Matrice fondamentale et calibration visuelle sur l'environnement. Vers une plus grande autonomie des système robotiques.

Quang-Tuan Luong

▶ **To cite this version:**

Quang-Tuan Luong. Matrice fondamentale et calibration visuelle sur l'environnement. Vers une plus grande autonomie des système robotiques.. Automatique / Robotique. Université Paris Sud - Paris XI, 1992. Français. NNT : . tel-00549134

HAL Id: tel-00549134
https://theses.hal.science/tel-00549134

Submitted on 21 Dec 2010

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.

N° d'ordre:

THÈSE

présentée à

L'UNIVERSITÉ DE PARIS-SUD
CENTRE D'ORSAY

pour obtenir le titre de

DOCTEUR EN SCIENCE

Spécialité

INFORMATIQUE

par

Quang-Tuan LUONG

Sujet de la thèse :

MATRICE FONDAMENTALE ET CALIBRATION VISUELLE SUR L'ENVIRONNEMENT

VERS UNE PLUS GRANDE AUTONOMIE DES SYSTEMES ROBOTIQUES

Soutenue le 18 Décembre 1992 devant le jury composé de :

MM.	**Jean-Pierre**	**JOUANNAUD**	Président
	Michel	**DHOME**	Rapporteurs
	Roger	**MOHR**	
	Olivier D.	**FAUGERAS**	Examinateurs
	Stephen J.	**MAYBANK**	

A mes parents, en témoignage de reconnaissance et d'affection.

Remerciements

Mes remerciements vont en premier lieu à **Olivier Faugeras** qui, après m'avoir accueilli dans son laboratoire, m'a trouvé un sujet de thèse, et m'a ouvert son foyer lorsque j'ai été blessé et seul. D'un point de vue strictement scientifique, je peux dire que je ne suis presque jamais ressorti de son bureau sans une nouvelle idée ou une réponse à une de mes questions, quelles qu'aient été ses préoccupations du moment.

Je voudrais ensuite remercier les membres du jury de thèse: **Jean-Pierre Jouannaud**, pour avoir accepté de me faire l'honneur de présider celui-ci, **Michel Dhome** et **Roger Mohr** qui ont rédigé les rapports dans des délais très brefs. Certaines des idées de **Steve Maybank** sont à l'origine de ce travail, et les discussions que nous avons eues avec lui ont été précieuses, aussi je lui suis particulièrement gré de s'être déplacé pour venir participer à ce jury. Ses remarques méticuleuses sur une première version du manuscrit m'ont également permis de corriger un certain nombre d'imprécisions.

Ma dette envers les autres membres du projet **ROBOTVIS**, que je cite dans le désordre, est vraiment considérable. Cela été l'une de mes plus grandes satisfactions que de pouvoir travailler dans une telle ambiance de gentillesse et de disponibilité. **Rachid Deriche** a joué un rôle important dans mon travail, en suggérant l'algorithme des huit points, et en attirant mon attention sur les problèmes que son emploi soulève, et sur une foule d'autres choses. Mes compagnons de bureau **Nassir Navab**, **Zhengyou Zhang**, et **Luc Robert** ont toujours été disponibles pour des discussions scientifiques souvent fructueuses, et pour répondre à des questions plus bêtes. Je voudrais leur associer **Théo Papadopoulo**, qui a contribué à ce travail en implémentant le critère euclidien pour le calcul de la matrice fondamentale, et le programme des continuations. Son expertise en Maple, ainsi que celle de Nassir m'ont aidé. Zhengyou m'a fourni beaucoup de connaissances directement utilisables, et s'est intéressé ces derniers mois à mon travail, effectuant entre autres une relecture complète de ce manuscrit. Sans les outils de Luc, je n'aurais pas pu avancer bien vite, et sans Luc lui même, beaucoup des résultats expérimentaux seraient absents. **Thierry Viéville** m'a toujours aidé à conserver le moral en me rassurant dans les périodes de doute, toujours prêt à fabriquer une solution sur mesure, son aide matérielle n'a pas été en reste. **Michel Schmitt** et **Régis Vaillant** m'ont grandement aidé à débuter, j'utilise d'ailleurs encore du code écrit par Régis. Les discussions que j'ai eues avec **Song De Ma** ont inspiré les travaux sur les plans. C'est grâce au programme de détection précise de points d'intérêts que m'a fourni **Thierry Blaszka** que j'ai pu traiter la plupart des images réelles. **Bernard Giai** a relu une grande partie du manuscrit. **Cyril Zeller** m'a aussi pas mal dépanné, ainsi que **Peter Sander**, qui a corrigé mon anglais. **Hervé Mathieu**, **Jean-Luc Szpyrka**, **Marie-Line Ramfos** et les "filles de la doc" ont totalement résolus pour moi tous mes problèmes d'acquisition d'images, de fonctionnement d'ordinateur, d'administration, et de documentation respectivement. Ce ne sont pas les moindres. Sans Mme Ferrer, la soutenance aurait sans doute eu lieu plus tard. Je voudrais également remercier les autres membres du projet, et de l'INRIA en général, trop nombreux pour être tous mentionnés, pour l'amitié qu'ils m'ont témoigné, et les jours que nous avons passés ensemble.

