilité jointe des quatre variables aléatoires A, B, C et D . Dans ce cas de figure nous avons $k(a, b, c, d) > 1$; nous supposons donc dans la suite que $\kappa > 1$.

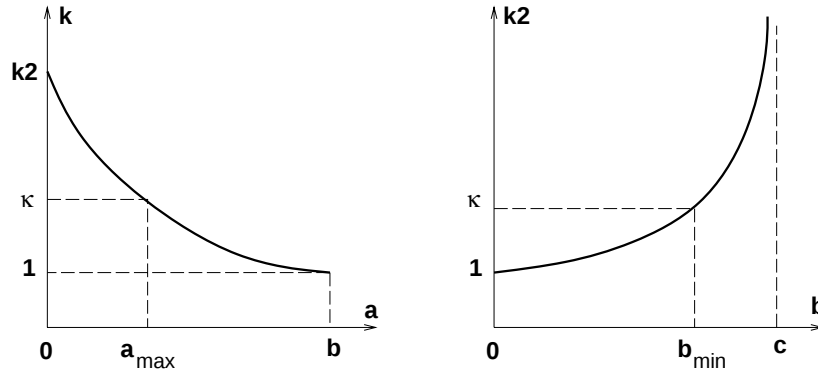


FIG. 4.2 -

Détermination de Ω_κ

Pour déterminer les bornes de l'intervalle Ω_κ , nous étudions tout d'abord les bornes pour a . Supposons que b, c et d sont fixes et vérifient la propriété (4.12). Le birapport k en fonction de $a \in [0, b]$ est tracé dans la figure 4.2. En utilisant l'hypothèse (4.12), on constate que la dérivée partielle

$$\frac{\partial k}{\partial a} = -\frac{(d-b)(d-c)}{(d-a)^2(c-b)}$$

est toujours négative. Pour $a = 0$, on obtient

$$k_2(b, c, d) = k(0, b, c, d) = \frac{c(d-b)}{d(c-b)}$$

et pour $a = b$ on obtient $k(b, b, c, d) = 1$. A présent, la condition $k(a, b, c, d) \geq \kappa$ ne peut être vérifiée pour une valeur de $a \in [0, b[$ seulement si $k_2(b, c, d)$ est supérieur à κ et $a \in [0, a_{\max}]$, où $a_{\max}$ vérifie $k(a_{\max}, b, c, d) = \kappa$. Après calculs, on obtient

$$a_{\max}(b, c, d, \kappa) = \frac{cb - cd + \kappa cd - \kappa db}{b - d + \kappa c - \kappa b}$$

Pour assurer que k_2 est supérieur à κ , nous fixons c et d et nous étudions k_2 en fonction de $b \in [0, c[$. Voir Figure 4.12. La dérivée partielle de k_2 par rapport à b est

$$\frac{\partial k_2}{\partial b} = \frac{c(d - c)}{d(c - b)^2}$$

En considérant l'hypothèse (4.12), on voit facilement que k_2 est strictement croissant. Il prend la valeur 1 pour $b = 0$ et $k_2 \rightarrow \infty$ quand $b \rightarrow c$. Pour tout $\kappa > 1$ pour toutes les valeurs de c et d telles que $0 \leq c < d \leq 1$ la condition $k_2(b, c, d) \geq \kappa$ est vérifiée pour $b \in [b_{\min}, c[$, où

$$b_{\min}(c, d, \kappa) = \frac{\kappa - 1}{\kappa - c/d} c$$

vérifie $k_2(b_{\min}, c, d) = \kappa$.

Pour récapituler, nous voulons trouver les valeurs de a, b, c et d avec $0 \leq a < b < c < d \leq 1$, dont le birapport $k(a, b, c, d)$ est supérieur à $\kappa > 1$. Pour tout choix de c, d , et κ avec $0 \leq c < d \leq 1$ and $\kappa > 1$ il est nécessaire et suffisant de prendre $b \in [b_{\min}, c[$ puis $a \in [0, a_{\max}]$. L'ensemble Ω_κ est donc

$$\Omega_\kappa = \left\{ (a, b, c, d) \left| \begin{array}{l} 0 \leq d \leq 1, \\ 0 \leq c < d, \\ b_{\min}(c, d, \kappa) \leq b < c, \\ 0 \leq a < a_{\max}(b, c, d, \kappa) \end{array} \right. \right\}$$

Calcul de la densité de probabilité

Nous calculons maintenant :

$$P_1 = P(K > \kappa, 0 \leq A < B < C < D \leq 1) = \iiint_{\Omega_\kappa} f_{(A,B,C,D)}(a, b, c, d) da db dc dd$$

En utilisant une distribution uniforme, ceci devient

$$P_1 = \int_0^1 \int_0^d \int_{b_{\min}}^c \int_0^{a_{\max}} 1 da db dc dd$$

Après des calculs, on (i.e. Maple) trouve que

$$P_1 = \frac{1}{24} (2\kappa(\kappa - 1) \ln\left(\frac{\kappa - 1}{\kappa}\right) + 2\kappa - 1)$$

On en déduit

$$F_0(\kappa) = \frac{1}{24} - P_1 = \frac{2}{24} \left(1 - \kappa(\kappa - 1) \ln \left(\frac{\kappa - 1}{\kappa} \right) - \kappa \right)$$

et

$$f_0(\kappa) = \frac{dF_0}{d\kappa}(\kappa) = \frac{2}{24} \left((2\kappa - 1) \ln \left(\frac{\kappa}{\kappa - 1} \right) - 2 \right)$$

Ces deux fonctions sont les bases sur lesquelles nous allons construire la fonction de répartition F_K et la densité de probabilité f_K de la variable aléatoire K .

Le cas général

Il existe 23 autres façons d'ordonner a, b, c et d . Les 24 ordonnancements peuvent être groupés en 6 groupes de 4. Dans chaque groupe, les birapports correspondants peuvent être calculés à partir du birapport associé à l'ordonnement (4.12) par une fonction définie sur l'intervalle $]1, \infty[$.

Les fonctions sont

$$\begin{array}{ll} g_1 :]1, \infty[\rightarrow]1, \infty[& g_1(x) = x \\ g_2 :]1, \infty[\rightarrow]1, \infty[& g_2(x) = 1 + 1/(x - 1) \\ g_3 :]1, \infty[\rightarrow]0, 1[& g_3(x) = 1/x \\ g_4 :]1, \infty[\rightarrow]0, 1[& g_4(x) = 1 - 1/x \\ g_5 :]1, \infty[\rightarrow]-\infty, 0[& g_5(x) = 1 - x \\ g_6 :]1, \infty[\rightarrow]-\infty, 0[& g_6(x) = -1/(x - 1) \end{array}$$

Supposons que X est une variable aléatoire avec une densité de probabilité $F_X(x) = P(X < x)$ et une distribution $f_X(x)$. On construit une nouvelle variable aléatoire $Y = g(X)$. Si g est une fonction strictement croissante alors

$$F_Y(y) = P(Y < y) = P(g(X) < y) = P(X < g^{-1}(y)) = F_X(g^{-1}(y))$$

Inversement, si g est strictement décroissante, alors

$$F_Y(y) = P(Y < y) = P(g(X) < y) = P(X > g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

Dans ces deux cas, la règle de transformation de la distribution obtenue est :

$$f_Y(y) = \frac{dF_Y}{dy}(y) = f_X(g^{-1}(y)) \left| \frac{d(g^{-1})}{dy}(y) \right| \quad (4.13)$$

En utilisant la règle de transformation (4.13) avec f_0 et les transformations g_1 à g_6 , il est possible de calculer la densité de probabilité jointe pour tous les autres ordonnancements de a, b, c et d . Il s'avère que les 6 expressions obtenues peuvent être représentées par 3 fonctions. Celles-ci sont montrées sur la figure 4.3. La densité

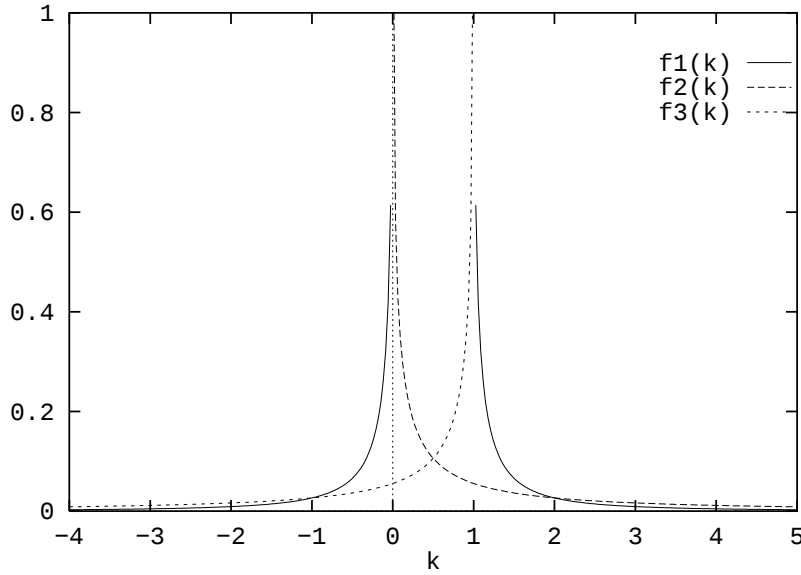


FIG. 4.3 -

de probabilité totale est obtenue en les additionnant. Le résultat de cette opération donne :

$$f_K(\kappa) = \begin{cases} f1(\kappa) + f3(\kappa) & \text{si } \kappa < 0 \\ f3(\kappa) + f2(\kappa) & \text{si } 0 < \kappa < 1 \\ f2(\kappa) + f1(\kappa) & \text{si } 1 < \kappa \end{cases}$$

$$f1(\kappa) = \frac{1}{3} \left((2\kappa - 1) \ln \left(\frac{\kappa}{\kappa - 1} \right) - 2 \right)$$

$$f2(\kappa) = \frac{1}{3} \left(\frac{(\kappa + 1) \ln(\kappa) + 2(1 - \kappa)}{(\kappa - 1)^3} \right)$$

$$f3(\kappa) = \frac{1}{3} \left(\frac{(\kappa - 2) \ln(1 - \kappa) - 2\kappa}{\kappa^3} \right)$$

La fonction de répartition est donnée par :

$$F_K(\kappa) = \begin{cases} 0 & \text{si } \kappa = -\infty \\ F1(\kappa) + F3(\kappa) & \text{si } -\infty < \kappa < 0 \\ 1/3 & \text{si } \kappa = 0 \\ 1/2 + F2(\kappa) + F3(\kappa) & \text{si } 0 < \kappa < 1 \\ 2/3 & \text{si } \kappa = 1 \\ 1 + F1(\kappa) + F2(\kappa) & \text{si } 1 < \kappa < \infty \end{cases}$$

$$F1(\kappa) = \frac{1}{3} \left(\kappa(1 - \kappa) \ln \left(\frac{\kappa - 1}{\kappa} \right) - \kappa + \frac{1}{2} \right)$$

$$F2(\kappa) = \frac{1}{3} \left(\frac{\kappa - \kappa \ln(\kappa) - 1}{(\kappa - 1)^2} \right)$$

$$F3(\kappa) = \frac{1}{3} \left(\frac{(1 - \kappa) \ln(1 - \kappa) + \kappa}{\kappa^2} \right)$$

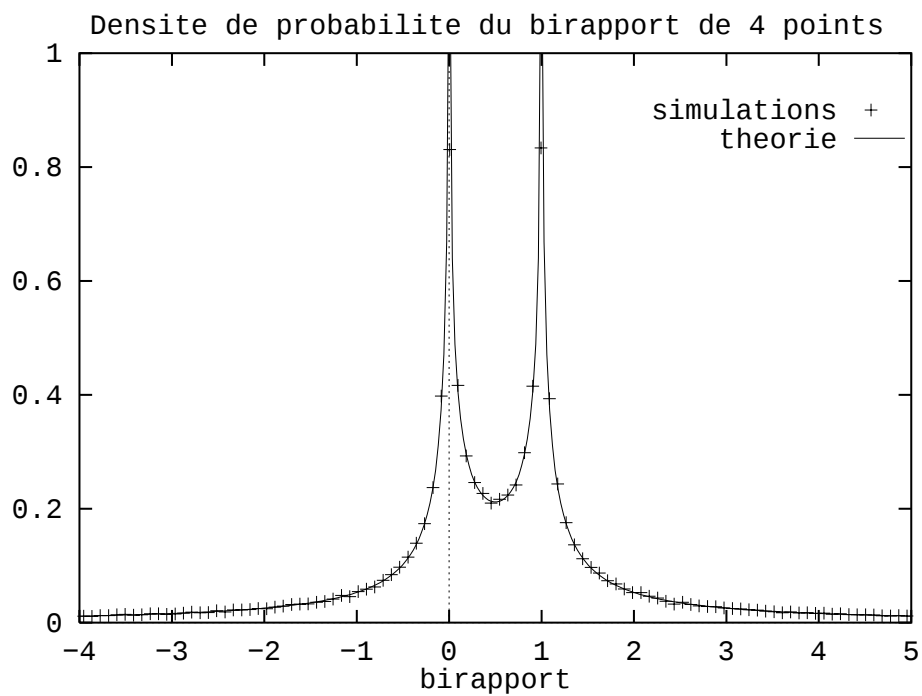


FIG. 4.4 -

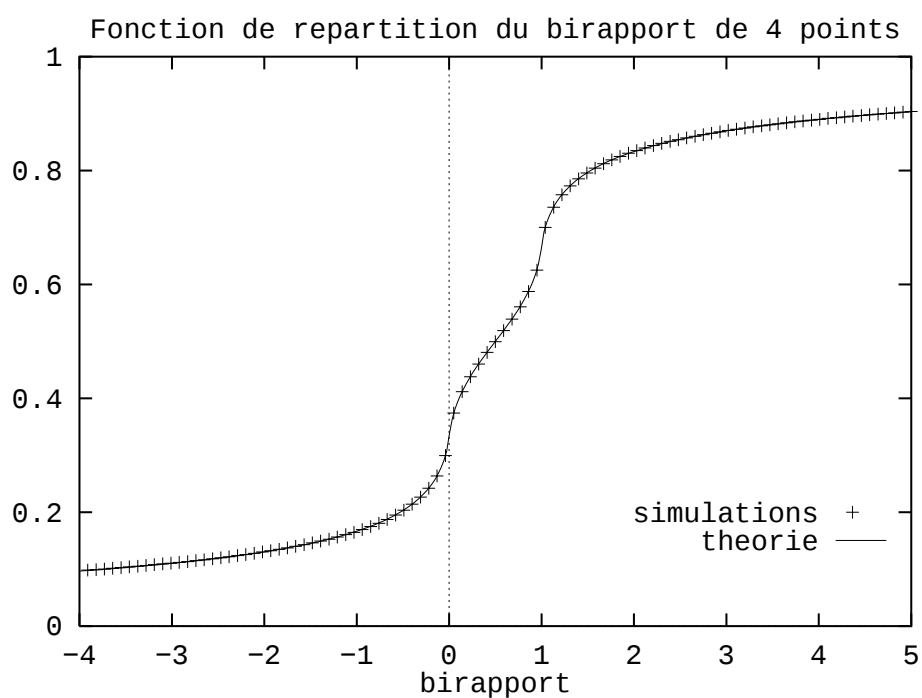


FIG. 4.5 -

Validation

Pour valider nos résultats, nous avons produit 300 000 réalisations des variables aléatoires uniformes A , B , C et D en utilisant un générateur de nombres pseudo-aléatoire. Le birapport de chaque quadruplet a été calculé. La densité de probabilité et la fonction de répartition ont été estimées à partir de ces données. Le résultat ainsi que la fonction analytique sont montrés sur la figure 4.4 et la figure 4.5.

Applications

Ce résultat peut être utilisé de plusieurs manières.

Tout d'abord, la densité de probabilité peut être utilisée pour calculer le facteur de discrimination associé à une configuration de quatre points alignés. Supposons que nous voulons identifier une configuration de quatre points dont le birapport est compris dans $[k - \Delta k, k + \Delta k]$, parmi N autres configurations de quatre points supposées conformes au modèle de distribution aléatoire que nous avons utilisé. Alors nous pouvons évaluer qu'en moyenne $(F_K(k + \Delta k) - F_K(k - \Delta k)) N$ de ces configurations seront dans l'intervalle $[k - \Delta k, k + \Delta k]$.

La fonction de répartition peut aussi être utilisée comme mesure de similarité entre deux configurations. Il faut noter que la comparaison de birapports est un point délicat : en effet le birapport est une notion projective, qui peut en théorie prendre des valeurs infinies. La différence relative est une mesure peu satisfaisante : par exemple, des birapports de 10^6 et $2 \cdot 10^6$ sont similaires alors que des birapports de 1 et 2 correspondent à des configurations très différentes, l'une étant une configuration dégénérée (deux points confondus), l'autre étant la configuration dite *harmonique*. De même, des birapports ayant des très grandes valeurs de signes opposées sont proches, car tous deux proches de l'infini.

Nous proposons comme mesure de distance entre deux birapports k_1 et k_2 la probabilité pour un birapport k d'être situé dans l'intervalle $[k_1, k_2]$. C'est-à-dire :

$$d_1(k_1, k_2) = F_K(k_2) - F_K(k_1)$$

La distance entre deux birapports est alors comprise entre 0 et 1. Pour tenir compte de la proximité des birapports à l'infini, il est préférable d'utiliser comme distance :

$$d_2(k_1, k_2) = \min(F_K(k_2) - F_K(k_1), 1 - F_K(k_2) - F_K(k_1))$$

Les valeurs obtenues sont alors comprises entre 0 et 1/2.

On peut remarquer toutefois que le calcul de la fonction densité par tirage aléatoire n'est valable que si le tirage respecte les contraintes présentes dans le cas réel : par exemple, si les quadruplets sont ordonnés, seule la restriction à l'intervalle $[1, \infty]$ de la densité de probabilité doit être prise en compte (voir figure 4.6). Si on impose une distance minimum entre coordonnées, la probabilité des configurations dégénérées est considérablement diminuée (voir figure 4.7).

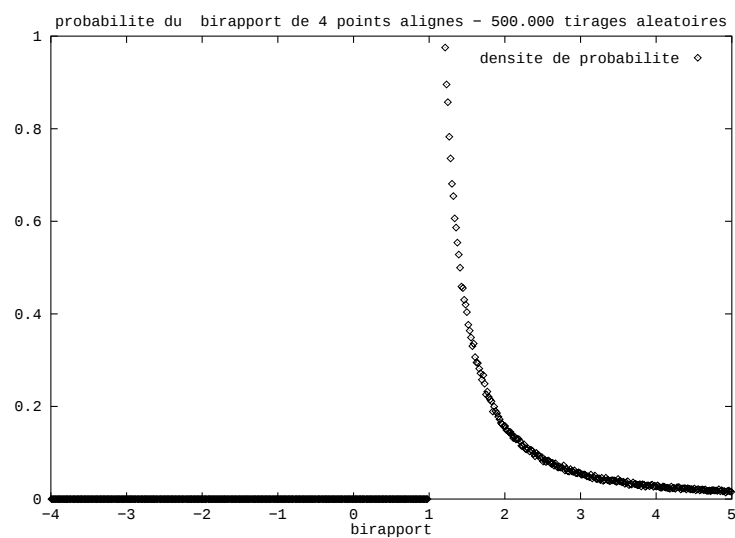


FIG. 4.6 - Densité de probabilité des birapports de quadruplets ordonnés, approximée par tirage aléatoire de 500,000 quadruplets.

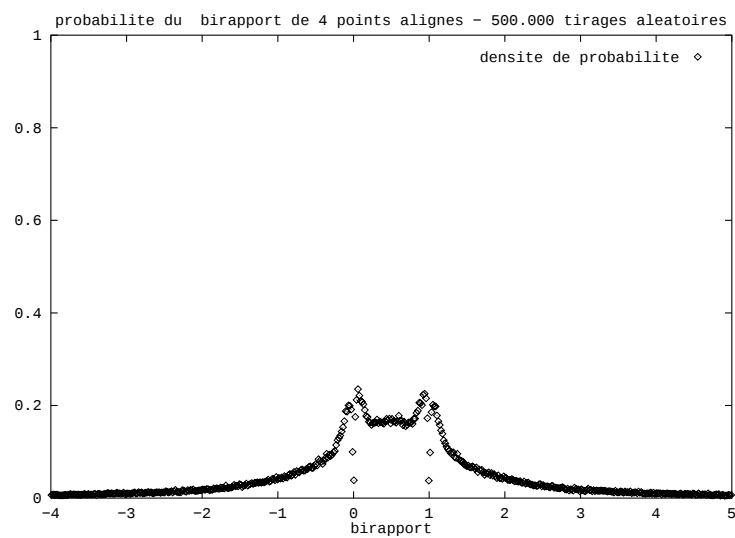


FIG. 4.7 - Densité de probabilité des birapports de quadruplets $[x_1, x_2, x_3, x_4]$ tels que $x_i \in [-10, 10]$ et $|x_i - x_j| > 1$. Approximation par tirage aléatoire de 500,000 quadruplets.

Approximation différentielle de l'incertitude

L'étude précédente nous permet d'évaluer le pouvoir de discrimination d'un birapport de quatre points donnés, mais elle ne donne aucune indication sur la sensibilité du birapport aux erreurs sur les points. Dans ce paragraphe, nous cherchons à évaluer l'incertitude sur le birapport de quatre points en fonction de l'incertitude sur ces points.

On considère quatre points de coordonnées x_1, x_2, x_3, x_4 mesurés avec une incertitude de dx_1, dx_2, dx_3, dx_4 où $dx_i > 0$. En linéarisant l'expression du birapport par son développement limité au premier ordre, nous pouvons donner une majoration de l'incertitude sur le birapport :

$$dk = \sum_{i=1}^4 \left| \frac{\partial k}{\partial x_i} \right| dx_i$$

Si l'erreur sur les coordonnées est de type gaussien, et si dx_i^2 représente la variance sur la coordonnée x_i , alors l'incertitude sur le birapport a pour variance :

$$dk^2 = \sum_{i=1}^4 \left(\frac{\partial k}{\partial x_i} \right)^2 dx_i^2$$

En supposant identiques les incertitudes sur chaque point, le jacobien du birapport donne directement la variance de l'incertitude sur le birapport.

$$dk = \sqrt{\sum_{i=1}^4 \left(\frac{\partial k}{\partial x_i} \right)^2} dx$$

Coelho [Coe 91] en déduit directement que $dk \propto k$. Une observation plus fine montre que cette relation linéaire se vérifie en dehors des zones dégénérées. Les zones dégénérées correspondent aux deux hyperplans de l'espace des coordonnées : $x_1 = x_4$ et $x_2 = x_3$ pour lesquels le birapport est infini (voir figure 4.8). Par contre, pour les grandes valeurs de k , $dk \propto k^2$ (voir plus loin).

Cependant, étant donné une valeur de l'incertitude sur les points et une valeur du birapport, il n'est pas possible d'en déduire directement l'incertitude sur le birapport. En effet, le facteur de proportionnalité entre dk et k ne dépend pas seulement des incertitudes sur les points mais aussi de la configuration des points eux-mêmes. Ceci est mis en évidence en considérant des configurations homothétiques (x_1, x_2, x_3, x_4) et $(\lambda x_1, \lambda x_2, \lambda x_3, \lambda x_4)$. Les dérivées partielles de k par rapport aux quatre coordonnées sont :

$$\begin{aligned} \frac{\partial k}{\partial x_1} &= \frac{(x_4 - x_2)(x_3 - x_4)}{(x_3 - x_2)(x_4 - x_1)^2} \\ \frac{\partial k}{\partial x_2} &= \frac{(x_3 - x_1)(x_4 - x_3)}{(x_4 - x_1)(x_3 - x_2)^2} \end{aligned}$$

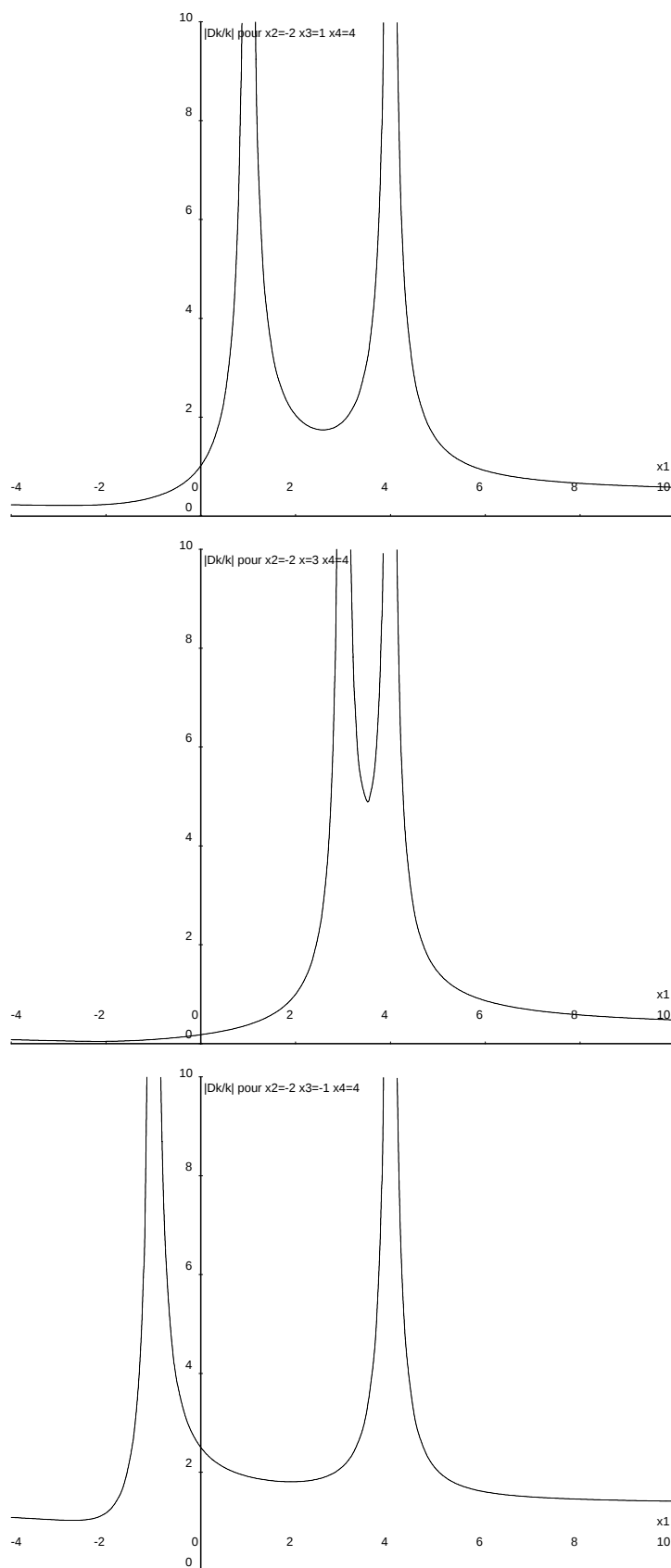


FIG. 4.8 - L'incertitude relative sur le birapport $k(x_1, x_2, x_3, x_4)$ en fonction de x_1 , et pour trois configurations de x_2, x_3, x_4 .

$$\begin{aligned}\frac{\partial k}{\partial x_3} &= \frac{(x_4 - x_2)(x_1 - x_2)}{(x_4 - x_1)(x_3 - x_2)^2} \\ \frac{\partial k}{\partial x_4} &= \frac{(x_3 - x_1)(x_2 - x_1)}{(x_3 - x_2)(x_4 - x_1)^2}\end{aligned}$$

Ceci implique

$$k(\lambda x_1, \lambda x_2, \lambda x_3, \lambda x_4) = \frac{(\lambda x_3 - \lambda x_1)(\lambda x_4 - \lambda x_2)}{(\lambda x_4 - \lambda x_1)(\lambda x_3 - \lambda x_2)} = k(x_1, x_2, x_3, x_4)$$

et

$$\frac{\partial k}{\partial x_i}(\lambda x_1, \lambda x_2, \lambda x_3, \lambda x_4) = \frac{1}{\lambda} \frac{\partial k}{\partial x_i}(x_1, x_2, x_3, x_4)$$

Des configurations ayant le même birapport n'engendrent donc pas nécessairement la même incertitude sur le birapport. Dans le cas de configurations homothétiques, l'incertitude sur le birapport varie de manière inversement proportionnelle aux distances entre les points. C'est un résultat attendu pour une grandeur qui dépend de rapports de distances algébriques. Une conséquence immédiate et conforme à l'intuition est que dans une même image, où les incertitudes sur les points peuvent être considérées comme identiques, il sera préférable de choisir des configurations comportant des points distants les uns des autres pour obtenir des birapports stables.

En conclusion, une configuration a un birapport d'autant plus stable qu'elle est éloignée des configurations dégénérées et que ses points sont distants les uns des autres. Un point important est qu'on ne peut pas évaluer correctement l'incertitude sur un birapport uniquement en connaissant sa valeur : accepter une marge d'erreur de 5% sur tout birapport ne donnera pas des résultats homogènes pour toutes les configurations. Par contre, à partir des coordonnées des points, on peut donner une estimation de l'incertitude sur le birapport de chaque configuration.

Etude de bornes

L'étude différentielle conduit à une approximation de l'incertitude comme une somme des incertitudes par rapport à chaque coordonnée. Cependant, on peut remarquer que le birapport $k(x_1, x_2, x_3, x_4)$ est monotone par rapport à chaque coordonnée. En effet la dérivée partielle de k par rapport à x_i a la forme suivante :

$$\frac{\partial k}{\partial x_i} = \frac{F(x_j, x_k, x_l)}{(x_j - x_i)^2} \quad \text{où } j, k, l \neq i$$

Elle est de même signe pour toutes les valeurs de x_i : k est donc monotone par rapport à x_i . De ce fait, les valeurs extrêmes de la fonction birapport sur l'intervalle $[x_1 - dx_1, x_1 + dx_1] \times [x_2 - dx_2, x_2 + dx_2] \times [x_3 - dx_3, x_3 + dx_3] \times [x_4 - dx_4, x_4 + dx_4]$ sont parmi les valeurs aux bornes : c'est-à-dire les 16 quadruplets appartenant à $\{x_1 - dx_1, x_1 + dx_1\} \times \{x_2 - dx_2, x_2 + dx_2\} \times \{x_3 - dx_3, x_3 + dx_3\} \times \{x_4 - dx_4, x_4 + dx_4\}$.

Ceci est valide si les incertitudes sur les points sont inférieures aux distances relatives entre les points : $dx_i < x_i - x_j$ pour tous i, j avec $i \neq j$.

Le minimum et le maximum de ces 16 valeurs définissent donc un intervalle d'incertitude autour du birapport de x_1, x_2, x_3, x_4 . Si on appelle ϵ_i le signe de $\frac{\partial k}{\partial x_i}$ alors

$$kmin = [x_1 - \epsilon_1 dx_1, x_2 - \epsilon_2 dx_2, x_3 - \epsilon_3 dx_3, x_4 - \epsilon_4 dx_4]$$

$$kmax = [x_1 + \epsilon_1 dx_1, x_2 + \epsilon_2 dx_2, x_3 + \epsilon_3 dx_3, x_4 + \epsilon_4 dx_4]$$

Cette évaluation n'utilise aucune approximation ou linéarisation. Elle donne une majoration exacte de l'incertitude sur le birapport.

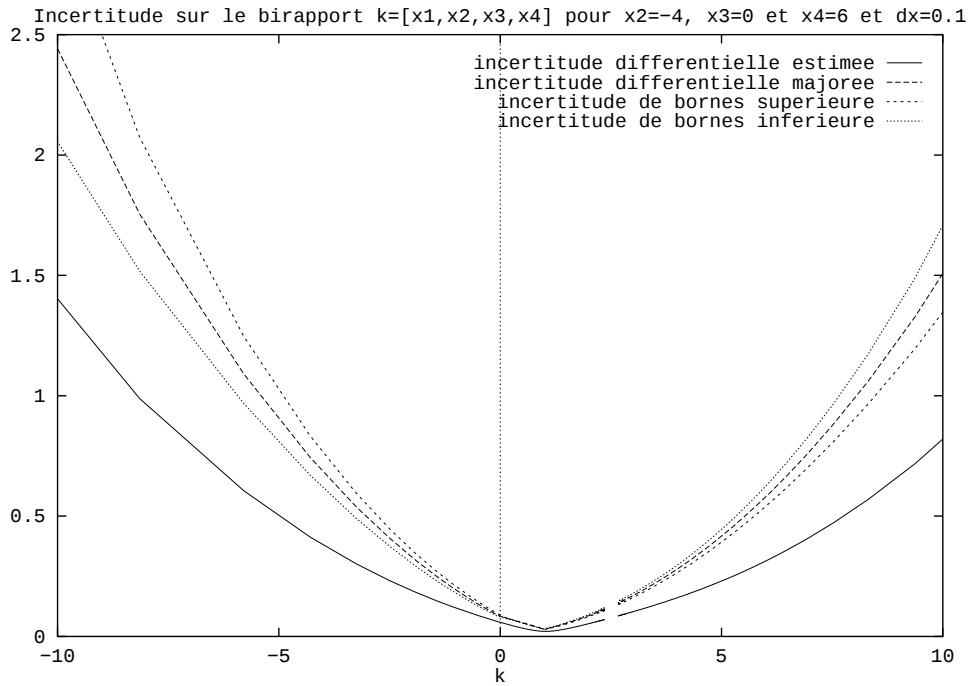


FIG. 4.9 - Comparaison des différents calculs d'incertitude du birapport

La figure 4.9 montre une comparaison des différents calculs d'incertitude sur le birapport. L'incertitude sur k est tracée en fonction de k , en faisant varier x_1 et en fixant x_2, x_3 et x_4 . L'incertitude sur les points a été fixée à 0.1, ce qui correspond à environ un dixième de la distance relative entre les points.

On observe que le minimum de l'incertitude intervient pour un birapport de valeur 1, ceci quelque soit le calcul de l'incertitude employé. Ceci reste vrai avec d'autres valeurs pour x_2, x_3 et x_4 . L'incertitude majorée issue de l'approximation linéaire peut parfois sous-estimer l'erreur maximum possible pour le birapport, mais qu'elle reste une bonne évaluation.

L'incertitude a une variation quasi linéaire en fonction de k autour de la valeur 1, elle a une variation quadratique en fonction de k pour les grandes valeurs de

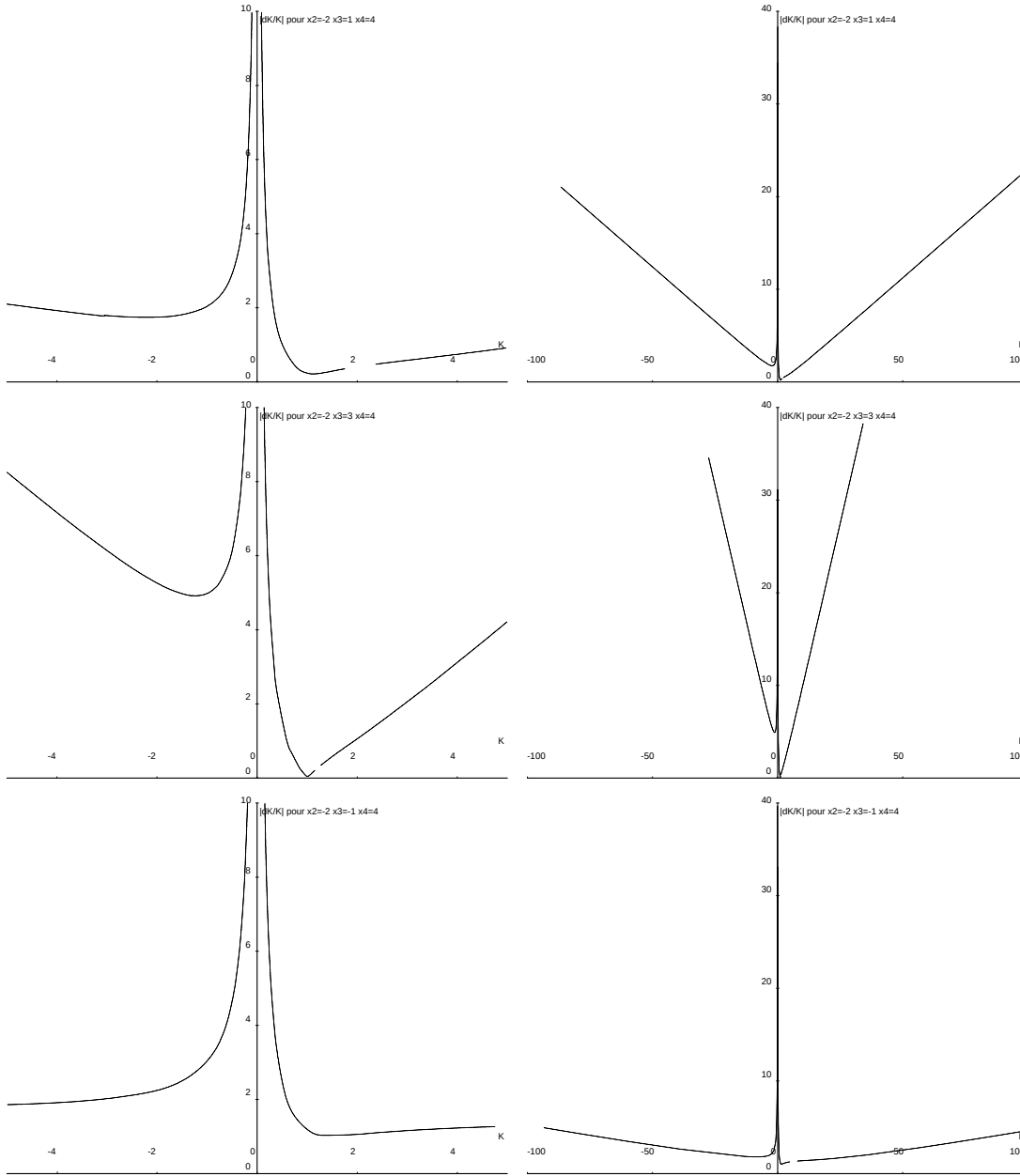


FIG. 4.10 - L'incertitude relative sur le birapport $k(x_1, x_2, x_3, x_4)$ en fonction de k , et pour trois configurations de x_2, x_3, x_4 , et pour x_1 variant.

k . Ceci a été vérifié en traçant $\left|\frac{dk}{k}\right|$, l'incertitude relative en fonction de k (voir figure 4.10). On obtient bien pour les grandes valeurs de $|k|$ une variation linéaire de $\left|\frac{dk}{k}\right|$ en fonction de k , c'est à dire une variation quadratique de $|dk|$ en fonction de k . Ceci implique que l'utilisation de l'approximation différentielle de l'incertitude donnée au paragraphe précédent engendre une sous-estimation de l'incertitude pour les grandes valeurs du birapport. Par contre, elle est très fiable pour les valeurs du

birapport situées dans l'intervalle $[0, 2]$.

4.3.2 Birapport de cinq points coplanaires

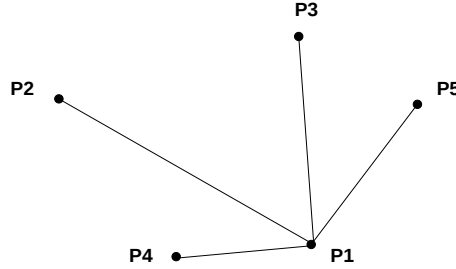


FIG. 4.11 - Le birapport associé à un quintuplet de points coplanaires

Soit un quintuplet de points coplanaires P_1, P_2, P_3, P_4, P_5 . Ces cinq points définissent un birapport :

$$k = [P_1; P_2, P_3, P_4, P_5]$$

Ce birapport correspond au birapport d'un faisceau de droites (voir figure 4.11) et il se calcule très facilement en fonction des coordonnées des points (voir la formule 1.2).