Table des matières

1 Calibration et Autocalibration **9**

1.1 Le modèle de la caméra 10

1.1.1 Un modèle projectif 10

1.1.2 Autres modèles 14

1.2 Qu'est-ce que l'autocalibration ? 17

1.2.1 Le problème de la calibration 17

1.2.2 Les méthodes classiques 18

1.2.3 Quels paramètres faut-il calculer ? 20

1.2.4 Evolutions de la méthode classique 22

1.2.5 Le principe de l'autocalibration 24

1.3 Comment obtenir les appariements ? 26

1.3.1 Les points d'intérêt 26

1.3.2 La mise en correspondance 28

1.4 Contributions au problème de la calibration 29

2 Relations géométriques et matrice fondamentale **33**

2.1 Matrice fondamentale et contraintes géométriques 34

2.1.1 La contrainte épipolaire 34

2.1.2 La matrice fondamentale 35

2.1.3 Lien avec la contrainte de Longuet-Higgins 37

2.1.4 Lien entre la géométrie épipolaire et le mouvement 39

2.2 La transformation épipolaire 40

2.2.1 Une remarque géométrique 40

2.2.2 Le cas des épipoles finis 41

2.2.3 Le cas général 44

2.3 La contrainte d'ordre épipolaire 46

3 Calcul des épipoles à partir de la transformation épipolaire 51
3.1 La méthode de Sturm 52
3.1.1 Théorie 52
3.1.2 Un exemple 55
3.1.3 Deux mouvements typiques 55
3.2 Utilisation de la méthode de Sturm 59
3.2.1 L'algorithme de Sturm: quelques détails 59
3.2.2 Sélection de l'épipole: la méthode exhaustive 59
3.2.3 Statistique robuste: utilisation d'une procédure de consensus aléatoire 60
3.3 Résolution approximative: la méthode des cubiques 66
3.3.1 Minimisation d'un critère algébrique 66
3.3.2 Une approche graphique 69
3.3.3 L'usage de la méthode des cubiques pour la recherche de candidats . 75
3.4 Une approche itérative: la méthode des birapports 76
3.4.1 Principe 76
3.4.2 Prise en compte de l'incertitude 77
3.4.3 L'ordre dans les birapports 78
3.4.4 Comparaison: résultats expérimentaux 80

4 Détermination de la matrice fondamentale 83
4.1 Le critère linéaire 84
4.1.1 L'algorithme des huit points 84
4.1.2 Implémentations du critère linéaire 85
4.1.3 Difficultés liées au critère linéaire 86
4.2 Critères non-linéaires 91
4.2.1 La distance aux droites épipolaires 91
4.2.2 La pondération par le gradient 91
4.2.3 La distance globale aux données: le critère euclidien 93
4.2.4 Méthodes robustes 94
4.3 Paramétrages de la matrice fondamentale 96
4.3.1 Une matrice définie à un facteur d'échelle près 96
4.3.2 Une matrice singulière 97
4.3.3 Une matrice fondamentale aux épipoles finis 98
4.4 Une comparaison expérimentale des méthodes 98
4.4.1 Approche statistique: données synthétiques 98
4.4.2 Quelques exemples réels 106

5 Une configuration spatiale particulière: les plans 113
5.1 Plans et matrice fondamentale 114
5.1.1 La correspondance entre les images d'un plan 114
5.1.2 Calcul de l'homographie 114
5.1.3 Relation entre les homographies de plans et la corrélation fondamentale 116
5.2 Stabilité du calcul de la matrice fondamentale en présence de structures planes 118
5.2.1 L'observation de base 118
5.2.2 Etude expérimentale: simulations statistiques 119
5.2.3 Et en pratique ? 120
5.3 Calcul de la matrice fondamentale à partir de plans 121
5.3.1 Une méthode de calcul direct de la matrice fondamentale 121
5.3.2 Interprétation géométrique: structures planes et détermination des épipoles 124
5.3.3 Une comparaison: résultats expérimentaux 128