Densité de probabilité

Dans le cas de cinq points coplanaires, l'expression du birapport fait intervenir dix variables et le calcul de l'expression analytique de la fonction de répartition n'a pas été possible. Une approximation de cette fonction par tirage pseudo-aléatoire montre qu'elle a la même allure que dans le cas de 4 points alignés (voir figure 4.12). On observe une asymétrie qui ne peut être imputée qu'au générateur de nombres pseudo-aléatoire. En effet, les valeurs 0 et 1 correspondent aux mêmes configurations, modulo un changement de numérotation des points.

Comme dans le cas du birapport de quatre points alignés, si on impose des contraintes sur les points, la densité de probabilité est modifiée. La figure 4.13 montre le résultat obtenu en imposant une distance minimum de 20 entre chaque paire de points, pour des coordonnées des points prises dans l'intervalle $[-500, 500]$. On filtre ainsi surtout des configurations dégénérées (de birapports 0 et 1).

On filtre aussi les configurations dégénérées en imposant un angle minimum pour tout dièdre formé de trois des cinq points : la figure 4.14 montre le résultat obtenu pour

$$(\widehat{P_i P_j P_k}) > 0.3 \text{ degrees } \forall i, j, k \in [1..5] \text{ tels que } i \neq j \neq k.$$

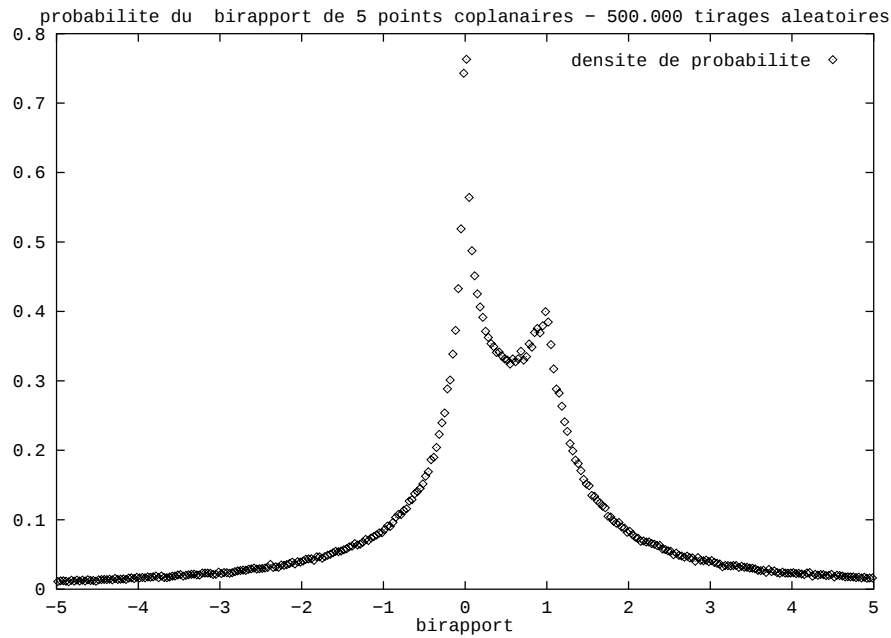


FIG. 4.12 - Approximation de la densité de probabilité du birapport de 5 points coplanaires, par tirage pseudo-aléatoire (500.000 échantillons)

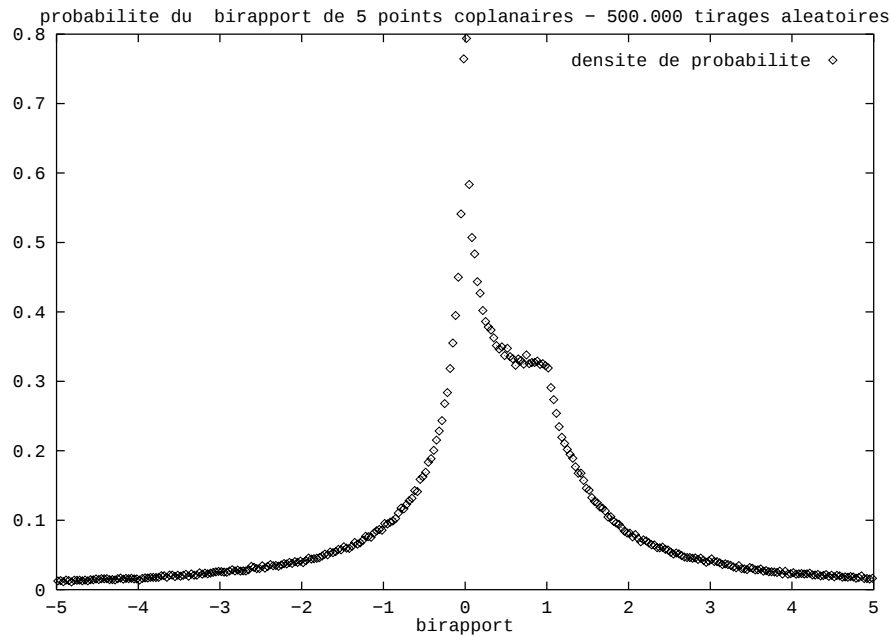


FIG. 4.13 - Approximation de la densité de probabilité du birapport de 5 points coplanaires, par tirage pseudo-aléatoire (500.000 échantillons), avec une distance minimum entre les points. Les points sont tirés dans l'intervalle $[-500, 500]$ et la distance minimum entre deux points d'un même quintuplet est 20.

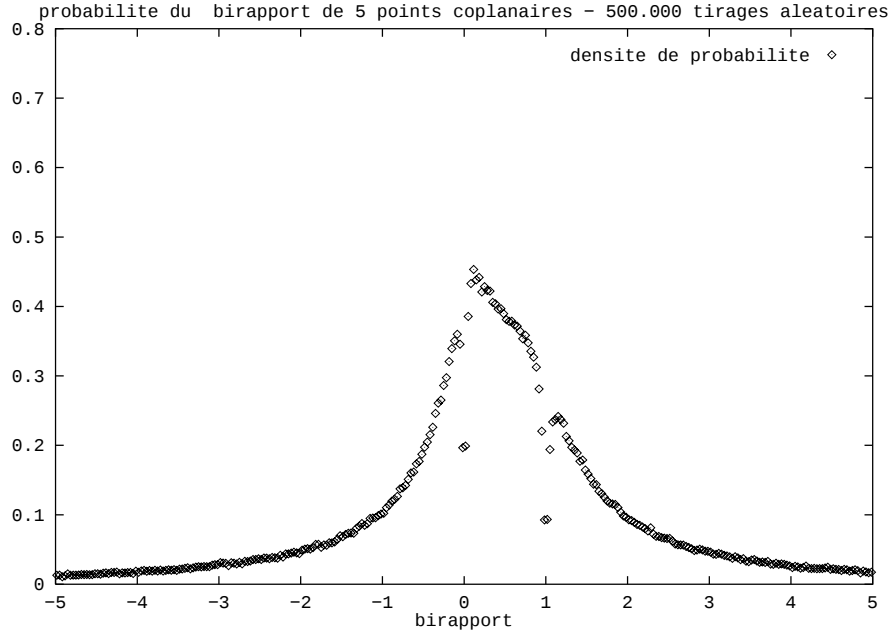


FIG. 4.14 - Approximation de la densité de probabilité du birapport de 5 points coplanaires, par tirage pseudo-aléatoire (500.000 échantillons), avec une contrainte de non alignement.

En conclusion, il semble donc légitime d'utiliser la fonction de répartition du birapport de quatre points alignés comme approximation de la fonction de répartition du birapport associé à cinq points coplanaires.

Evaluation de l'incertitude

Nous avons évalué l'incertitude sur le birapport associé à cinq points coplanaires en fonction de l'incertitude sur les coordonnées des points. Nous avons là aussi utilisé une approximation différentielle. Si on note (x_i, y_i) les coordonnées du point P_i , et dx_i et dy_i l'incertitude sur ces coordonnées, alors l'incertitude sur k est donnée par :

$$dk \simeq \sqrt{\sum_{i=1}^5 \left(\frac{\partial k}{\partial x_i} \right)^2 dx_i^2 + \left(\frac{\partial k}{\partial y_i} \right)^2 dy_i^2}$$

Comme les résultats obtenus sont similaires au cas de quatre points alignés, nous montrons les courbes d'erreur pour un seul cas de figure, celui présenté sur la figure 4.15. Seule la coordonnées x_1 varie et les autres coordonnées sont fixées.

La figure 4.16 montre l'incertitude sur k quand x_1 varie. On observe ici quatre configurations dégénérées. Elles surviennent lorsque le point P_1 est aligné successivement avec P_2 et P_5 , P_2 et P_4 , P_5 et P_3 , P_4 et P_3 .

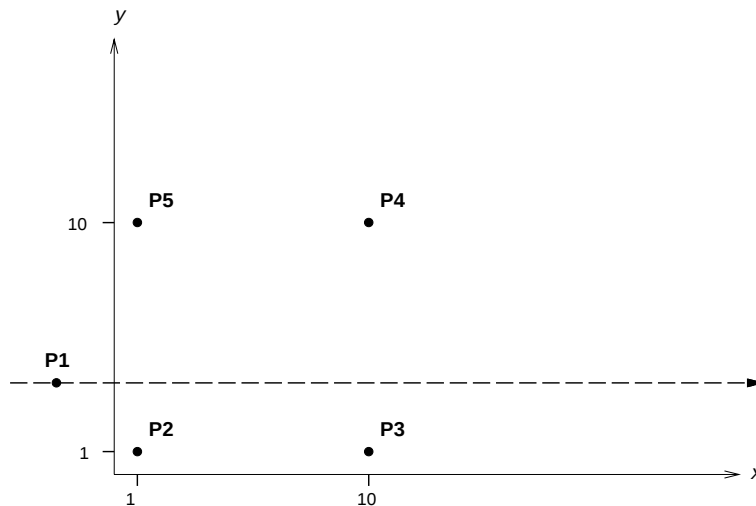


FIG. 4.15 - Disposition des points utilisés pour les calculs d'incertitude

On peut aussi observer sur les figures 4.17 et 4.18 que l'incertitude sur le birapport varie de manière quadratique par rapport au birapport, dès que celui-ci est en dehors de l'intervalle $[0, 2]$.

En conclusion, cette étude théorique nous a permis d'identifier les configurations les plus instables, en nous donnant des moyens de les filtrer (par des contraintes sur les distances entre les points, ou les angles formés par les points). Nous disposons aussi, grâce au calcul de la fonction de répartition du birapport, d'une mesure a priori du pouvoir discriminant d'un invariant, ainsi que d'une mesure de similarité entre deux invariants. Enfin, nous avons à notre disposition des fonctions approxi-mant l'incertitude obtenue sur un birapport en fonction des coordonnées des points utilisés pour le calcul du birapport et des incertitudes sur ces coordonnées.

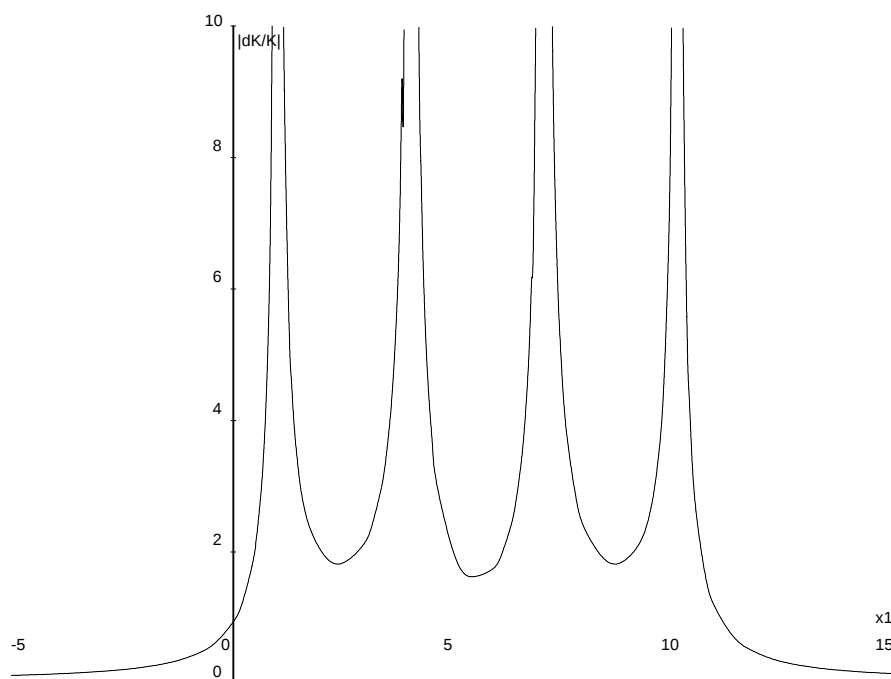


FIG. 4.16 - Incertitude relative unitaire (pour $dx = dy = 1$) sur le birapport de 5 points coplanaires en fonction de x_1 pour $y_1 = 4$, $x_2 = 1$, $y_2 = 1$, $x_3 = 10$, $y_3 = 1$, $x_4 = 10$, $y_4 = 10$, $x_5 = 1$, $y_5 = 10$.

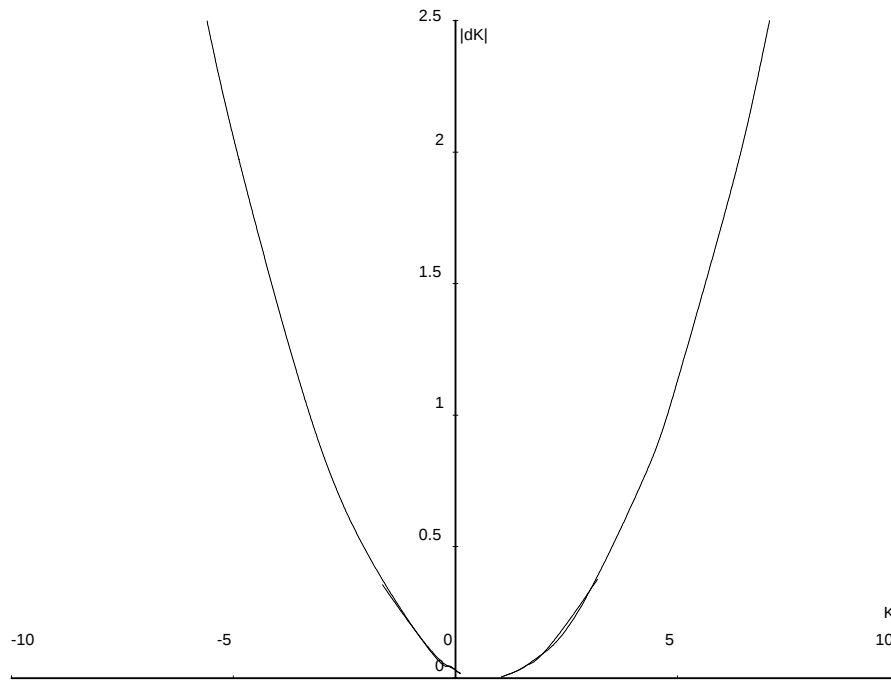


FIG. 4.17 - Incertitude sur le birapport de 5 points coplanaires en fonction du birapport pour $y_1 = 4$, $x_2 = 1$, $y_2 = 1$, $x_3 = 10$, $y_3 = 1$, $x_4 = 10$, $y_4 = 10$, $x_5 = 1$, $y_5 = 10$ et x_1 variant, et pour $dx = dy = 0.1$.

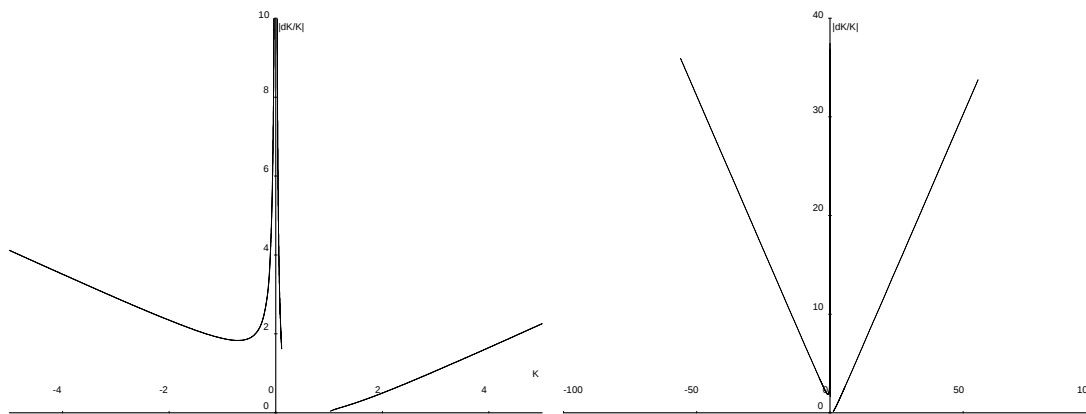


FIG. 4.18 - Incertitude relative unitaire sur k en fonction de k pour $y_1 = 4$, $x_2 = 1$, $y_2 = 1$, $x_3 = 10$, $y_3 = 1$, $x_4 = 10$, $y_4 = 10$, $x_5 = 1$, $y_5 = 10$ et x_1 variant

4.4 Les invariants de cinq points coplanaires

Etant donnés cinq points non trois à trois alignés de $\mathbb{P}^2$, il existe deux invariants projectifs indépendants de ces cinq points, comme nous l'avons vu au paragraphe 1.4.1. Nous proposons dans ce paragraphe différents moyens de calcul d'un couple d'invariants projectifs indépendants.

4.4.1 Invariants de cinq points ordonnés

Si on dispose de cinq points ordonnés P_1, P_2, P_3, P_4, P_5 , on peut choisir comme invariants associés au quintuplet les coordonnées projectives de P_5 par rapport aux points P_1, P_2, P_3, P_4 .

4.4.2 Invariants de cinq points non ordonnés

Dans le cas qui nous préoccupe, on recherche à comparer des *ensembles de cinq points* et non des *quintuplets*. L'invariant projectif le plus simple calculable à partir de cinq points est le birapport d'un des faisceaux de droites définis par ces cinq points. (voir figure 4.11). Il y a cinq façons de choisir le point sommet du faisceau ; le point choisi, il reste 4! façon de choisir l'ordre des quatre droites du faisceau. Il y a donc a priori 120 birapports calculables à partir de cinq points coplanaires donnés, dont 2 seulement sont indépendants. Vouloir les comparer tous serait très coûteux. On cherche donc des invariants pouvant se calculer indépendamment de l'ordre des points.

Invariants d'un faisceau de droites non ordonnées

Etant donné un faisceau de quatre droites, les 4! possibilités pour choisir l'ordre des droites ne génèrent en fait que six valeurs de birapport différentes. Si $[l_1, l_2, l_3, l_4] = k$ alors le birapport d'une permutation quelconque de l_1, l_2, l_3, l_4 est l'une des six valeurs :

$$\begin{array}{ccc} k & 1-k & 1-\frac{1}{k} \\ \frac{1}{k} & \frac{1}{1-k} & \frac{k}{k-1} \end{array} \quad (4.14)$$

Ceci nous permet de construire un invariant projectif indépendant de l'ordre des droites du faisceau :

- Une première solution est de calculer une fonction symétrique de ces six valeurs. Les plus élémentaires sont la somme, qui prend la valeur constante 0 et le produit, qui prend la valeur constante 1. On peut donc utiliser la somme

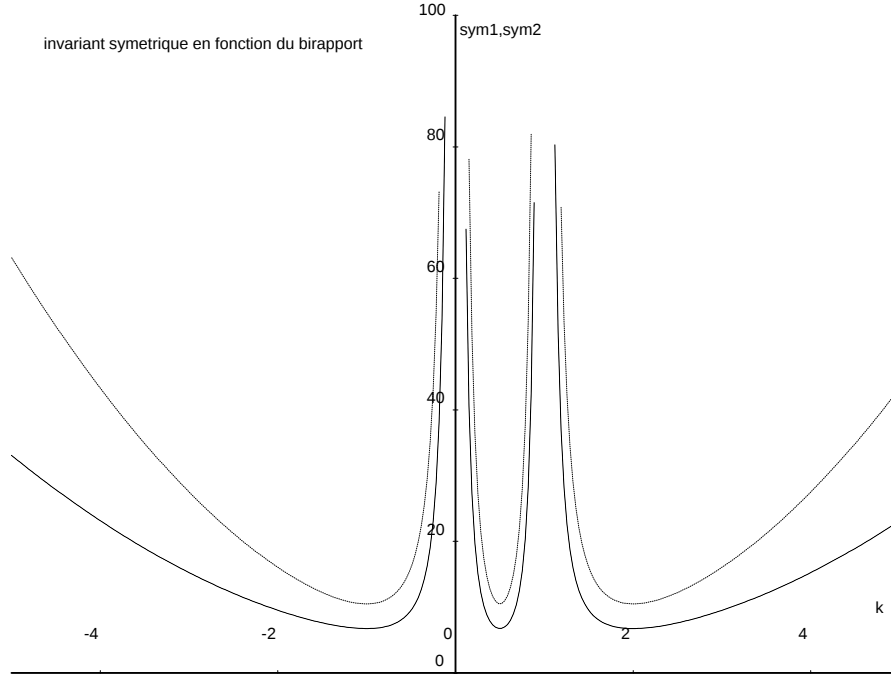


FIG. 4.19 - Deux fonctions indépendantes de l'ordre des droites d'un faisceau. La courbe inférieure correspond à la fonction proposée par Maybank. La courbe supérieure correspond à la somme des carrés des 6 valeurs.

du carré des valeurs :

$$k^2 + \frac{1}{k^2} + (1-k)^2 + \frac{1}{(1-k)^2} + \left(1 - \frac{1}{k}\right)^2 + \left(\frac{k}{k-1}\right)^2 \quad (4.15)$$

ou la fonction rationnelle donnée par Maybank [May 91] :

$$j(k) = \frac{(k^2 - k + 1)^3}{k^2(k-1)^2} \quad (4.16)$$

$j(k)$ prend la même valeur pour toutes les fonctions de k données en 4.14 et $j(k) = j(k')$ implique que k' est l'une de ces six fonctions de k .

Les fonctions symétriques 4.15 et 4.16 sont en fait très proches (voir figure 4.19). On peut aussi utiliser les fonctions symétriques proposées par Lenz et Meer [Len 92].

- Une deuxième solution s'appuie sur la remarque suivante : les différentes valeurs $k, 1-k, \dots$ occupent des intervalles disjoints qui forment une partition de $\mathbb{R}$, comme le montre la figure 4.20.

On peut ainsi ramener la valeur calculée dans un intervalle fixé : on prend comme invariant celle parmi les six valeurs qui appartient à l'intervalle voulu.

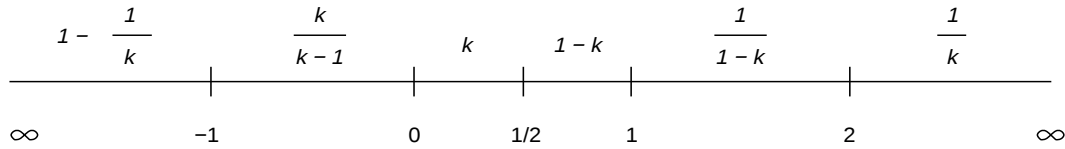


FIG. 4.20 - Les six valeurs du birapport occupent des intervalles fixes formant une partition de $\mathbb{R}$

Ceci revient à modifier l'ordre des droites du faisceau pour faire le calcul du birapport.

Par exemple supposons qu'on choisisse de se ramener à l'intervalle $[0, \frac{1}{2}]$. Si la valeur k du birapport du faisceau appartient à l'intervalle $[2, \infty]$, on prend comme invariant $\frac{1}{k}$. Si k appartient à $[\frac{1}{2}, 1]$, on prend comme invariant $1 - k$, et ainsi de suite.

On peut donc associer à un faisceau de quatre droites un invariant projectif indépendant de l'ordre des droites. Cependant à partir de cinq points, il reste cinq possibilités pour choisir le point sommet du faisceau. On dispose donc de cinq invariants a_1, a_2, a_3, a_4, a_5 dont deux seulement sont indépendants.

Génération de 2 invariants indépendants à partir de 5 invariants

Deux possibilités sont envisageables pour générer deux invariants indépendants à partir des cinq invariants a_1, a_2, a_3, a_4, a_5 :

- On peut choisir deux invariants distincts parmi les cinq, par exemple le minimum et le maximum des cinq valeurs, ou les deux valeurs maximales. Cependant, on ne tient alors pas compte des trois autres valeurs calculées.
- Les polynômes symétriques de premier et de deuxième ordre en les cinq invariants fournissent deux invariants prenant en compte les cinq valeurs :

$$I_1 = \sum_{i=1}^5 a_i$$

$$I_2 = \sum_{i \neq j} a_i a_j$$

Il y a une relation bijective entre le couple (I_1, I_2) et les cinq valeurs a_1, a_2, a_3, a_4, a_5 .

En résumé, chaque groupe de cinq points non trois à trois alignés peut être caractérisé par deux invariants projectifs indépendants de l'ordre des cinq points. On dispose de plusieurs choix pour calculer ces invariants.

4.4.3 Invariants de cinq points : cas de la conservation de l'enveloppe convexe

On peut proposer une dernière caractérisation projective si on fait l'hypothèse de conservation de l'enveloppe convexe : les points de l'enveloppe convexe d'un groupe modèle se projettent sur les points de l'enveloppe convexe du groupe image. Cette hypothèse est valide dans la plupart des cas réels.

On suppose aussi que l'enveloppe convexe peut être orientée, et que cette orientation est conservée entre la scène et les images. On suppose donc que les cinq points sont toujours observés depuis le “même côté” du plan les contenant. Cette hypothèse n'est pas toujours vérifiée pour des objets plans et il existe des polyèdres simples pour lesquels les points peuvent être visibles depuis chaque côté du plan (voir l'exemple d'une pyramide de base pentagonale sur la figure 4.21).

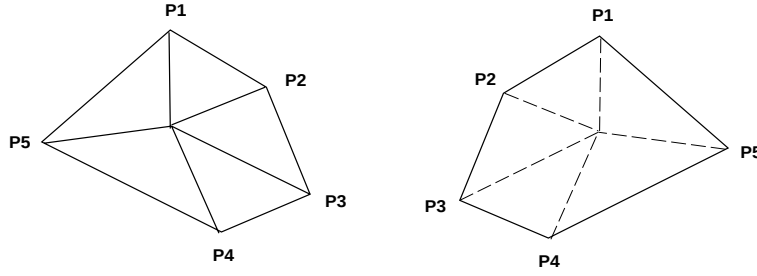


FIG. 4.21 - Les points $P_1 P_2 P_3 P_4 P_5$ peuvent être visibles depuis les deux côtés du plan les contenant.

Un groupe de cinq points est alors caractérisé par trois invariants : un premier invariant entier est le nombre de points de l'enveloppe convexe du groupe de cinq points (3, 4 ou 5).

On se donne tout d'abord un sens de parcours (arbitraire) de l'enveloppe convexe. Suivant le nombre de points de l'enveloppe convexe, les invariants projectifs sont calculés comme suit (voir figure 4.22) :

3 points

On numérote les points : P_4 et P_5 sont les points intérieurs à l'enveloppe convexe. P_1 est le point isolé par rapport à la droite (P_4, P_5) . P_2 et P_3 sont les points de l'enveloppe convexe tels que P_1, P_2, P_3 soit l'enveloppe convexe orientée. P_4 et P_5 sont distingués par les angles α_4 et α_5 , tels que $\alpha_4 \leq \alpha_5$, c'est-à-dire de sorte à conserver l'ordre des droites du faisceau issu de P_1 . Les deux invariants choisis sont alors :

$$k1 = [P_1; P_2, P_3, P_4, P_5]$$

$$k2 = [P_2; P_3, P_4, P_5, P_1]$$

4 points

On numérote les points : P_5 est le point intérieur. On définit P_1, P_2, P_3, P_4 ,

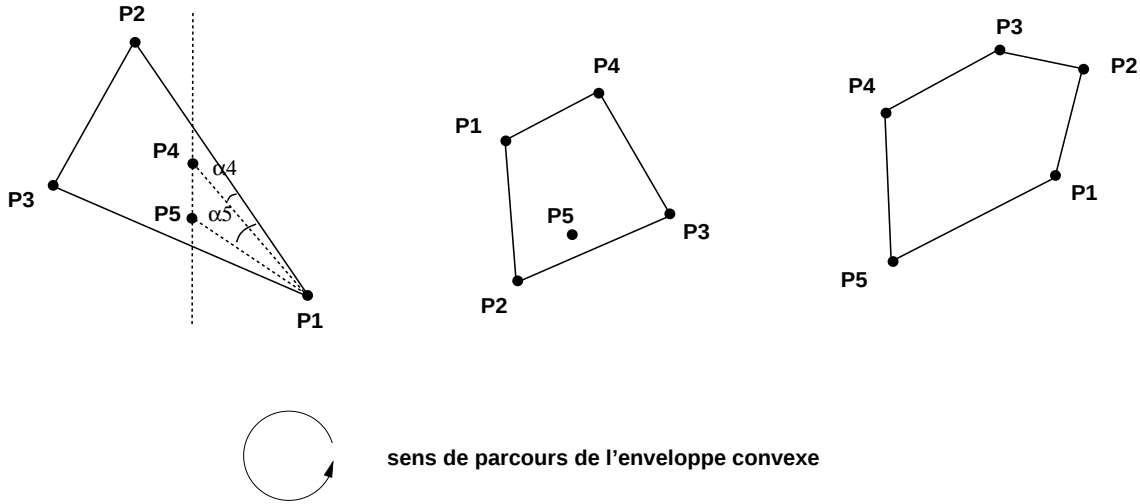


FIG. 4.22 - Numérotation des points dans les trois cas de figure : enveloppe convexe à 3, 4 ou 5 points

comme la permutation circulaire de l'enveloppe convexe orientée générant la valeur maximum du birapport $k_1 = [P_1; P_2, P_3, P_4, P_5]$. Les deux invariants projectifs sont alors :

$$k1 = [P_1; P_2, P_3, P_4, P_5]$$

$$k2 = [P_2; P_3, P_4, P_1, P_5]$$

5 points

On numérote les points : on définit P_1, P_2, P_3, P_4, P_5 , comme la permutation circulaire de l'enveloppe convexe orientée générant la valeur maximum du birapport $k_1 = [P_1; P_2, P_3, P_4, P_5]$. Les deux invariants projectifs sont alors :

$$k1 = [P_1; P_2, P_3, P_4, P_5]$$

$$k2 = [P_2; P_3, P_4, P_5, P_1]$$

On peut remarquer que $[P_1; P_2, P_3, P_4, P_5] = [P_1; P_5, P_4, P_3, P_2]$. La restriction concernant l'orientation de l'enveloppe convexe peut donc être levée pour les configurations dont l'enveloppe convexe comprend quatre ou cinq points : on obtiendra les mêmes invariants quelque soit le demi espace depuis lequel on visualise les cinq points. Cette restriction doit être conservée pour les configurations dont l'enveloppe convexe comprend trois points, mais les cas réels où de telles configurations peuvent être visibles depuis les deux demi espaces concernent essentiellement les objets plans ou transparents.

En conclusion, les hypothèses utilisées pour le calcul de ce dernier type d'invariant sont valides en pratique et elles nous permettent, comme dans [Rao 92], de diminuer la combinatoire liée au non ordonnancement des points.

4.5 Simulations

4.5.1 Données utilisées pour les test

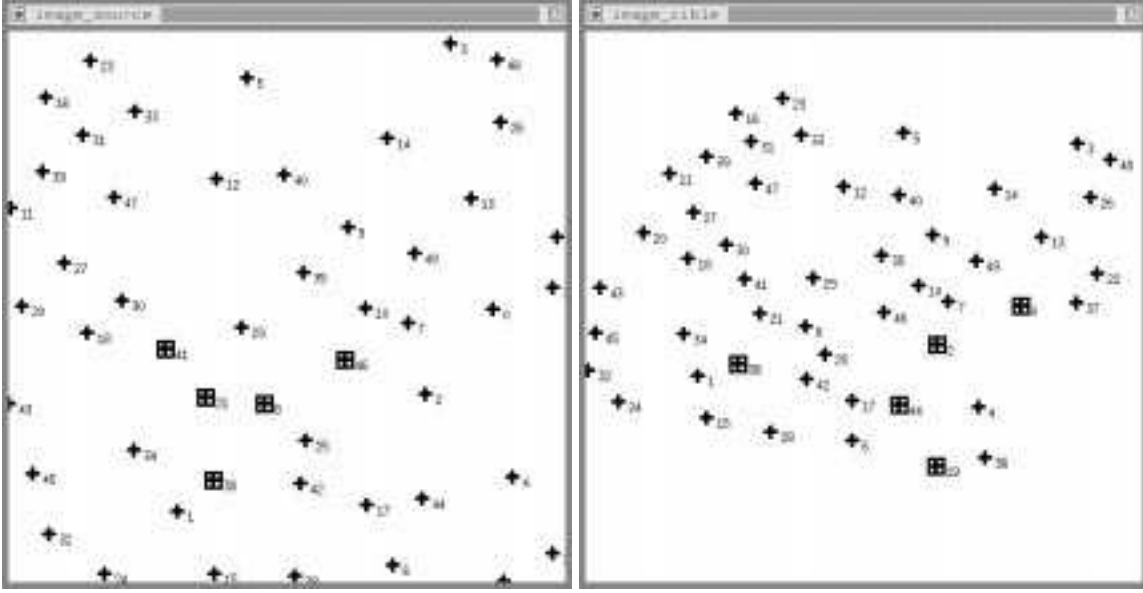


FIG. 4.23 - Les deux images utilisées pour les tests : on peut voir dans l'image 1 (à gauche) l'un des 12 quintuplets sélectionnés comme modèles et dans l'image 2 (à droite), un groupe sélectionné comme candidat pour ce modèle.

Quantifier la qualité des invariants ne peut se faire que pour :

- une base de modèles donnée,
- une incertitude donnée sur les coordonnées des points de l'image,
- une image donnée, c'est à dire une transformation projective donnée.

Nous utilisons donc la configuration de test suivante.

Nous construisons deux images (voir figure 4.23). L'image 1 est de taille 512×512 *pixel* et elle contient 50 points, répartis aléatoirement selon une distribution uniforme. On impose une distance supérieure à 40 *pixels* entre les coordonnées de chaque paire de points.

L'image 2 est générée en appliquant une transformation homographique à l'image 1, puis en rajoutant sur chaque point une erreur aléatoire gaussienne de moyenne nulle. Les points de l'image 2 sont enfin discrétisés au plus proche voisin.

Nous construisons notre base de modèles à partir de l'image 1, en sélectionnant des configurations qui vérifient une contrainte de non alignement :

$$\sin(\widehat{P_i P_j P_k}) > 0.3 \quad \forall i \neq j \neq k \in [1..5]$$

Ceci nous permet d'éviter de prendre comme modèles des configurations trop instables, comme l'a montré l'étude théorique. Nous trouvons ainsi 12 configurations qui forment notre base de modèles. Les invariants projectifs caractérisant ces 12 configurations sont calculés et ces dernières sont stockées dans une table indexée par les invariants.

Nous recherchons ensuite dans l'image 2 les groupes de cinq points susceptibles d'être les transformés d'un groupe modèle. Pour cela, nous considérons tous les groupes de cinq points de l'image 2 (ils sont 1.500.000) et nous calculons les invariants projectifs associés à chaque groupe, ainsi que la marge d'erreur sur ces invariants. Nous cherchons alors dans la table s'il existe un modèle caractérisé par des invariants similaires, c'est à dire dont la valeur est compatible avec la marge d'erreur. Si c'est le cas, le groupe de l'image 2 est retenu comme candidat pour le groupe modèle.

Pour chaque groupe modèle, la qualité de la reconnaissance est déterminée, d'une part par la sélection ou le rejet du véritable candidat (correspondant au projeté du modèle), et d'autre part par le taux de discrimination, c'est à dire le rapport entre le nombre total de groupes considérés dans l'image 2 et le nombre de faux candidats retenus.

4.5.2 Comparaison de la stabilité

Le taux de discrimination dépend de plusieurs facteurs. Les premiers sont les qualités intrinsèques des invariants : leur stabilité et leur densité de probabilité.

Si une configuration est caractérisée par un invariant instable, la configuration constituant le vrai candidat sera caractérisée par un invariant d'une valeur éloignée. Dans le meilleur des cas, toutes les configurations dont les invariants sont plus proches de ceux du modèle seront des faux candidats. Le nombre de ces faux candidats dépend de la densité de probabilité de l'invariant autour de la valeur associée au groupe modèle. Un invariant très instable mais avec une densité très faible, peut donner le même nombre de faux candidats qu'un invariant très stable mais d'une valeur correspondant à une forte densité.

Le troisième facteur qui intervient dans le taux de discrimination est l'évaluation de la tolérance sur l'invariant. Si elle est sous-estimée, le vrai candidat risque d'être rejeté. Si elle est sur-estimée, un nombre superflu de faux candidats sera retenu.

Pour pouvoir comparer la qualité des invariants indépendamment d'un choix pour l'évaluation de la tolérance, avons fait l'expérimentation suivante.

Nous nous plaçons dans l'espace des invariants. Comme chaque configuration de cinq points est caractérisée par deux invariants, cet espace est un plan. L'axe horizontal correspond au premier invariant k et l'axe vertical au deuxième invariant l . Un point $P(x, y)$ de cet espace représente une configuration caractérisée par les invariants $k = x$ et $l = y$. Pour chacune des 12 configurations modèles de l'image 1, nous avons calculé les invariants qui lui sont associés, ainsi que ceux associés à son

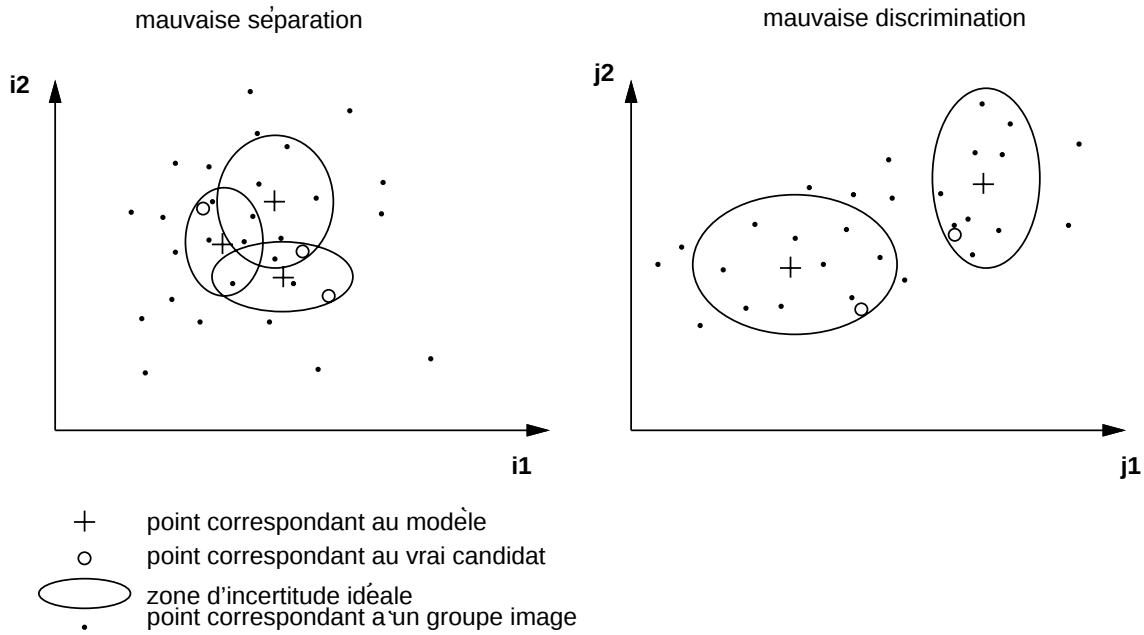


FIG. 4.24 - Visualisation des qualités de séparation et de discrimination dans l'espace des invariants

projeté dans l'image 2 (le vrai candidat).