6 L'influence des mouvements sur la stabilité 133
6.1 Une caractérisation probabiliste de la stabilité 134
6.1.1 Le calcul de l'incertitude: cas linéaire 134
6.1.2 Le calcul de l'incertitude: cas non-linéaire 135
6.1.3 Une validation expérimentale 136
6.2 Surface critique et ambiguïté 139
6.2.1 La surface critique 141
6.2.2 La transformation quadratique entre deux images ambigües 143
6.2.3 Calcul direct de la transformation quadratique 144
6.2.4 Calcul de la transformation quadratique à partir d'une matrice fondamentale 147
6.2.5 Exemples numériques 148
6.3 Caractérisation qualitative de l'instabilité: l'influence du type de mouvement 150
6.3.1 L'amplitude de la translation 150
6.3.2 La direction de la translation 153
6.3.3 L'angle de rotation: le cas des translations pures 154
6.4 Caractérisation qualitative de l'instabilité: l'ambiguïté 155
6.4.1 Lien théorique entre instabilité et ambiguïté 156
6.4.2 Une expérience à partir d'une surface critique 158
6.4.3 Une validation statistique globale 160

7 Le calcul des paramètres intrinsèques: théorie 163
7.1 Calculer les paramètres intrinsèques à l'aide de la contrainte de rigidité . . . 164
7.1.1 Le principe: utiliser une caméra en mouvement 164
7.1.2 Les contraintes de rigidité 165

7.1.3 Equivalence des équations de Trivedi et des conditions de Huang-Faugeras . . . 166
7.2 Les équations de Kruppa . . . 168
7.2.1 La conique absolue . . . 168
7.2.2 Les équations de Kruppa . . . 168
7.2.3 Lien avec la matrice essentielle et les conditions de Huang-Faugeras . 172
7.2.4 Combien de mouvements sont nécessaires ? . . . 175
7.2.5 Cas dégénérés . . . 176
7.3 Obtention des paramètres intrinsèques . . . 177
7.3.1 Lien entre la conique absolue et les paramètres intrinsèques . . . 177
7.3.2 Liens entre paramètres intrinsèques et coefficients de Kruppa . . . 180

8 Le calcul des paramètres intrinsèques: expériences 183
8.1 Une étude théorique: approche algébrique . . . 184
8.2 Une approche semi-analytique . . . 184
8.2.1 La méthode des continuations . . . 186
8.2.2 Implémentation et résultats . . . 187
8.3 Une approche itérative . . . 190
8.3.1 Avantages pratiques . . . 191
8.3.2 Minimisation globale . . . 192
8.3.3 Filtrage récursif . . . 194
8.4 Résultats . . . 200
8.4.1 Une évaluation des méthodes . . . 200
8.4.2 L'autocalibration est utilisable . . . 202
8.4.3 Quelques cas réels . . . 203

9 Le calcul des paramètres extrinsèques et la reconstruction 207
9.1 Le calcul du mouvement d'une caméra . . . 208
9.1.1 Méthodes de détermination du mouvement en coordonnées normalisées 208
9.1.2 Sensibilité vis-à-vis de l'imprécision sur les paramètres intrinsèques. . 214
9.2 Autocalibration par calcul simultané du mouvement et des paramètres intrinsèques . . . 217
9.2.1 Effet de l'ajustement des paramètres avec un seul mouvement . . . 217
9.2.2 Minimisation globale avec des mouvements multiples . . . 218
9.3 Méthodes pour le calcul des paramètres extrinsèques en stéréovision . . . 223
9.3.1 Les paramètres extrinsèques d'un système binoculaire . . . 223
9.3.2 Une méthode monoculaire pour calibrer un système binoculaire . . . 224
9.3.3 Détermination des matrices de projection et utilisation de mouvements multiples . . . 228
9.4 Exemples réels . . . 230

9.4.1 Calcul des paramètres intrinsèques: une série d'images avec variation de la focale 230
9.4.2 Reconstructions métriques 230

Chapitre 1

Calibration et Autocalibration

Le problème de la *calibration* des caméras est de déterminer la relation analytique qui lie les coordonnées tridimensionelles d'un point à celles, bidimensionelles, de son image par la caméra. Notre hypothèse de base est que cette relation est linéaire projective, ce qui nous permet de disposer des outils de la géométrie projective, que nous exploiterons abondamment. Dans l'introduction, nous présentons donc le modèle de caméra que nous utilisons et nous discutons des paramètres à estimer. Ceux qui nous intéressent le plus sont les cinq paramètres dits intrinsèques, qui, attachés à une caméra, sont indépendants de sa position et de son orientation. Quelques unes des approches, classiques aussi bien que récentes, du problème de la calibration sont passées en revue, et leur caractéristiques sont détaillées. Elles ne nous satisfont pas car elles nécessitent, malgré quelques progrès, l'utilisation d'une mire de calibration. En contraste, le but principal de notre approche que nous appelons *auto-calibration* est d'effectuer le calcul des paramètres intrinsèques d'une caméra en utilisant uniquement les informations disponibles dans des images quelconques prises par cette caméra. Ces informations provenant d'appariements de points, nous évoquons les méthodes pouvant être utilisées en amont de la nôtre pour obtenir des points d'intérêt et les mettre en correspondance. Nous situons pour finir nos contributions personnelles.