Nous traçons dans l'espace des invariants le point correspondant au modèle, $P_1(x_1, y_1)$ et le point correspondant au vrai candidat $P_2(x_2, y_2)$. Soit l'ellipse de centre P_1 , contenant P_2 et admettant $(P_1 P_2)$ pour diagonale, c'est à dire :

$$\frac{\text{axe horizontal}}{\text{axe vertical}} = \frac{|x_2 - x_1|}{|y_2 - y_1|}$$

Cette ellipse peut être considérée comme un intervalle d'erreur "idéal", c'est à dire comme le plus petit intervalle d'erreur centré sur le modèle contenant le vrai candidat.

Nous avons tracé ces intervalles d'erreur pour les différentes paires d'invariants proposés dans le paragraphe 4.4. Ceci nous permet de comparer sur les mêmes données la stabilité des différents invariants projectifs, ainsi que leur pouvoir séparateur. En effet la proximité ou l'intersection de deux intervalles d'erreur signifie que les invariants utilisés ne permettent pas de distinguer les deux modèles.

Les résultats sont montrés sur les figures 4.25 à 4.32. Pour chaque type d'invariant la figure de gauche présente le résultat obtenu sans ajout de bruit : les erreurs sont dues uniquement à la discrétisation. Dans la figure de droite, un bruit gaussien de moyenne nulle et d'écart type 1 *pixel* a été appliqué aux points de l'image 2 avant la discrétisation. Ces résultats mettent en évidence la fragilité engendrée par l'utilisation des fonctions et des polynômes symétriques. En effet, ces derniers opèrent un moyennage sur les valeurs des invariants, ce qui a pour effet d'augmenter

la densité de probabilité autour de chaque valeur, alors que les incertitudes sont cumulées. On obtient donc une mauvaise séparation des modèles, ainsi qu'une très grande sensibilité au bruit de mesure. Par contre les invariants utilisant la conservation de l'enveloppe convexe aboutissent à une bonne séparation des différents modèles, et sont relativement stables.

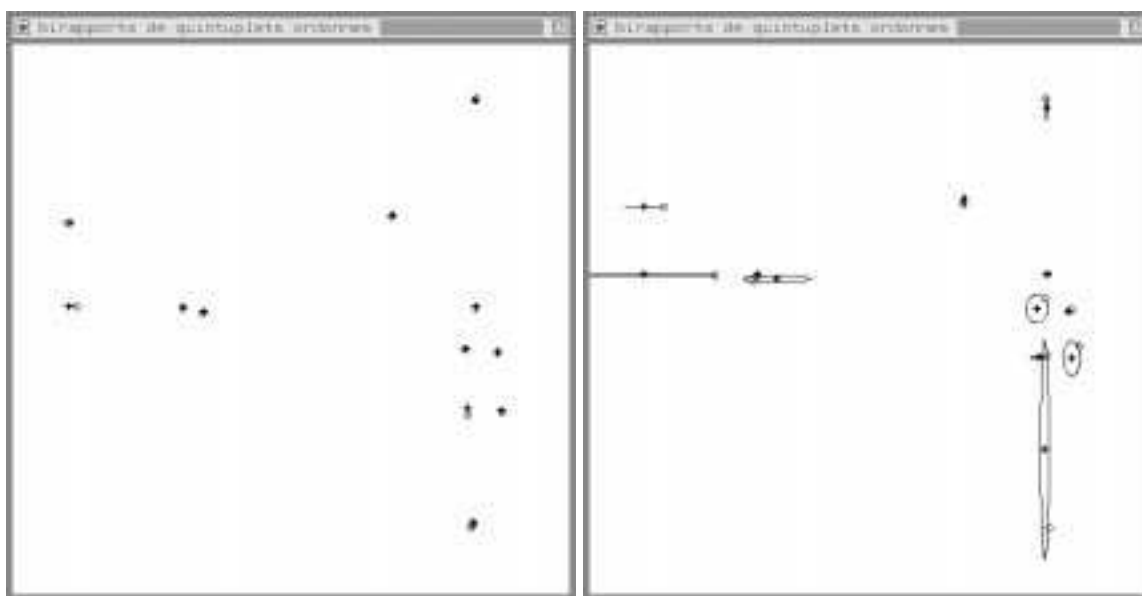


FIG. 4.25 - invariants de cinq points ordonnés : deux birapports

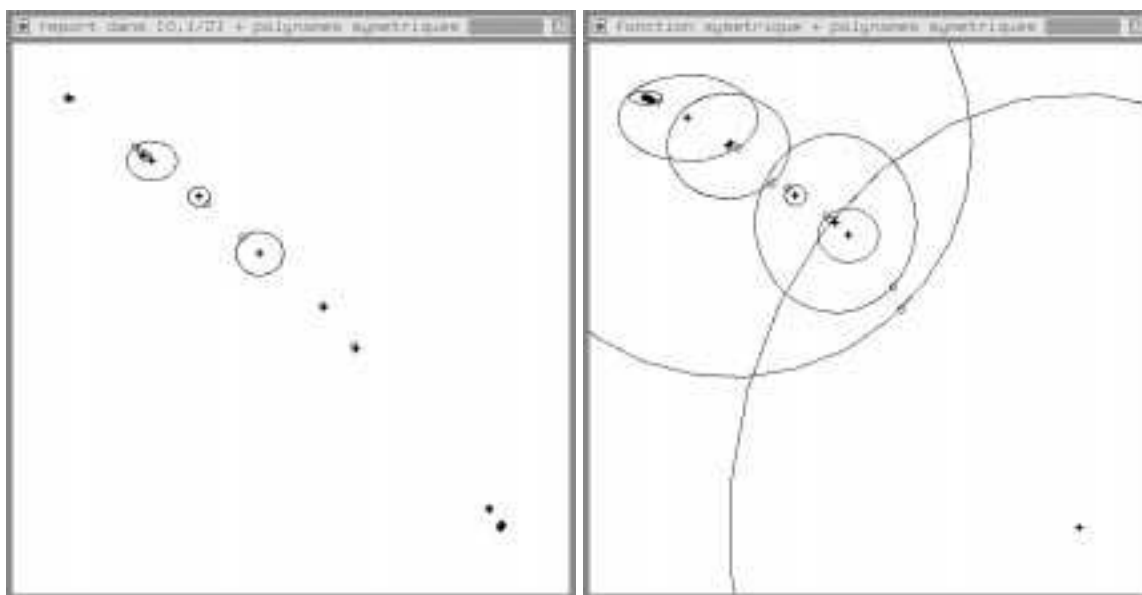


FIG. 4.26 - invariants calculés en utilisant la fonction de symétrisation puis les polynômes symétriques

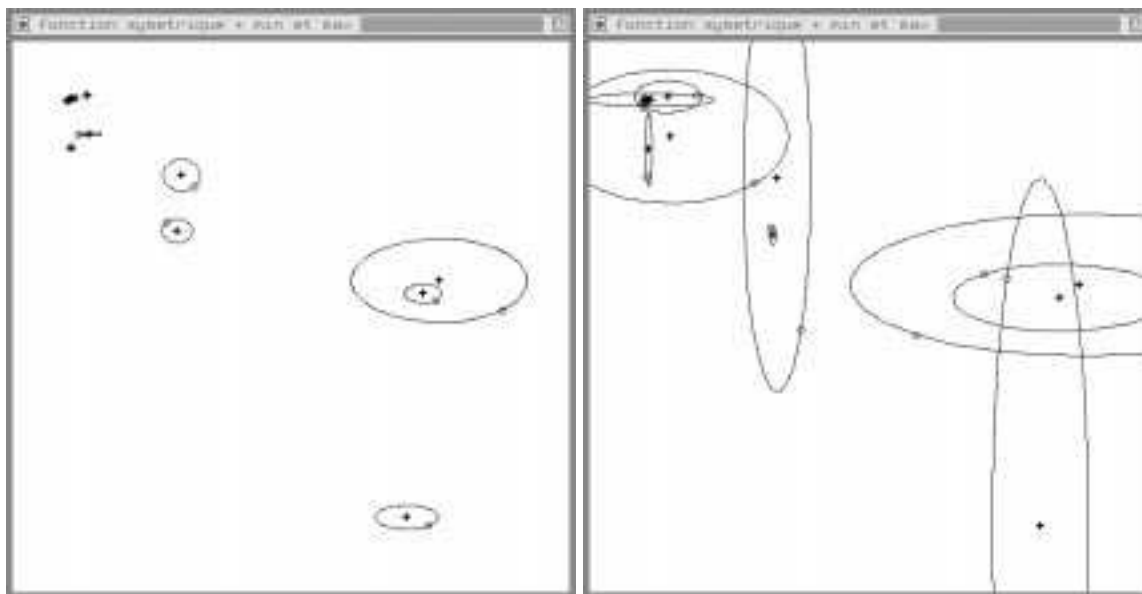


FIG. 4.27 - invariants calculés en utilisant la fonction symétrique puis minimum et maximum des 5 valeurs

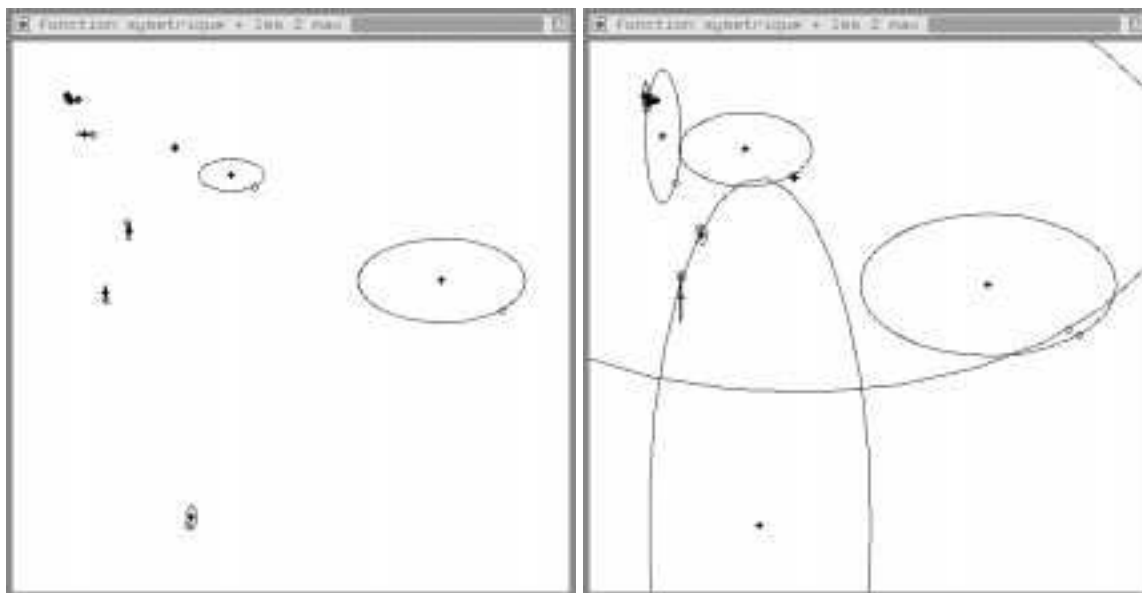


FIG. 4.28 - invariants calculés en utilisant la fonction symétrique puis les deux plus grandes des 5 valeurs

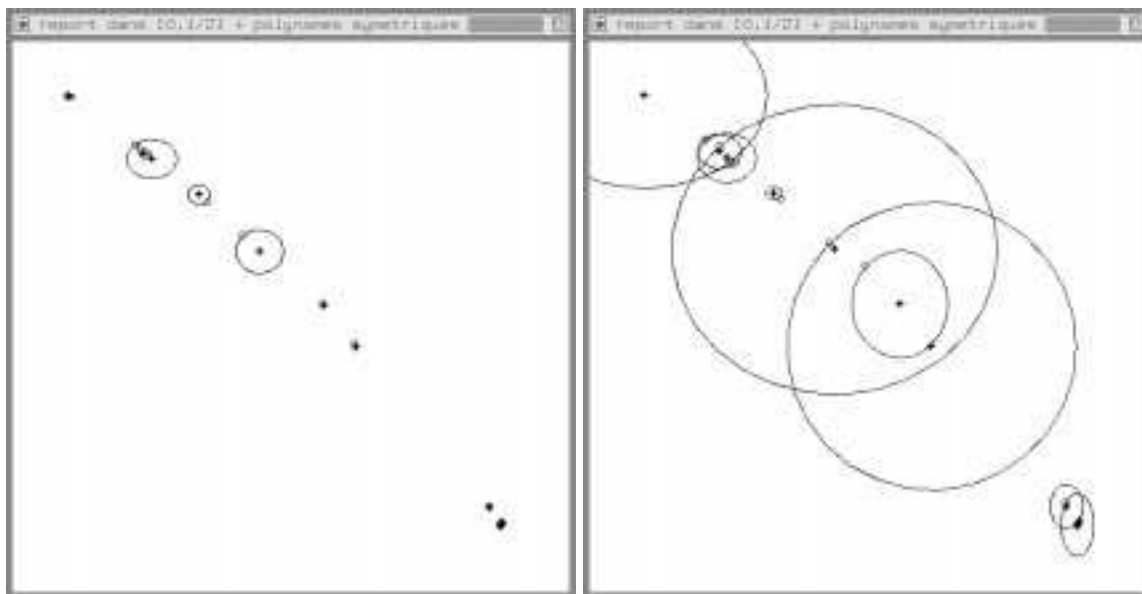


FIG. 4.29 - invariants calculés en reportant les birapport dans l'intervalle $[0, 1/2]$ puis en utilisant les polynômes symétriques

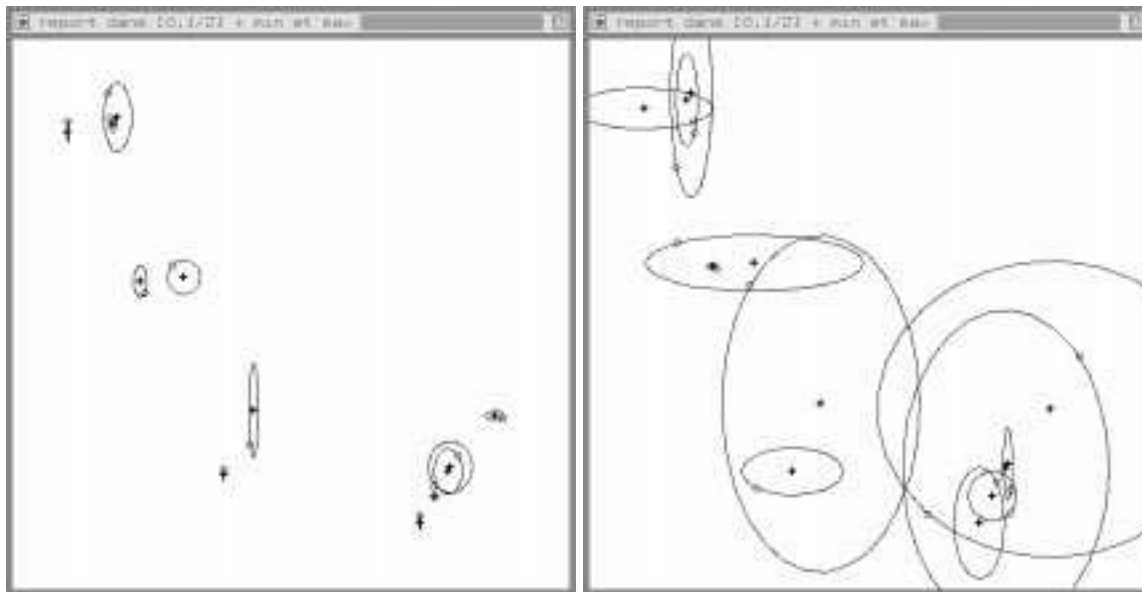


FIG. 4.30 - invariants calculés en reportant les birapport dans l'intervalle $[0, 1/2]$ puis en prenant le minimum et le maximum des cinq valeurs

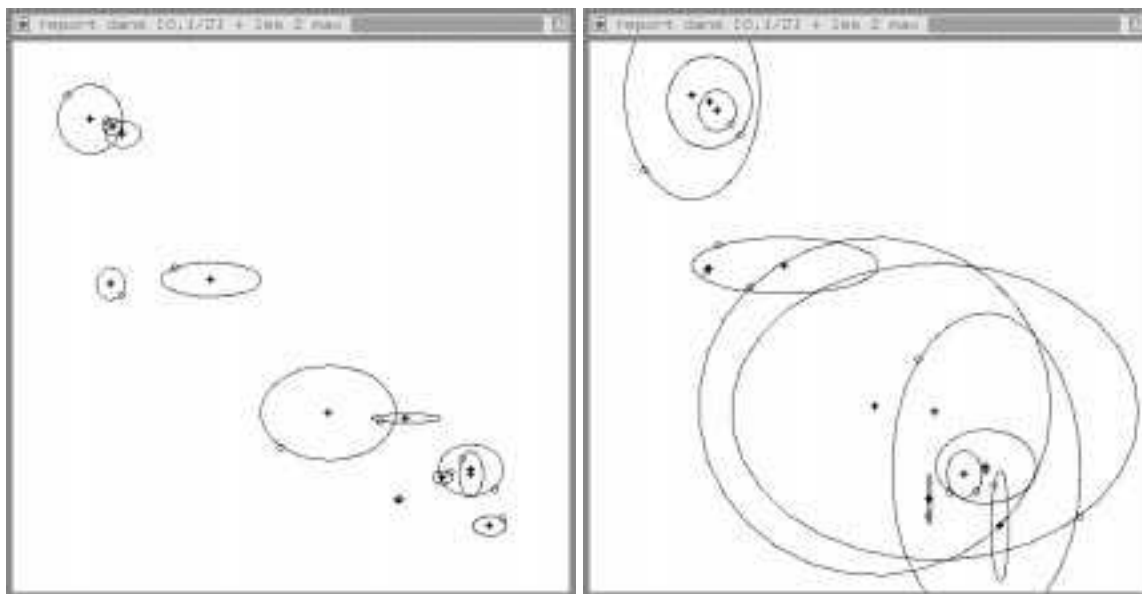


FIG. 4.31 - invariants calculés en reportant les birapport dans l'intervalle $[0, 1/2]$ puis prenant les deux plus grandes des 5 valeurs

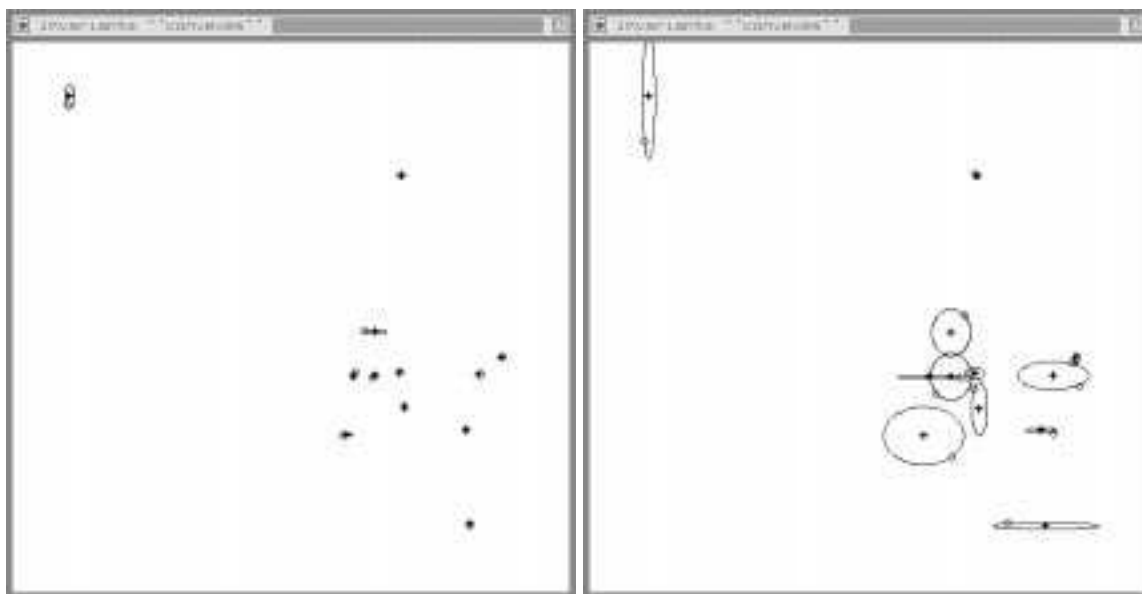


FIG. 4.32 - invariants calculés en supposant qu'il y a conservation de l'enveloppe convexe

no	DONNEES				RESULTATS		
	invariants utilisés		bruit	tol.	trouv.	faux cand.	fact. discr.
1	dans $[0,1/2]$	min et max	discr.	3,5 %	9	50.000	< 30
2	convexes	2 premiers	discr.	7 %	12	5.500	< 300
3	convexes	5	discr.	8 %	12	1.000	1.500
4	convexes	5	discr.	tol1	12	1.000	1.500
5	convexes	5	discr.	tol2	12	400	5.000
6			discr. + 0.5		12	1.000	1.500
7			discr. + 1		12	1.800	800

TAB. 4.1 - Taux de discrimination

4.5.3 Comparaison des performances

Dans ce paragraphe, nous présentons les performances obtenues lors des tests de reconnaissance, tels qu'ils ont été décrits plus haut.

Des premiers essais ont été effectués, sans bruitez les données (seules les erreurs d'arrondi interviennent). Comme attendu, les résultats obtenus sont parfaits : le vrai candidat de chaque configuration modèle est reconnu, et aucun faux candidat n'est sélectionné.

Dans les tests suivants les points de l'image 2 sont bruités. Le tableau 4.1 présente pour chaque test effectué les données du test : les invariants utilisés, le bruit rajouté sur les points (soit uniquement le bruit de discrétisation, soit discrétisation ainsi que bruit gaussien d'amplitude exprimée en *pixels*), et la tolérance acceptée sur les invariants. Les résultats de chaque test sont présentés en donnant le nombre de vrais candidats sélectionnés sur l'ensemble des 12 modèles, le nombre moyen de faux candidats par modèle, et enfin le taux de discrimination moyen par modèle, c'est à dire :

$$\begin{aligned}
 \text{taux discr.} &= \frac{1}{12} \sum_{i=1}^{i=12} \frac{\text{nombre total de candidats}}{\text{nombre de faux candidats pour le modele } i} \\
 &= \frac{1.500.000}{12} \sum_{i=1}^{i=12} \frac{1}{\text{nombre de faux candidats pour le modele } i}
 \end{aligned}$$

Ce taux de discrimination représente le nombre moyen de configurations rejetées pour une configuration retenue. Il donne donc une évaluation des capacités de filtrage des invariants.

Dans un premier temps, seule l'erreur de discrétisation est appliquée sur l'image 2, et la tolérance sur les invariants est fixée en pourcentage de la valeur des invariants. Ce pourcentage est ajusté a priori : on calcule pour chaque modèle les invariants associés au vrai candidat, et on choisit la tolérance minimum telle que tous les vrais candidats soient sélectionnés au moment de la reconnaissance. Ce calcul n'est bien sûr pas possible dans un cas réel, puisqu'on ne connaît justement pas les vrais

candidats! Cependant, ceci permet d'évaluer les performances dans le meilleur des cas, c'est-à-dire pour un nombre minimum de faux candidats.

Les résultats obtenus avec les invariants autres que ceux utilisant la convexité confirment ce que nous avons vu au paragraphe précédent : les invariants sont très instables et même avec des tolérances très larges sur les invariants, tous les vrais candidats ne sont pas retenus. Nous donnons pour comparaison le résultat pour les deux invariants obtenus en ramenant les birapports de faisceau dans l'intervalle $[0, 1/2]$, puis en prenant le minimum et la maximum des cinq valeurs (test 1). Tous les autres tests présentés ont été effectués avec les invariants utilisant l'hypothèse de la conservation de l'enveloppe convexe.

Les tests 2 et 3 montrent les résultats obtenus en utilisant soit deux invariants projectifs indépendants, soit cinq invariants redondants. Les cinq invariants sont les invariants projectifs de faisceaux, obtenus en prenant successivement chaque point comme sommet. En théorie, les cinq invariants ne rajoutent rien à la caractérisation, puisque seulement deux d'entre eux sont indépendants. Cependant, en présence des erreurs, les performances obtenues avec cinq invariants sont nettement meilleures qu'avec deux invariants.

Dans le test 4, la tolérance n'est plus fixée a priori. Comme dans un cas réel, elle a été estimée automatiquement : pour chaque configuration de l'image 2, une tolérance sur les invariants associés à cette configuration est calculée en utilisant l'approximation différentielle donnée au paragraphe 4.3.2. On recherche ensuite parmi les modèles une configuration dont les invariants appartiennent à l'intervalle de tolérance, c'est-à-dire vérifiant la condition suivante :

$$|invariant_{candidat} - invariant_{modele}| < \sqrt{\sum_{i=1}^5 \left(\frac{\partial k}{\partial x_i} \right)^2 + \left(\frac{\partial k}{\partial y_i} \right)^2} dx_i \quad \text{où } dx_i = 0.5$$

L'incertitude de 0.5 *pixels* sur les coordonnées des points est une évaluation de l'incertitude générée par la discrétisation au plus proche voisin.

On aboutit à des performances similaires à celles obtenues avec une tolérance fixe, mais qui doit être ajustée à la main. Ceci valide notre approximation de l'incertitude.

Le test 5 utilise un raffinement de la tolérance : la tolérance sur chaque invariant est plus large, mais on impose un seuil sur la moyenne des tolérances calculées sur les cinq invariants :

$$|invariant_{candidat} - invariant_{modele}| < \sqrt{\sum_{i=1}^5 \left(\frac{\partial k}{\partial x_i} \right)^2 + \left(\frac{\partial k}{\partial y_i} \right)^2} dx_i$$

où $dx_i = 0.75 + \text{amplitude du bruit}$ et

$$\text{Moyenne} \left(\frac{|invariant_{candidat} - invariant_{modele}|}{|invariant_{modele}|} \right) < S\% \quad \text{où } S \text{ est ajusté}$$

Les tests 6 et 7 montrent l'effet d'un bruit de mesure gaussien de moyenne nulle et d'écart type 0.5 pixel puis 1 pixel . Le taux de discrimination décroît rapidement, montrant la sensibilité des invariants aux bruits de mesure.

4.6 Conclusion

L'objectif de ce chapitre était l'étude de la stabilité des invariants projectifs en présence de bruit dans les images, ainsi que l'évaluation de leur potentiel comme indices de reconnaissance.

Dans une première approche théorique sur le birapport, nous dérivons la densité de probabilité et la fonction de répartition du birapport de quatre points alignés. Ce résultat nous permet d'évaluer le pouvoir discriminant d'un birapport en fonction de sa valeur. On observe que les birapports de valeur proche de 0 ou 1 ont une forte probabilité, et sont donc a priori moins discriminants. Ces valeurs correspondent aux valeurs dites dégénérées (points confondus dans le cas de quatre points alignés ou points alignés dans le cas de cinq points coplanaires); elles peuvent être filtrées en imposant des contraintes sur la distance et l'alignement des points. Par ailleurs, nous proposons une mesure de similarité entre birapports, basée sur la fonction de répartition, qui remplace avantageusement la distance euclidienne, mesure mal adaptée pour une notion projective telle que le birapport.

Nous calculons ensuite par différentiation une approximation de l'incertitude sur le birapport en présence d'une incertitude sur les points. La valeur du birapport elle-même ne suffit pas pour déterminer son incertitude. Mais si nous connaissons les coordonnées des points utilisés pour le calcul du birapport, ainsi que les incertitudes sur ces coordonnées, nous pouvons évaluer l'incertitude sur le birapport de quatre points alignés ou de cinq points coplanaires. On observe cependant que l'incertitude sur le birapport varie en fonction du birapport de manière linéaire pour les petites valeurs ($k \in [0, 2]$) et de manière quadratique pour les grandes valeurs. Les configurations engendrant de grandes valeurs pour le birapport sont donc également à éviter. Ce sont d'ailleurs aussi les configurations dégénérées.

Nous considérons ensuite les configurations de cinq points coplanaires et nous proposons différents types d'invariants projectifs caractérisant ces configurations et indépendants de l'ordre des points. En particulier, sous l'hypothèse de conservation de l'enveloppe convexe, on peut identifier les points et se ramener au cas où les points sont ordonnés.

Les tests effectués avec ces différents invariants montrent que l'utilisation de fonctions symétriques conduit à une mauvaise séparation des différentes configurations et donc à un faible taux de discrimination. L'ordonnancement des points, en supposant la conservation de l'enveloppe convexe, par exemple, est donc un préalable nécessaire pour l'utilisation des invariants projectifs. Le calcul de l'incertitude sur le birapport s'avère efficace pour estimer la tolérance à accepter sur chaque invariant : il donne des résultats similaires à ceux obtenus avec une tolé-

rance fixe, optimisée a posteriori. Un autre résultat intéressant est que l'utilisation d'un nombre redondant d'invariants, même si elle n'apporte en théorie aucune caractérisation supplémentaire, aboutit en pratique à une nette amélioration des performances. Par contre les performances se dégradent rapidement dès lors qu'on augmente l'amplitude du bruit sur les coordonnées des points.

Cette étude montre donc que les invariants projectifs permettent un important filtrage. La sélection des candidats pourrait encore être améliorée en considérant la compatibilité des points mis en correspondance par chaque association entre une configuration modèle et une configuration candidate. Meer [Mee 92] propose un algorithme pour prendre en compte cette information, avec une combinatoire raisonnable. Cependant, les invariants projectifs ne semblent pas pouvoir être utilisés comme seuls indices de reconnaissance. Ils se présentent plutôt comme une caractérisation complémentaire, associée à d'autres caractérisations, topologiques par exemple. Les recherches envisagées dans l'équipes s'orientent donc vers une coopération de ces deux types de caractérisation, regroupant les travaux de H. Sossa [Sos 92a] sur la caractérisation par des invariants topologiques et les études menées sur les invariants affines et les invariants projectifs tridimensionnels par P. Gros et L. Quan [Gro 92] ainsi que les conclusions de notre propre étude.

Conclusion

Dans cette thèse, nous avons étudié les possibilités offertes par l'utilisation de la géométrie projective et de ses invariants, pour le positionnement et la reconnaissance en vision par ordinateur.

Nous avons d'abord montré le lien existant entre les transformations projectives de l'espace dans le plan et les modèles de projection utilisés pour approximer la transformation de l'espace dans le plan effectuée par une caméra. Nous montrons qu'une transformation projective non dégénérée de l'espace dans le plan est soit une projection perspective, soit une projection parallèle, soit la composition d'une projection orthogonale de l'espace dans le plan suivie d'une transformation projective plane (homographie).

Les propriétés de la géométrie projectives et en particulier ses invariants peuvent être appliquées au positionnement relatif. Nous avons montré que des problèmes fondamentaux de la vision artificielle tels que

- comment rétroprojeter un point image?
- comment projeter un point de l'espace?
- comment déterminer la géométrie épipolaire?

peuvent être résolus par des constructions géométriques directes en utilisant uniquement des propriétés de la géométrie projective.

Cette méthode de positionnement relatif a plusieurs avantages:

- Un étalonnage numérique explicite, souvent coûteux, est ici évité; seuls des points de référence sont nécessaires.

- La méthode n'implique que des calculs simples, presque toujours le calcul arithmétique de birapports.
- La méthode fournit les même possibilités que l'étalonnage traditionnel, et parfois plus, par exemple des inférences géométriques à partir des données de référence partielles.
- On utilise uniquement des propriétés des invariants projectifs purs.
- La méthode permet une appréciation intuitive du processus d'étalonnage ; les phénomènes d'instabilité peuvent être facilement isolés, nous donnant ainsi des clés pour l'élaboration de méthodes plus robustes et plus stables.

De plus nous avons montré par une étude comparative que la précision du positionnement relatif était comparable à celle obtenue avec une caméra étalonnée. Dans le cas où la position des caméras n'est connue qu'approximativement, une méthode employant partiellement les invariants projectifs est moins sensible aux erreurs. Notre étude dans un cas simple a abouti aux conclusions suivantes. Les méthodes géométriques sont limitées par la présence de configurations dégénérées : pour certaines positions de la caméra (qui induisent, par exemple, des alignements entre les points de référence et le centre de projection), les méthodes géométriques ne sont pas utilisables. Par contre, en dehors de ces configurations, la précision obtenue avec la méthode géométrique correspond à une précision sur l'orientation des lignes de vue inférieure à 0.2 degrés, c'est à dire une précision sur l'orientation de la caméra inférieure à 0.2 degrés (ceci sans tenir compte des erreurs d'étalonnage sur les paramètres intrinsèques).

Ce travail a été précurseur dans l'exploration du positionnement relatif à l'aide des propriétés de la géométrie projective et il a été suivi de plusieurs travaux qui ont permis de généraliser cette méthode. Dans notre équipe, en particulier, a été développée tout récemment une méthode de positionnement relatif robuste à partir d'une séquence d'images. L'appariement des points dans les images, étape que nous avons supposée résolue dans notre travail, est effectuée de manière automatique, par un suivi des points et à l'aide, encore une fois, de la géométrie projective pour la détermination de la géométrie épipolaire. De plus, un nombre de points de référence redondant peut être pris en compte.

Des méthodes similaires ont été développées indépendamment par Faugeras et Maybank [Fau 92c] [May 92] et Hartley [Har 92b] [Har 92a].

La deuxième partie de notre travail concerne la stabilité des invariants projectifs. Nous donnons des résultats théoriques, en particulier l'expression analytique de la fonction de probabilité du birapport de quatre points. Nous proposons une mesure de similarité pour les birapports, ainsi qu'une évaluation a priori de l'incertitude sur les birapports de quatre points alignés ou de cinq points coplanaires, étant données des coordonnées des points et les incertitudes sur ces coordonnées. Nous identifions aussi les configurations les plus instables et les moins discriminantes et

nous donnons des moyens pour filtrer ces configurations.

Nous proposons ensuite différents invariants de cinq points coplanaires, tenant compte du fait que les points ne peuvent souvent pas être ordonnés.

Les tests effectués sur des données simulées et bruitées montrent les performances de ces invariants dans un processus de reconnaissance. Nous en déduisons aussi des conséquences pratiques sur le choix des invariants : des invariants utilisant des fonctions symétriques mènent à une mauvaise séparation des configurations modèles et sont à éviter. La redondance des invariants, par contre, apporte en pratique de meilleures performances.

Cette étude est assez générale pour que ses résultats puissent être utilisés pour toute méthode utilisant des invariants projectifs s'exprimant à l'aide de birapports. Elle donne aussi des indications sur les aspects méthodologiques concernant un processus de reconnaissance à l'aide des invariants projectifs.

En conclusion, les approches utilisant les invariants projectifs pour le positionnement relatif et pour la reconnaissance se multiplient. Elles se présentent comme une alternative lorsque les techniques classiques utilisant l'étalonnage ne peuvent pas être employées. Si les modèles de projection plus simples (projection parallèle ou projection orthogonale) sont valides, on dispose aussi des invariants affines ou euclidiens, moins généraux mais plus robustes. Lorsque ces modèles ne sont plus assez précis et que la projection perspective doit être considérée, les invariants projectifs restent les seules caractéristiques géométriques utilisables.

Dans le cas de caméras mobiles, ce type d'approche semble particulièrement intéressant : on peut montrer que dans le cas où les paramètres intrinsèques des caméras restent fixés, on aboutit à une représentation euclidienne de la scène observée [May 92]. Dans le cas contraire (utilisation de plusieurs caméras, caméra subissant des mises au points...), seule une représentation projective peut être déterminée. Pour pouvoir l'appliquer à la robotique mobile, il faudrait employer une description projective des commandes de déplacement des robots. Peut-on envisager que des robots se déplacent un jour dans des espaces projectifs?

Annexe A

Transformations projectives et projections : démonstration

Nous allons démontrer le théorème 1 que nous avons énoncé au paragraphe 1.3.3.

Une transformation projective de $\mathbb{P}^3$ dans $\mathbb{P}^2$ est décrite par une matrice 3×4 définie à un facteur multiplicatif près. Nous commençons par établir la forme des matrices associées aux projections perspectives et aux projections parallèles de l'espace dans le plan. Une matrice de projection décrit les coordonnées d'un point image m dans le repère image $\mathcal{R}_i$ par rapport aux coordonnées du point scène M correspondant dans un repère scène $\mathcal{R}_s$. Ces deux repères sont définis comme suit :

- $R_s = (O, \vec{U}, \vec{V}, \vec{W})$ est un repère euclidien quelconque,
- $R_i = (o, \vec{u}, \vec{v})$ est un repère affine quelconque du plan image. On note $\alpha = \widehat{(\vec{u}, \vec{v})}$ l'angle orienté entre les deux vecteur de base.

Étant donné M un point de la scène et m son image par une projection, les coordonnées (x, y) de m dans $\mathcal{R}_i$ s'expriment en fonction des coordonnées (X, Y, Z) de

$$-v' = V'$$

La projection perspective de centre C sur le plan image $\mathcal{P}$ prend une forme très simple si on l'exprime dans les repères $\mathcal{R}'_s$ et $\mathcal{R}'_i$. Soit un point M de coordonnées (X', Y', Z') dans $\mathcal{R}'_s$. M a pour image le point m de coordonnées (x', y') dans $\mathcal{R}'_i$. M et m sont liés par la matrice de projection perspective suivante :

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}_{\mathcal{R}'_i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}_{\mathcal{R}'_s} \quad \text{où } \lambda \in \mathbb{R}^*$$

Nous allons en déduire la forme de la matrice de projection dans les repères $\mathcal{R}_s$ et $\mathcal{R}_i$.

La matrice de passage entre les repères $\mathcal{R}'_i$ et $\mathcal{R}_i$ exprime les coordonnées (x, y) d'un point m dans le repère $\mathcal{R}_i$ en fonction de ses coordonnées (x', y') dans le repère $\mathcal{R}'_i$. Si (u_0, v_0) sont les coordonnées de c dans $\mathcal{R}_i$, on a :

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{\mathcal{R}_i} = \begin{bmatrix} \frac{1}{u} & \frac{1}{u \tan \alpha} & u_0 \\ 0 & -\frac{1}{v \sin \alpha} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}_{\mathcal{R}'_i} \quad \text{où } \lambda \in \mathbb{R}^*$$

d'où la matrice de projection perspective entre $\mathcal{R}'_s$ et $\mathcal{R}_i$:

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{\mathcal{R}_i} = \begin{bmatrix} \frac{1}{u} & \frac{1}{u \tan \alpha} & u_0 & 0 \\ 0 & -\frac{1}{v \sin \alpha} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}_{\mathcal{R}'_s} \quad \text{où } \lambda \in \mathbb{R}^*$$

On pose :

$$a = \frac{1}{u} \quad b = \frac{1}{u \tan \alpha} \quad c = -\frac{1}{v \sin \alpha}$$

Ce changement de variables est valide car il existe une bijection entre les triplets $(a, b, c) \in \mathbb{R}^* \times \mathbb{R} \times \mathbb{R}^*$ et les triplets $(u, v, \alpha) \in \mathbb{R}^* \times \mathbb{R}^* \times]0, \Pi[$.

Les coordonnées d'un point dans les deux repères $\mathcal{R}_s$ et $\mathcal{R}'_s$ sont liées par la relation :

$$\lambda \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}_{\mathcal{R}'_s} = \begin{bmatrix} & R & & -R t \\ & & & \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_{\mathcal{R}_s} \quad \text{où } \lambda \in \mathbb{R}^*$$

où R est une matrice de rotation et $t = (x_c, y_c, z_c)^t$ sont les coordonnées de C dans $\mathcal{R}_s$.

On en déduit enfin la forme de la matrice de projection perspective P entre les repères $\mathcal{R}_i$ et $\mathcal{R}_s$:

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{\mathcal{R}_i} = \begin{bmatrix} a & b & u_0 & 0 \\ 0 & c & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} R & -Rt \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_{\mathcal{R}_s} \quad \text{où } \lambda \in \mathbb{R}^*$$

c'est à dire

$$P_1 = \begin{bmatrix} \mathcal{M}_1 R & -\mathcal{M}_1 R t \end{bmatrix}$$

où

$$\mathcal{M}_1 = \begin{bmatrix} a & b & u_0 \\ 0 & c & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$t = (x_c, y_c, z_c)$$

On peut remarquer que a et c étant non nuls, la matrice $\mathcal{M}_1$ est de rang 3. R étant une matrice de rotation, $\mathcal{M}_1 R$ est aussi de rang 3.