1.1 Le modèle de la caméra

1.1.1 Un modèle projectif

Une caméra, comme tout autre système de prise de vues, donne des images planes d'un monde physique que nous ne pouvons appréhender que comme un espace tridimensionnel. Les premiers systèmes de formation d'images connus et les plus simples sont les *sténopés*: l'appareil est constitué d'une simple boite noire, hermétique à la lumière, sauf sur la face avant qui est percée d'une ouverture ponctuelle C. Les rayons lumineux pénètrent dans la boite par l'ouverture et forment une image inversée de la scène sur le plan $\mathcal{R}$ qui constitue l'arrière de la boite, où se trouve placée une surface sensible. Une caméra ou un appareil photographique modernes fonctionnent suivant un principe qui n'est pas différent, la simple ouverture ayant été remplacée par des lentilles afin de permettre d'obtenir une zone de netteté plus étendue. Adopter le modèle du sténopé pour rendre compte de leur propriétés géométriques est la solution la plus simple, qui est la plus généralement utilisée, et conduit à un modèle géométrique qui se révèle mathématiquement riche, tout en offrant une fidélité qui est très bonne.

Le coeur du modèle est constitué par une projection perspective dont les éléments, représentés figure 1.1 sont par conséquent:

- un plan $\mathcal{R}$: la *rétine*
- un point C de l'espace n'appartenant pas à $\mathcal{R}$: le *centre optique*
- une *distance focale* f, distance entre le point C et le plan $\mathcal{R}$

La projection m d'un point de l'espace M est la trace du *rayon optique* Cm sur la rétine.

Introduisons un système de coordonnées pour l'espace tridimensionnel, attaché à la caméra. Le repère de ce système, que nous appelons *repère de la caméra*, a pour origine le centre optique C, pour axes (X_c, Y_c, Z_c). Les deux premiers vecteurs (X_c, Y_c) constituent une base orthonormée du *plan focal* $\mathcal{F}$, qui est le plan parallèle au plan rétinien et contenant le centre optique, le troisième vecteur Z_c, orthogonal au plan rétinien, définit *l'axe optique*. Si de plus, les unités du repère de la caméra sont choisies de sorte que la distance focale f soit égale à 1 dans ce repère, nous dirons que ce dernier est *normalisé*. En adoptant le *repère normalisé du plan rétinien*, qui est le repère centré au *point principal* c, intersection de l'axe optique avec le plan rétinien, et ayant pour axes (X_c, Y_c), la relation entre les coordonnées tridimensionnelles X, Y, Z d'un point M et les coordonnées bidimensionnelles u, v de son image m par la projection perspective est:

$$\frac{1}{Z} = \frac{u}{X} = \frac{v}{Y} \tag{1.1}$$

En utilisant les coordonnées projectives pour l'espace et pour le plan rétinien, l'équa-

tion (1.1) peut se réécrire:

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathbf{P}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{1.2}$$

La matrice $\mathbf{P}$ est dite *matrice de projection perspective*. Elle apparaît ici sous sa forme la plus simple car nous avons choisi deux repères privilégiés pour l'espace, tout comme pour le plan rétinien. La relation entre les coordonnes projectives de l'espace et celles du plan image est ici linéaire. Il est simple de montrer que cette propriété demeure vérifiée quels que soient les repères rétiniens et repères du monde choisis. C'est la propriété la plus importante de notre modélisation, car elle signifie que nous aurons toujours la possibilité d'utiliser tous les puissants outils de la géométrie projective. En effet, la projection de l'espace $\mathbf{IR^3}$ vers la rétine $\mathbf{IR^2}$ peut être décomposée en trois applications linéaires:

- une application
$$\pi_{3\to 3} : [X, Y, Z]^t \to [X, Y, Z, 1]^t$$
plongement de l'espace euclidien $\mathbf{IR^3}$ dans l'espace projectif $\mathcal{P}^3$

- une application linéaire p de $\mathcal{P}^3$ dans le plan projectif $\mathcal{P}^2$ qui constitue la projection elle-même.