A.2 Matrice de projection orthographique

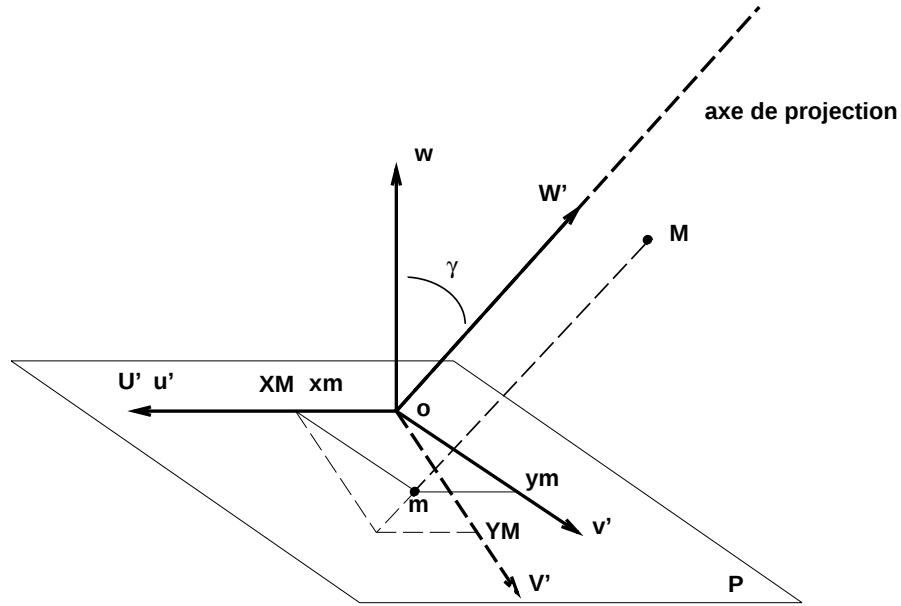


FIG. A.2 - Les différents repères utilisés pour la projection orthographique.

Pour établir la forme d'une matrice de projection orthographique, nous introduisons à nouveau des repères intermédiaires $\mathcal{R}'_s$ et $\mathcal{R}'_i$ (voir figure A.2). Ces repères

sont définis comme suit :

- $\mathcal{R}'_s = (o, \vec{U}', \vec{V}', \vec{W}')$ est un repère euclidien d'origine o , point origine de $\mathcal{R}_i$ et dont les vecteurs unitaires sont déterminés par :
 - $\vec{W}'$ est colinéaire à l'axe de projection,
 - $\vec{U}'$ est dans le plan de projection et orthogonal à $\vec{W}'$,
 - $\vec{V}'$ est tel que $(\vec{U}', \vec{V}', \vec{W}')$ soit orthonormé direct.
- $\mathcal{R}'_i = (o, \vec{u}', \vec{v}')$ est un repère affine du plan image, d'origine o . Les vecteurs unitaires du repère sont :
 - $\vec{u}' = \vec{U}'$
 - $\vec{v}'$ est orthogonal à $\vec{u}'$ et de norme $v' = \frac{1}{\sin \gamma}$ où γ est l'angle entre la normale au plan de projection et l'axe de projection. (voir figure A.3)

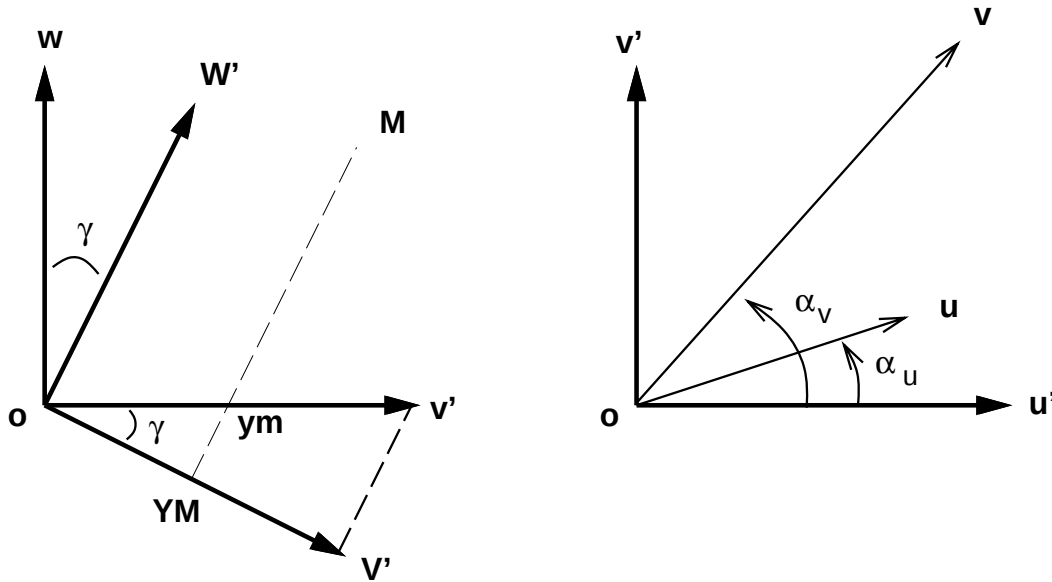


FIG. A.3 - Projections sur les plans $(o, \vec{v}, \vec{w})$ et $(o, \vec{u}, \vec{v})$

La matrice de projection orthographique a la forme suivante dans les repères $\mathcal{R}'_s$ et $\mathcal{R}'_i$:

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}_{\mathcal{R}'_i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}_{\mathcal{R}'_s} \quad \text{où } \lambda \in \mathbb{R}^*$$

On en déduit la forme de la matrice P_2 de projection orthographique exprimée dans les repères $\mathcal{R}_s$ et $\mathcal{R}_i$.

La matrice de changement de repère affine entre $\mathcal{R}'_i$ et $\mathcal{R}_i$ dépend des angles α_u et α_v que font les vecteurs $\vec{u}$ et $\vec{v}$ avec le vecteur $\vec{u}'$, de la norme de $\vec{u}$ et $\vec{v}$ et de l'angle γ (voir figure A.3).

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{\mathcal{R}_i} = \frac{1}{u v \cos(\alpha_u - \alpha_v)} \begin{bmatrix} \sin \alpha_v v & \frac{-\cos \alpha_v v}{\sin \gamma} & 0 \\ -\sin \alpha_u u & \frac{-\cos \alpha_u u}{\sin \gamma} & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}_{\mathcal{R}'_i} \quad \text{où } \lambda \in \mathbb{R}^*$$

On pose

$$\begin{aligned} d &= \frac{\sin \alpha_v}{u \cos(\alpha_u - \alpha_v)} \\ e &= \frac{-\cos \alpha_v}{u \cos(\alpha_u - \alpha_v) \sin \gamma} \\ f &= \frac{-\sin \alpha_u}{v \cos(\alpha_u - \alpha_v)} \\ g &= \frac{\cos \alpha_u}{v \cos(\alpha_u - \alpha_v) \sin \gamma} \end{aligned}$$

Le changement de repère entre $\mathcal{R}'_s$ et $\mathcal{R}_s$ est comme dans le cas de la projection perspective un déplacement et il s'exprime sous la même forme. On en déduit l'expression de la projection orthographique dans les repères $\mathcal{R}_s$ et $\mathcal{R}_i$.

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{\mathcal{R}_i} = \begin{bmatrix} d & e & 0 & 0 \\ f & g & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} R & -R t \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_{\mathcal{R}_s} \quad \text{où } \lambda \in \mathbb{R}^*$$

c'est à dire

$$P_2 = \begin{bmatrix} \mathcal{M}_2 R & \mathcal{N}_2 u \end{bmatrix}$$

où R est une matrice de rotation de dimensions 3×3 et où

$$\begin{aligned} \mathcal{M}_2 &= \begin{bmatrix} d & e & 0 \\ f & g & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \mathcal{N}_2 &= \begin{bmatrix} d & e & 0 & 0 \\ f & g & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ u &= (-R t, 1) \end{aligned}$$

On peut remarquer que quelque soit la matrice de rotation R , la matrice $\mathcal{M}_2 R$ présente une dernière ligne nulle.

A.3 Matrice de transformation projective 3D-2D

Nous avons établi dans les paragraphes précédents la forme des matrices de projection perspective et orthographique. Il nous faut maintenant étudier si toute matrice de transformation projective se ramène soit à une matrice de projection perspective, soit à une matrice de projection orthographique.

On se donne une matrice P de transformation projective de $\mathbb{P}^3$ dans $\mathbb{P}^2$, supposée de rang 3. Une matrice P de rang 2 correspond à un cas dégénéré : une projection de $\mathbb{P}^3$ dans $\mathbb{P}^1$ ou de $\mathbb{P}^3$ sur une droite de $\mathbb{P}^2$. Nous excluons donc ce cas de figure.

Considérons d'abord les trois premières colonnes d'une matrice de transformation projective, qui forment une matrice A de taille 3×3 .

A.3.1 A de rang 3

Supposons d'abord que la matrice A est de rang 3. Alors il existe une matrice de projection perspective P_1 telle que $P_1 = P$.

Nous devons d'abord montrer qu'il existe une matrice de la forme $\mathcal{M}_1$ et une matrice de rotation R telles que $\mathcal{M}_1 \times R = A$. On peut se ramener au cas où la dernière ligne de A est $(0, 0, 1)$, grâce au lemme suivant :

Lemme 1 *Soit A , une matrice de taille 3×3 dont la dernière ligne est non nulle. Il existe une matrice de rotation R_1 et une matrice B de dernière ligne $(0, 0, 1)$ telles que $B \times R_1 = A$.*

Il nous suffit donc de montrer que pour toute matrice B de dernière ligne égale à $(0, 0, 1)$, il existe une matrice de rotation $R' = R \times R_1^{-1}$ et une matrice de la forme $\mathcal{M}_1$ telles que $\mathcal{M}_1 \times R' = B$. En développant ce produit, on obtient :

$$\mathcal{M}_1 \times R' = \begin{bmatrix} a r'_1 + b r'_2 + u_0 r'_3 \\ c r'_2 + v_0 r'_3 \\ r'_3 \end{bmatrix}$$

En posant

$$\begin{aligned} r'_1 &= (\cos\theta, \sin\theta, 0) \\ r'_2 &= (-\sin\theta, \cos\theta, 0) \\ r'_3 &= (0, 0, 1) \end{aligned}$$

on obtient pour chacun des six éléments non nuls de B une équation en fonction de a, b, c, u_0, v_0 et θ . La résolution de ces six équations fournit a, b, c, u_0, v_0 et θ en fonction des coefficients non nuls de B :

$$\tan\theta = \frac{b_{21}}{b_{22}}$$

$$\begin{aligned}
a &= b_{11}\cos\theta + b_{12}\sin\theta \\
b &= b_{12}\cos\theta - b_{11}\sin\theta \\
c &= \frac{-b_{21}}{\sin\theta} \\
u_0 &= b_{13} \\
v_0 &= b_{23}
\end{aligned}$$

Il existe donc une matrice de rotation R et une matrice de projection perspective $\mathcal{M}_1$ telles que $\mathcal{M}_1 \times R = A$.

Considérons à présent la dernière colonne de la matrice $P [p_{14}, p_{24}, p_{34}]^t$. On cherche $\lambda \in \mathbb{R}^*$ et t tels que $\lambda[p_{14}, p_{24}, p_{34}]^t = -\mathcal{M}_1 R t$. a et c étant non nuls, la matrice $M_1 R$ est inversible. Il existe donc (x_c, y_c, z_c) vérifiant cette égalité pour tout triplet $[p_{14}, p_{24}, p_{34}]$, λ étant choisi arbitrairement.

Une matrice P de dimensions 3×4 dont la sous-matrice A formée de ses trois premières colonnes est de rang 3 peut donc être interprétée comme une matrice de projection perspective.

A.3.2 A de rang 2

Supposons maintenant que la matrice A est de rang 2.

A de dernière ligne nulle

Supposons que la dernière ligne de A est nulle. Alors il existe une matrice de projection orthographique $\mathcal{M}_2$ et une matrice de rotation R telles que $\mathcal{M}_2 \times R = A$. Pour le démontrer nous utilisons le lemme suivant :

Lemme 2 Soit A , une matrice 3×3 de rang 2. Il existe une matrice de rotation R_1 et une matrice B de dernière colonne nulle telles que $A \times R_1 = B$.

La démonstration de ce lemme est donnée plus loin.

A ayant une dernière ligne nulle, $A \times R_1$ a aussi une dernière ligne nulle. B est donc une matrice de dernière ligne et de dernière colonne nulles ; c'est la forme d'une matrice $\mathcal{M}_2$. Nous avons donc trouvé des matrices $\mathcal{M}_2 = B$ et $R = R_1^{-1}$ qui vérifient la condition recherchée $\mathcal{M}_2 \times R = A$.

Il reste à considérer la dernière colonne de P . On cherche $\lambda \in \mathbb{R}^*$ et t tels que $\lambda[p_{14}, p_{24}, p_{34}]^t = \mathcal{N}_2 u$ avec $u = (-Rt, 1)$. La matrice N_2 étant de rang 3, pour tout triplet $[p_{14}, p_{24}, p_{34}]$, il existe (au moins) un vecteur u tel que $[p_{14}, p_{24}, p_{34}]^t = \mathcal{N}_2 u$.

La dernière coordonnée de u est non nulle. En effet, le contraire impliquerait que $p_{34} = 0$, ce qui signifierait une matrice P dernière ligne nulle et donc de rang 2. En utilisant le facteur λ , on peut donc trouver u de dernière coordonnée égale à 1.

R étant inversible, pour toute valeur de u , il existe t tel que $u = (Rt, 1)$.

On a donc montré que dans le cas où la sous-matrice A est une matrice de dernière ligne nulle, la matrice P peut être interprétée comme une matrice de projection orthographique.

A de dernière ligne non nulle

Dans le cas où A est de rang 2 et de dernière ligne non nulle, il n'existe aucune projection perspective ni orthographique, associée à la matrice de transformation projective P . A quelle transformation géométrique correspond alors une telle matrice?

Il s'agit en fait d'une projection orthogonale sur le plan de projection, suivie d'une transformation projective (homographie) dans le plan de projection. Une telle transformation a la forme suivante :

$$\begin{aligned} P_3 &= \lambda H \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} & R & -Rt \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \lambda \begin{bmatrix} \mathcal{M}_3 R & \mathcal{N}_3 u \end{bmatrix} \quad \text{où } \lambda \in \mathbb{R}^* \end{aligned}$$

où H est une matrice de taille 3×3 correspondant à une homographie plane, R est une matrice de rotation de dimensions 3×3 et où

$$\begin{aligned} \mathcal{M}_3 &= \begin{bmatrix} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & 0 \\ h_{31} & h_{32} & 0 \end{bmatrix} \\ \mathcal{N}_3 &= \begin{bmatrix} h_{11} & h_{12} & 0 & h_{13} \\ h_{21} & h_{22} & 0 & h_{23} \\ h_{31} & h_{32} & 0 & h_{33} \end{bmatrix} \\ u &= (-Rt, 1) \end{aligned}$$

Il nous faut maintenant démontrer que toute matrice P dont la sous-matrice A est de rang 2 et de dernière ligne non nulle peut se décomposer sous cette forme. Il faut donc chercher H , R , et t , tels que $P = [\mathcal{M}_3 R \quad \mathcal{N}_3 u]$.

En utilisant le lemme 2, on peut trouver R_1 matrice de rotation et B de dernière colonne nulle telles que $A \times R_1 = B$. On choisit alors $R = R_1^{-1}$, et une matrice H dont les deux premières colonnes sont égales aux deux premières colonnes de B . On a alors $\mathcal{M}_3 R = A$.

Considérons à présent la dernière colonne de P : $p_4 = [p_{14}, p_{24}, p_{34}]$. On cherche u et $\mathcal{N}_3$ tels que

$$p_4 = \mathcal{N}_3 u = -\mathcal{M}_3 R t + h_3$$

où $h_3 = [h_{13}, h_{23}, h_{33}]^t$ est la dernière colonne de la matrice H . Avec $t = [0, 0, 0]$ et $h_3 = p_4$, cette condition est vérifiée.

Pour toute matrice P dont la sous matrice A est de rang 2 et de dernière ligne non nulle, on a donc trouvé une homographie H et un changement de repère (R, t) tels que la matrice P représente une projection orthogonale sur un plan, suivie d'une homographie H dans ce plan.

A.3.3 A de rang 1

Il est facile de constater que si A est de rang 1, alors P est nécessairement de rang inférieur ou égal à 2. Il s'agit donc d'un cas de projection dégénérée.

A.4 Démonstration des lemmes

Démonstration du lemme 1

Soit A une matrice 3×3 quelconque. On cherche R matrice de rotation et B matrice de dernière ligne $(0, 0, 1)$ telles que $B \times R = A$. En développant le produit, on obtient :

$$a_{3i} = r_{3i} \quad i = 1, 2, 3$$

et

$$\begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{bmatrix} = \begin{bmatrix} R^t & 0 \\ 0 & R^t \end{bmatrix} \times \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{21} \\ b_{22} \\ b_{23} \end{bmatrix}$$

On peut normaliser la matrice A en imposant $\|a_{31}^2 + a_{32}^2 + a_{33}^2\| = 1$. Ceci détermine la troisième ligne de la matrice de rotation R . On choisit la première et la deuxième ligne de R telles que R soit une matrice de rotation (orthonormée). Les six coefficients de B sont alors déterminés par un système de six équations linéaires en les coefficients de A . La matrice du système est inversible de déterminant non nul. Il existe donc une et une seule solution pour B . On a donc trouvé deux matrices R et B vérifiant les conditions demandées et telles que $B \times R = A$. $\square$

Démonstration du lemme 2

Soit A une matrice 3×3 de rang 2. On cherche R matrice de rotation et B matrice de dernière colonne nulle telles que $A \times R = B$. A étant de rang 2, il existe un vecteur v non nul, de norme unité, appartenant au noyau de A , c'est-à-dire tel

que :

$$Av^t = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Soit R une matrice de rotation dont la dernière colonne est le vecteur v . Alors $A \times R$ est une matrice de dernière colonne nulle. $\square$

Annexe B

Intersection de deux coniques

Nous donnons ici la solution explicite à l'intersection de deux coniques ayant trois points en commun. Les coniques sont définies comme dans le paragraphe 1.5.1 et nous choisissons un repère de référence construit à partir des trois points d'intersection, A , B et C : ils sont respectivement l'origine et les extrémités de nos vecteurs de base. Un tel changement de repère est toujours possible et simplifie notablement la formule complexe que nous obtenons.