- une application
$$\pi_{2\to 2} : [x, y, z]^t \to [\frac{x}{z}, \frac{y}{z}]^t$$
de $\mathcal{P}^2$ dans le plan euclidien $\mathbf{IR^2}$ (la rétine) composée d'un changement identique de représentant dans $\mathcal{P}^2$ et de l'injection canonique.

Aussi, nous pouvons résumer l'ensemble des hypothèses que nous faisons en disant que:

> *Nous adoptons un modèle de caméra pour lequel la relation entre les coordonnées dans le repère du monde et les coordonnées image est linéaire projective.*

Dans le cas le plus général, illustré par la figure 1.1, la matrice $\mathbf{P}$ est une matrice 3×4 quelconque, définie seulement à un facteur multiplicatif près. A condition de satisfaire la contrainte: $\|[P_{31}, P_{32}, P_{33}]\| \neq 0$, qui assure que le centre optique n'est pas situé dans le plan rétinien, elle peut se décomposer de manière unique sous la forme [152, 39]:

$$\mathbf{P} = \mathbf{A} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathbf{P}_c} \mathbf{D} \tag{1.3}$$

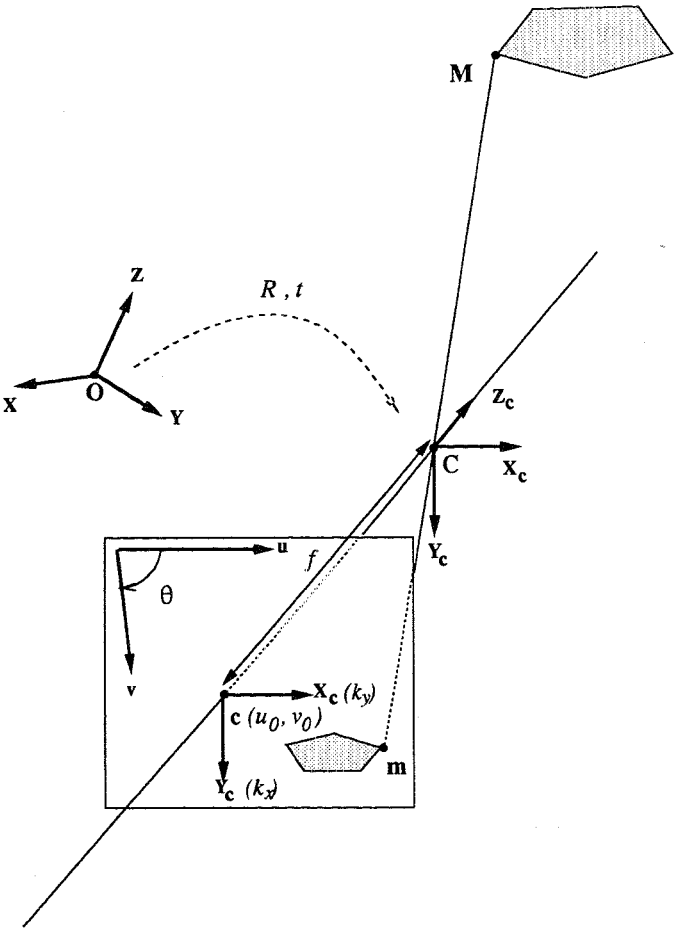


Fig. 1.1: Le modèle de la caméra

- La matrice 4×4 $\mathbf{D}$ est une matrice de déplacement de la forme:

$$\mathbf{D} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{pmatrix} \tag{1.4}$$

Elle représente le déplacement faisant passer du repère du monde au repère de la caméra. Cette matrice comprend donc 6 paramètres: 3 pour la rotation et 3 pour la translation. Ces paramètres, représentant la position et l'orientation de la caméra, sont dits *extrinsèques.*

- La matrice 3×4 $\mathbf{P}_c$ est la matrice de projection dans le repère normalisé de la caméra vers le repère rétinien normalisé, déjà rencontrée précédemment.

- La matrice 3×3 $\mathbf{A}$ représente un changement de coordonnées rétiniennes, faisant passer du repère rétinien normalisé au repère dans lequel nous mesurons effectivement les points des images finalement obtenues, que nous appelons par conséquent *repère image* ou *repère pixels.* Par convention, nous emploierons les notations $\mathbf{m}$ pour désigner les coordonnées normalisées et les notations $\mathbf{p}$, $\mathbf{q}$, pour désigner les coordonnées pixelliques. Plusieurs formes de $\mathbf{A}$ peuvent être utilisées. Par exemple dans [152, 165], l'angle θ est défini de manière quelque peu différente. Nous employons la forme proposée par [39] qui présente l'avantage de traiter de manière symétrique les deux axes:

$$\mathbf{A} = \begin{pmatrix} \alpha_u & -\alpha_u \cot\theta & u_0 \\ 0 & \alpha_v \sin\theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.5}$$