La première conique C_λ est définie par le birapport λ et A , B , C et un quatrième point $D = (x_d, y_d)$ (voir section 1.5.1). La deuxième conique C_μ est définie par le birapport μ , A , B , C et un quatrième point $F = (x_f, y_f)$. À l'aide d'un logiciel de calcul formel (Maple), nous avons exprimé x et y , les coordonnées du quatrième point d'intersection I de C_λ et C_μ .

$$\begin{aligned}x &= \frac{Z (\mu x_f (1 - y_d - x_d) - \lambda x_d (1 - x_f - y_f))}{\Delta} \\y &= \frac{Z (\mu y_f (1 - y_d - x_d) - \lambda y_d (1 - x_f - y_f) + x_d y_f - y_d x_f - y_f + y_d)}{\Delta}\end{aligned}$$

où

$$Z = \mu x_f y_d (\lambda - 1) + \lambda x_d y_f (1 - \mu)$$

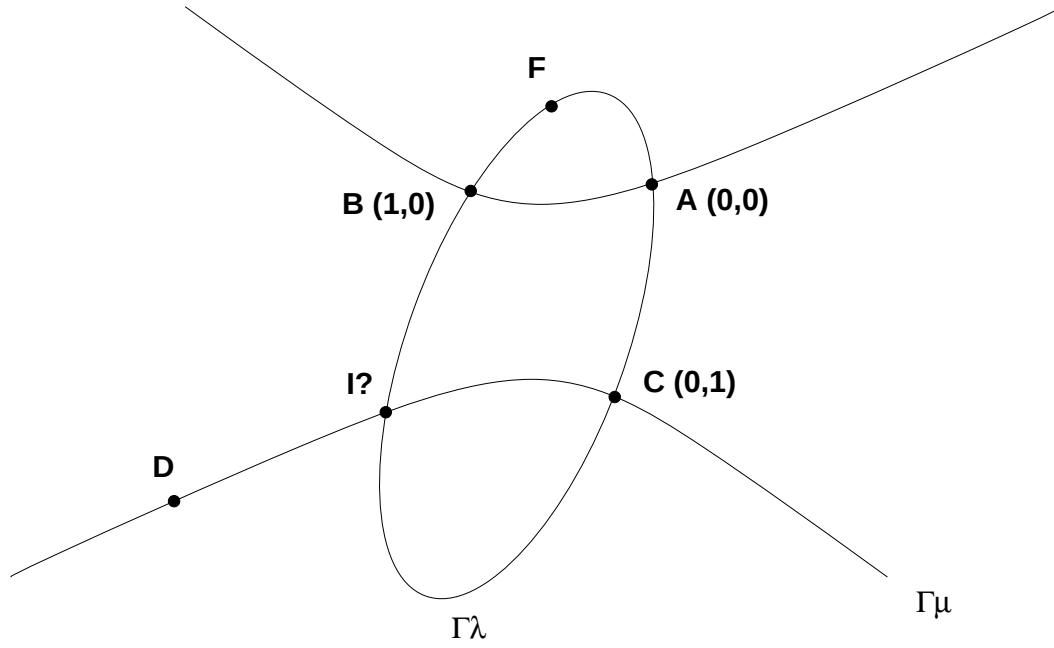


FIG. B.1 - Détermination de la quatrième intersection de deux coniques à partir du théorème de Chasles :

$[I; A, B, C, D] = \lambda$ et $[I; A, B, C, F] = \mu$

$$\begin{aligned}
 \Delta = & \lambda \mu [-x_f (1 - y_d) + (1 - y_f) x_d] \\
 & [\lambda (\mu - 1) y_f (1 - y_d) - y_d (\lambda - 1) (1 - y_f) \mu \\
 & + 2 (\mu - 1) (\lambda - 1) (y_d (1 - x_f) - (1 - x_d) y_f)] \\
 & - (\mu - 1) (\lambda - 1) [y_d (1 - x_f) - (1 - x_d) y_f] \\
 & [\lambda (\mu - 1) x_d (1 - x_f) - \mu (\lambda - 1) (1 - x_d) x_f] \\
 & + [\lambda (\mu - 1) y_f x_d - \mu (\lambda - 1) x_f y_d] \\
 & [\lambda (\mu - 1) (y_f x_d - (1 - x_f) (1 - y_d)) \\
 & + \mu (\lambda - 1) ((1 - y_f) (1 - x_d) - x_f y_d)]
 \end{aligned}$$

Bibliographie de l'auteur

Chapitres d'ouvrages

- **Roger Mohr, Luce Morin, Claude Inglebert et Long Quan**
Geometric solutions to some 3D vision problems, *Integration and Control in Real Time Active Vision*, édité par J.L. Crowley, E. Granum et R. Storer, ESPRIT BRA Series, Springer-Verlag, 1991.
- **Roger Mohr, Luce Morin et Enrico Grosso**
Relative positioning with uncalibrated cameras, *Geometric Invariance in Computer Vision*, édité par J.L Mundy, A. Zisserman, MIT Press, pages 440–460, 1992.

Conférences internationales avec comité de lecture et actes

- **Roger Mohr et Luce Morin**
Relative positioning from geometric invariants, *Proceedings of the Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, USA*, pages 139–144, juin 1991.
- **Roger Mohr, Luce Morin et Enrico Grosso**
Relative positioning with poorly calibrated cameras, *Proceeding of the DARPA–ESPRIT workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland*, pages 7–45, mars 1991.

Conférences nationales avec comité de lecture et actes

- **Luce Morin, Roger Mohr et Enrico Grosso**
Positionnement relatif à partir d'invariants projectifs, *Actes du 8ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Lyon – Villeurbanne, France*, volume 1, pages 311–321, janvier 1991.
- **Luce Morin et Roger Mohr**
Positionnement relatif tridimensionnel à partir d'images, *Deuxièmes journées nationales du GRECO-PRC communication homme-machine*, EC2 Editeur, pages 397–408, janvier 1991.
- **R. Mohr and L. Morin**
Geometric solutions to some 3D vision problems, *Journées Raisonement Géométrique: de la perception vers l'action*, pages 253–277, P.R.C - G.D.R. Intelligence Artificielle, septembre 1991.

Rapports techniques

- **Roger Mohr, Luce Morin et Enrico Grosso**
Relative Positioning with Poorly Calibrated Cameras, rapport technique RT 64 IMAG, LIFIA, avril 1991.
- **Kalle Åström et Luce Morin**
Random cross ratios, rapport technique RT 88 IMAG - 14 LIFIA, octobre 1992.

Bibliographie

- [Åst 92] K. Åström et L. Morin. Random cross ratios. Rapport no. RT 88 IMAG - 14 LIFIA, LIFIA–IRIMAG, Octobre 1992.
- [Aya 87] N. Ayache et F. Lustman. Fast and reliable passive stereovision using three cameras. *Proceedings of International Workshop on Machine Vision and Machine Intelligence, Tokyo, Japan*, 1987.
- [Bal 82] D.H. Ballard et C.M. Brown. *Computer Vision*. Prentice Hall, 1982.
- [Bar 91] E.B. Barrett, P.M. Payton, N.N. Haag, et M.H. Brill. General methods for determining projective invariants in imagery. *Computer Vision, Graphics and Image Processing: Image Understanding*, 53(1):46–65, Janvier 1991.
- [Bar 92] E.B. Barrett, M.H. Brill, N.N. Haag, et P.M. Payton. Invariant linear methods in photogrammetry and model-matching. *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA*, page to appear, 1992.
- [Bat 91] H. Batatia, A. Ayache, et C. Krey. Reconnaissance d'objets polyédriques à partir d'une seule image, utilisant des invariants projectifs. *Actes du 8ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Lyon – Villeurbanne, France*, volume 1, pages 323–328. AFCET, Novembre 1991.
- [Bou 93] B. Boufama, R. Mohr, et F. Veillon. Euclidian constraints for uncalibrated reconstruction. *Proceedings of the 4th International Conference on Computer Vision, Berlin, Germany*, pages 466–470, Mai 1993.

- [Bra 92] P. Brand. Correction des Erreurs dues à la Mise en Correspondance de Points et aux Distortions de Caméras, Juillet 1992.
- [Bri 83] M.H. Brill et E.B. Barret. Closed-form extension of the anharmonic ratio to N-space. *Computer Vision, Graphics and Image Processing*, 23:92–98, 1983.
- [Bur 90] J.B. Burns, R. Weiss, et E.M. Riseman. View variation of point set and line segment features. *Proceedings of DARPA Image Understanding Workshop, Pittsburgh, Pennsylvania, USA*, pages 650–659, 1990.
- [Bur 91] J.B. Burns, R. Weiss, et E.M. Riseman. View variation of point-set and line-segment features. *Proceeding of the DARPA–ESPRIT workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland*, pages 55–108, Mars 1991.
- [Coe 91] C. Coelho, A. Heller, J. Mundy, D. Forsyth, et A. Zisserman. An experimental evaluation of projective invariants. *Proceeding of the DARPA–ESPRIT workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland*, pages 273–293, Mars 1991.
- [Der 87] R. Deriche. Using Canny’s criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1(2):167–187, 1987.
- [Dud 73] R. Duda et P.Hart. *Pattern Classification and Scene Analysis*. Wiley-Interscience, 1973.
- [Fau 87] O.D. Faugeras et G. Toscani. Camera calibration for 3D computer vision. *Proceedings of International Workshop on Machine Vision and Machine Intelligence, Tokyo, Japan*, 1987.
- [Fau 92a] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? G. Sandini, éditeur, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 563–578. Springer-Verlag, Mai 1992.
- [Fau 92b] O.D. Faugeras. *3D Computer Vision*. M.I.T. Press, 1992.
- [Fau 92c] O.D. Faugeras, Q.T. Luong, et S.J. Maybank. Camera Self-Calibration: Theory and Experiments. G. Sandini, éditeur, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 321–334. Springer-Verlag, Mai 1992.
- [For 90a] D. Forsyth, J. Mundy, et A. Zisserman. Transformational invariance - a primer. *Proceedings of the British Machine Vision Conference, Oxford, England*, pages 1–6, Septembre 1990.

- [For 90b] D. Forsyth, J.L. Mundy, A. Zisserman, et C.M. Brown. Projectively invariant representation using implicit algebraic curves. O. Faugeras, éditeur, *Proceedings of the 1st European Conference on Computer Vision, Antibes, France*, pages 427–436. Springer-Verlag, Avril 1990.
- [For 91a] D. Forsyth, J. L. Mundy, A. Zisserman, et C. Rothwell. Invariant descriptors for 3D object recognition and pose. *Proceeding of the DARPA-ESPRIT workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland*, pages 171–208, Mars 1991.
- [For 91b] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, et C. Rothwell. Invariant descriptors for 3D object recognition and pose. *IEEE Transactions on PAMI*, 13(10):971–991, Octobre 1991.
- [For 92] D.A. Forsyth, J.M. Mundy, A. Zisserman, et C.A. Rothwell. Recognising rotationally symmetric surfaces from their outlines. G. Sandini, éditeur, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 639–647. Springer-Verlag, Mai 1992.
- [Gro 91] P. Gros. Modélisation d'images par appariement et correction. *Actes du 8ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Lyon – Villeurbanne, France*, volume 3, pages 1335–1344, Novembre 1991.
- [Gro 92] P. Gros et L. Quan. Projective Invariants for Vision. Rapport no. RT 90 IMAG - 15 LIFIA, LIFIA-IRIMAG, Grenoble, France, Décembre 1992.
- [Har 92a] R. Hartley, R. Gupta, et T. Chang. Stereo from uncalibrated cameras. *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA*, pages 761–764, 1992.
- [Har 92b] R. I. Hartley. Estimation of relative camera positions for uncalibrated cameras. G. Sandini, éditeur, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 579–587. Springer-Verlag, 1992.
- [Hor 89] R. Horaud, B. Conio, O. Le Boulleux, et B. Lacolle. An analytic solution for the perspective 4-point problem. *Computer Vision, Graphics and Image Processing*, 47:33–44, 1989.
- [Hut 91] D.P. Huttenlocher. Fast affine point matching: an output-sensitive method. *Proceedings of the Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, USA*, pages 263–268, 1991.
- [Koe 89] J.J. Koenderink et A. J. van Doorn. Affine structure from motion. Rapport, Utrecht University, Utrecht, The Netherlands, Octobre 1989.

- [Lar 91a] S. Larabi. *Utilisation de la Géométrie Projective en Vision Stéréoscopique Binoculaire*. Thèse de doctorat, Institut National Polytechnique de Toulouse, France, Avril 1991.
- [Lar 91b] S. Larabi, B. Thiesse, et P. Marthon. Utilisation de la géométrie projective en vision stéréoscopique binoculaire : appariement de stéréogrammes de segments. *Actes du 8ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Lyon – Villeurbanne, France*, volume 1, pages 329–340. AFCET, Novembre 1991.
- [Lav 91] J.M. Lavest, G. Rives, et M. Dhome. Utilisation d’un objectif à focale variable en vision monoculaire en vue de la reconstruction 3D. *Actes du 8ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle, Lyon – Villeurbanne, France*, volume 1, pages 293–301. AFCET, Novembre 1991.
- [Lee 90] C.H. Lee et T. Huang. Finding point correspondences and determining motion of a rigid object from two weak perspective views. *Computer Vision, Graphics and Image Processing*, 52:309–327, 1990.
- [Len 92] R. Lenz et P. Meer. Point Configuration Invariants Under Simultaneous Projective and Permutation Transformations. Rapport no. TR - A - 0153, ATR Auditory and Visual Perception Research Laboratories, Septembre 1992.
- [LH 81] H.C. Longuet-Higgins. A computer program for reconstructing a scene from two projections. *Nature*, volume 293, pages 133–135. XX, Septembre 1981.
- [Liu 90] Y. Liu, T.S. Huang, et O.D. Faugeras. Determination of camera location from 2D to 3D line and point. *IEEE Transactions on PAMI*, 12(1):28–37, Janvier 1990.
- [Mar 78] D. Marr. Representing visual information — A computational approach. *Computer Vision Systems*, pages 61–80, 1978.
- [May 91] S.J. Maybank. The projection of two non-coplanar conics. *Proceeding of the DARPA-ESPRIT workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland*, pages 47–53, Mars 1991.
- [May 92] S.J. Maybank et O.D. Faugeras. A theory of self calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.
- [Mee 92] P. Meer. Error Tolerant Method for Invariance Based Feature Correspondence. Rapport no. TR - A - 0149, ATR Auditory and Visual Perception Research Laboratories, Août 1992.
- [Möb 85] A.F. Möbius. *Gesammelte Werke*, volume 1. Hirzel, Leipzig, 1885.

- [Moh 91] R. Mohr, L. Morin, et E. Grosso. Relative positioning with poorly calibrated cameras. *Proceeding of the DARPA-ESPRIT workshop on Applications of Invariants in Computer Vision, Reykjavik, Iceland*, pages 7–45, Mars 1991.
- [Moh 92] R. Mohr, L. Quan, F. Veillon, et B. Boufama. Relative 3D reconstruction using multiples uncalibrated images. Rapport no. RT 84-I-IMAG LIFIA 12, LIFIA-IRIMAG, 1992.
- [Mun 92] J.L. Mundy et A. Zisserman, éditeurs. *Geometric Invariance in Computer Vision*. MIT Press, Cambridge, Massachusetts, USA, 1992.
- [Peu 92a] B. Peuchot. Accurate subpixel detectors. *14th annual International Conference IEEE EMBS, Paris.*, pages 1958–1959, Octobre 1992.
- [Peu 92b] B. Peuchot. Camera virtual equivalent model - 0.01 pixel detectors. *14th Annual International Conference IEEE EMBS. Satellite Symposium on 3D Advanced Image Processing in Medecine, Rennes.*, pages 41–45, Novembre 1992.
- [Peu 92c] B. Peuchot et M. Saint-André. CCD Camera calibration virtual equivalent model. *14th Annual International Conference IEEE EMBS, Paris.*, pages 1960–1961, Octobre 1992.
- [Qua 92] L. Quan, P. Gros, et R. Mohr. Invariants of a pair of conics revisited. *Image and Vision Computing*, 10(5):319–323, Juin 1992.
- [Rao 92] N. S. Rao, W. Wu, et C. W. Glover. Algorithms for recognizing planar polygonal configurations using perspective images. *IEEE Trans. on Robotics and Automation*, 8(4):480–485, Août 1992.
- [Rot 91] C.A. Rothwell, A. Zisserman, D.A. Forsyth, et J.L. Mundy. Using projective invariants for constant time library indexing in model based vision. *Proceedings of the British Machine Vision Conference, Glasgow, Scotland*, pages 62–70. Springer-Verlag, Septembre 1991.
- [San 92] A. Sanfeliu, A. Llorens, et W. Emde. Sensibility, relative error and error probability of projective invariants of planar surfaces of 3D objects. *Proceedings of the 11th International Conference on Pattern Recognition, The Hag, Netherland*, Août 1992.
- [Sem 52] J.G. Semple et G.T. Kneebone. *Algebraic Projective Geometry*. Oxford Science Publication, 1952.
- [Sha 92] A. Shashua. Projective Structure from Two Uncalibrated Images : Structure from Motion and Recognition. Rapport no. A.I. Memo No. 1363, Massachusetts Institute of Technology, Septembre 1992.

- [Sko 88] T. Skordas. *Mise en correspondance et reconstruction stéréo utilisant une description structurelle des images*. Thèse de doctorat, Institut National Polytechnique de Grenoble, France, 1988.
- [Sos 92a] H. Sossa. *Reconnaissance d'Objets Polyédriques dans une Base de Modèles*. Thèse de doctorat, Institut National Polytechnique de Grenoble, France, Décembre 1992.
- [Sos 92b] H. Sossa et R. Horaud. Model indexing based on a graph-theoretic approach. B. Neumann, éditeur, *Proceedings of the 10th European Conference on Artificial Intelligence, Wien, Austria*, pages 826–828. John Wiley & Sons, Ltd., Août 1992.
- [Sos 92c] H. Sossa et R. Horaud. Model indexing: the graph-hashing approach. *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA*, Juin 1992.
- [Spa 90] G. Sparr et L. Nielsen. Shape and mutual cross-ratios with applications to the interior, exterior and relative orientation. *Proceedings of the 1st European Conference on Computer Vision, Antibes, France*, pages 607–609. Springer-Verlag, Avril 1990.
- [Ste 90] Ian Stewart. D'où a été prise la photo? *Pour la science*, 148:106–111, Février 1990.
- [Tho 87] D.W. Thompson et J.L. Mundy. Three-dimensional model matching from an unconstrained viewpoint. *Proceedings of IEEE International Conference on Robotics and Automation, Raleigh, USA*, pages 208–220, 1987.
- [Tom 91] C. Tomasi et T. Kanade. Factoring image sequences into shape and motion. *Proceedings of IEEE Workshop on Visual Motion, Princeton, New Jersey*, pages 21–28, Los Alamitos, California, USA, Octobre 1991. IEEE Computer Society Press.
- [Tri 87] C. Trip. Where is the camera? *Mathematical Gazette*, 71:8–14, 1987.
- [Tsa 87] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.
- [Yua 89] J.S.C. Yuan. A general photogrammetric solution for the determining object position and orientation. *IEEE Trans. on Robotics and Automation*, 5(2):129–142, 1989.

L'un des objectifs de la vision par ordinateur est la restitution des caractéristiques tridimensionnelles d'objets à partir d'une ou plusieurs images de ces objets, soit pour déterminer leur forme et leur position, soit pour les identifier. Les méthodes classiques de positionnement s'appuient sur l'étalonnage préalable des caméras, technique délicate et parfois inutilisable, comme dans le cas de caméras mobiles. Nous montrons comment l'utilisation des propriétés de la géométrie projective permet d'éviter un étalonnage explicite et aboutit à un positionnement relatif des objets observés. Des expérimentations sur des scènes réelles contenant des objets polyédriques simples permettent de valider la méthode et d'évaluer la précision du positionnement obtenu. La multiplication des techniques utilisant les invariants projectifs, tant pour le positionnement que pour la reconnaissance, nous a ensuite conduit à étudier la stabilité de ces derniers en présence de bruit dans les images. Une étude théorique nous permet de proposer une mesure de similarité entre invariants projectifs, ainsi que des moyens pour identifier et filtrer les valeurs instables. Nous considérons ensuite différentes caractérisations des ensembles de cinq points coplanaires par des invariants projectifs et nous comparons leurs performances dans le cadre d'un processus de reconnaissance sur des données simulées et bruitées.

mots-clés : vision par ordinateur, géométrie projective, positionnement relatif, autocalibrage, reconnaissance d'objets, invariants, stabilité, discrimination

One of the main goals in computer vision is to infer three-dimensional characteristics of some observed objects by analyzing one or several images of them. This allows either to determine the shape and position of these objects or to recognize them. Classical positioning methods depend on a previous calibration of the cameras. The calibration process is tedious and unstable, and sometimes even impossible, as it is the case with a moving camera. We show that using the properties of projective geometry, one can avoid explicit calibration and obtain a relative positioning of the observed objects. We validate our method by doing some experiments on real images of simple polyhedric scenes, evaluating also its precision. The development of techniques that use projective invariants for positioning or recognition leads us to study the stability of projective invariants when computed from noisy images. A theoretical study allows us to propose a similarity measure for projective invariants, and some methods to identify and eliminate unstable configurations. We then consider different characterizations of five coplanar points with projective invariants and we compare their performance in a recognition process using simulated noisy data.

key-words : computer vision, projective geometry, relative positioning, self-calibration, object recognition, invariants, stability, discrimination