Les paramètres qui y figurent sont dits *intrinsèques.* Ils sont indépendants de la position de la caméra, et sont déterminés uniquement par les propriétés optiques de la caméra et par celles du système de numérisation. Ces paramètres sont:

  - α_u, α_v: facteurs d'échelle sur les axes horizontaux et verticaux. Ils sont fonction de la distance focale, de la taille des éléments photosensibles et des pixels, et des délais utilisés dans le système d'acquisition. Ils sont tous les deux non nuls.
  - u_0, v_0: coordonnées pixel de l'intersection de l'axe optique avec le plan rétinien. Habituellement, les pixels sont numérotés partir de l'un des coins de l'image, de sorte que u_0, v_0 correspondent normalement au centre de l'image. En pratique, il y a souvent un petit décalage dû à une imprécision de centrage.
  - θ: angle entre les deux axes du repère image. En pratique, cet angle, qui est celui entre les lignes et les colonnes sur le CCD est très bien contrôlé et peut être considéré égal à $\frac{\pi}{2}$. Son rôle a par conséquent surtout un intérêt théorique, qui est d'assurer la généricité du modèle de caméra, c'est à dire une décomposition unique de toute matrice 3×4 $\mathbf{P}$ sous la forme (1.3).

Etant donné la forme de notre modèle, il n'est pas possible de séparer les paramètres k_u et k_v, nombre de pixels par unité de longueur suivant chacun des axes, de la distance

focale focale f, ce qui explique que nous n'ayons considéré que les facteurs d'échelle $\alpha_u = fk_u$ et $\alpha_v = fk_v$, qui représentent donc la taille, en pixels horizontaux et verticaux, de la distance focale.

Nous obtenons ainsi bien 5 paramètres intrinsèques, ce qui nous fait un total de 11 paramètres pour décrire une caméra. C'est également le nombre de paramètres indépendants d'une matrice de projection perspective, puisque celle-ci, de taille 3×4, comprend 12 éléments qui ne sont définis qu'à un facteur d'échelle près. Cependant, il nous arrivera parfois de *fixer* θ à $\frac{\pi}{2}$ c'est-à-dire de considérer une matrice $\mathbf{A}$ de la forme:

$$\mathbf{A} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.6}$$

Nous appelons ce modèle: modèle simplifié à quatre paramètres, ou modèle avec contrainte de perpendicularité. C'est celui que nous utiliserons en pratique.

1.1.2 Autres modèles

Prise en compte de la distorsion Trois hypothèses implicites sont contenues dans le modèle de caméra précédent:

1. L'image se forme sur une surface qui est effectivement plane.
2. Les pixels sont régulièrement espacés dans le plan image.
3. La caméra n'a pas de distorsion optique.

Il est généralement considéré que les deux premières hypothèses sont assez bien vérifiées: les éléments photosensibles sont très régulièrement répartis sur un capteur CCD. La distorsion du plan image dans une caméra CCD est surtout due aux imprécisions de synchronisation lors de la conversion A/D (résultant en un décalage entre lignes de pixels successives) et à la transmission du signal vidéo. Des investigations photogrammétriques récentes ont montré [24, 11, 87] que ces distorsions sont faibles (0.1 pixel), et peuvent être rendues négligeables par une conversion A/D synchrone. Cependant, certaines caméras de faible qualité (utilisées par Toscani [152] par exemple) présentent une distorsion tangentielle importante et irrégulière qui justifie une correction locale sur toute l'image, technique ayant permis par ailleurs à Peuchot [127] d'atteindre une très grande précision sans qu'il soit nécessaire de spécifier complètement le modèle.

La distorsion optique n'est pas négligeable. Celle-ci transforme les coordonnées-image normalisées idéales x et y en coordonnées:

$$\begin{cases} x' = x + \delta_x(x, y) \\ y' = y + \delta_y(x, y) \end{cases}$$

Une littérature relativement abondante traite de divers modèles de distorsion, et de leur utilité pratique. Il n'est néanmoins pas très facile d'en comparer les conclusions, car elles dépendent beaucoup du matériel utilisé pour l'expérimentation, lequel est de qualité extrêmement variable. Une investigation systématique sur une optique de bonne qualité est présentée par Lavest [83], qui a utilisé les courbes de distorsion fournies par le constructeur Angénieux. Celles-ci font apparaître que la distorsion dépend de manière complexe de la distance au point principal, de la focale, et de la distance de mise au point et n'est pas systématiquement toujours dans le même sens. Un exemple est présenté par cet auteur pour montrer qu'à défaut de prendre en compte la distorsion, les erreurs de reconstruction sont dix fois plus importantes. Nous pensons que ce résultat est dû à l'utilisation d'un zoom[1], et surtout à la tâche considérée, la stéréovision axiale[2].

Les éléments de distorsion considérés le plus souvent sont:

- La distorsion radiale (symétrique autour de l'axe optique), toujours représentée par un polynôme:
$$\begin{cases} \delta_x^r = x(K_1 r^2 + K_2 r^4 + K_3 r^6 + \dots \\ \delta_y^r = y(K_1 r^2 + K_2 r^4 + K_3 r^6 + \dots \end{cases}$$
- Le décentrage [15]:
$$\begin{cases} \delta_x^c = P_1(r^2 + 2x^2) + 2P_2 xy \\ \delta_y^c = P_2(r^2 + 2y^2) + 2P_1 xy \end{cases}$$

avec $r = \sqrt{x^2 + y^2}$. La plupart des auteurs [160, 69, 57, 12, 47] trouvent qu'une précision de reprojection de l'ordre de 0.1 à 0.5 pixel est atteinte en considérant uniquement le premier terme de distorsion radiale, et qu'une telle précision est tout à fait suffisante pour des applications métriques tridimensionnelles à partir de caméras CCD. L'approximation consistant à considérer que l'erreur de décentrement est absorbée par le décentrement du point principal est assez bonne. La distorsion radiale augmentant avec l'angle de champ, elle est peu importante au centre de l'image, même pour des caméras souffrant de distorsion. Il a été remarqué [141] qu'en pratique elle peut souvent y être modélisée par une fonction linéaire $\delta^r = K_0 r$, auquel cas l'approximation consistant à considérer que la distorsion radiale est absorbée par la distance principale est assez bonne dans la zone centrale, fait utilisé par [69, 171]. En revanche, si elle demeure inférieure au pixel sur toute l'image avec certaines caméras [18, 69], auquel cas la correction n'est pas nécessaire, elle peut atteindre avec d'autres caméras une dizaine de pixels sur les bords, et doit alors être corrigée. Beyer[12], utilisant des méthodes photométriques, trouve une amélioration de la précision d'un facteur 6 en introduisant K_1, puis d'un facteur 6 avec K_2, K_3, P_1, P_2. Weng et ses collègues [171] trouvent une amélioration d'un facteur 4 avec K_1, puis seulement d'un facteur 1.2 avec quatre termes de distorsion tangentielle. La différence semble provenir de leur méthodologie de vision par ordinateur qui ne leur permet pas des localisations suffisamment précises et redondantes. Il faut en effet se souvenir que nous ne sommes pas seulement limités par la précision de calibration, mais

[1] Il est connu que les objectifs à focales fixes souffrent moins de distorsion.

[2] Il a été récemment montré [124] qu'elle est beaucoup plus sensible que la stéréovision latérale.

également par celle de mesure des points image. A moins de pouvoir détecter des points avec une précision subpixellique, il n'est pas nécessaire de corriger la distorsion si celle-ci n'entraîne que des erreurs de l'ordre du pixel. Il est aisé de vérifier ce dernier point en mesurant des droites du bord de l'image.

Au delà du modèle du sténopé Plusieurs auteurs [66, 79, 84, 85, 149, 150] ont trouvé que le modèle du sténopé ne permet pas de décrire avec une bonne précision des systèmes optiques complexes tels que les objectifs à focale variable. Ils ont alors utilisé un modèle général de lentille épaisse, pour lequel le centre optique est remplacé par deux points nodaux. La présentation la plus détaillée de ce modèle est fournie par Lavest, Rives et Dhome [84, 85], qui mettent en évidence, par une analyse optique des approximations conduisant au modèle sténopé et par de nombreuses expérimentations, le fait que dans le cas d'un objectif à focale variable le déplacement fictif du centre optique n'est pas proportionnel à la variation de distance focale. Ce déplacement est proportionnel à la variation du plan principal objet, qui est bien plus importante. Toutefois, malgré ces conclusions négatives, l'approximation par le modèle sténopé demeure valide à condition de ne plus considérer la distance entre l'objet et l'image comme constante, ce qui a surtout des conséquences en stéréovision axiale.

Une autre approche est une modélisation par deux plans [54, 169]: l'idée est qu'en partant d'un certain nombre de correspondances entre points de l'image et points tridimensionnels situés sur deux plans, les coordonnées tridimensionelles de tout point des deux plans peuvent être obtenues par interpolation à partir des coordonnées de son image. le rayon optique est obtenu en joignant les deux points situés sur les deux plans de référence. Cette technique généralise également le modèle du sténopé, et permet de prendre en compte les distorsions.

Propriétés géométriques Le principal inconvénient de l'utilisation d'un modèle avec correction de la distorsion, ou d'un modèle non-sténopé est de perdre la linéarité projective entre les coordonnées tridimensionnelles et le repère pixel. Par exemple, les épipolaires ne seront plus des droites. La linéarité projective est l'hypothèse de base pour l'utilisation du puissant formalisme de la géométrie projective dont des travaux récents, trop nombreux pour être tous cités (voir la bibliographie de l'ouvrage collectif [119]) ont montré le riche potentiel en vision. Nous utiliserons abondamment dans cette thèse les outils de base de la géométrie projective. Ceux-ci sont présentés dans plusieurs ouvrages classiques de mathématiques [137, 23, 49], et également rappelés dans des chapitres de livres consacrés à la vision par ordinateur [39, 75, 119], ce qui nous dispense de les exposer dans ce manuscript.

Nous préférons privilégier les propriétés mathématiques, tant que nous obtenons un modèle qui nous permet d'obtenir une précision suffisante pour des applications métriques tridimensionnelles à partir de caméras CCD. De toutes manières, étant donné la sensibilité des résultats obtenus actuellement par les méthodes d'*auto-calibration*, nous pensons que des perfectionnements de celles-ci sont une étape préliminaire à l'utilisation de modèles plus perfectionnés. Les résultats obtenus avec une méthode d'auto-calibration ne prenant en compte qu'un modèle linéaire projectif pourraient alors servir pour initialiser une méthode fondée sur un modèle plus complet, qui reste encore à découvrir.

1.2 Qu'est-ce que l'autocalibration ?

Nous commençons par expliquer ce en quoi consiste la *calibration*, qui constitue la problématique générale dans laquelle se situe cette thèse. Nous détaillons ensuite quelques unes des approches classiques de ce problème. Bien qu'elles permettent d'obtenir une très bonne précision, elles ne sont pas entièrement satisfaisantes car elles nécessitent l'utilisation d'une mire de calibration. Les travaux permettant de se passer d'une telle mire étant encore très rares en vision par ordinateur, nous évoquons ceux allant dans cette direction, avant de spécifier les caractéristiques originales de notre approche, que nous appelons *autocalibration*.

1.2.1 Le problème de la calibration

Le problème de la *calibration*[3] est de déterminer la relation mathématique qui lie les coordonnées tridimensionnelles d'un point à celles, bidimensionnelles, de son image par la caméra. Une fois que nous nous sommes donnés un modèle de la caméra, comme nous l'avons fait précédemment, le problème de la calibration est d'en déterminer les paramètres numériques.

Il s'agit d'un problème extrêmement important, car ce n'est qu'une fois qu'une caméra a été calibrée que nous pouvons nous servir des images qu'elle produit pour inférer des propriétés géométriques tridimensionnelles du monde réel. De plus, pour modéliser et prédire la performance de n'importe quel algorithme de vision par ordinateur, une connaissance détaillée des senseurs est nécessaire.

Le principal but de notre thèse est de présenter une méthode de calcul des paramètres intrinsèques. Ces paramètres, uniquement attachés aux caractéristiques de la caméra, sont importants car ils permettent d'effectuer des mesures métriques à l'aide d'une caméra. Leur connaissance permet de se ramener dans les repères normalisés, ce qui est l'hypothèse de base adoptée dans tous les travaux de détermination du mouvement et de la structure tridimensionnelle. En revanche, la connaissance des paramètres extrinsèques d'une caméra unique n'est pas utile dans les tâches pour lesquelles il n'existe pas un repère privilégié du monde. Dans ce cas, l'information pertinente est le positionnement relatif, soit de la caméra mobile au cours du temps, soit de celui du système de caméras, si nous en considérons plusieurs. Nous montrons également comment il est possible de déterminer ces paramètres extrinsèques, une fois que les paramètres intrinsèques sont obtenus.

Cependant le calcul des paramètres des caméras n'est pas indispensable pour l'obtention d'une partie de l'information tridimensionnelle, qui est non-métrique. La caractérisation de cette information, ainsi que le développement d'algorithmes permettant de la reconstituer, forment une ligne de recherche récente extrêmement prometteuse [70] [38] [115] [60] [132] [139]. Il s'agit d'une forme de calibration *faible*, à laquelle une partie importante de cette thèse contribue également, puisque nous présentons un grand nombre d'algorithmes permettant de calculer la *matrice fondamentale* qui est à la base de cette forme de calibration.

[3] Les termes *étalonnage* et *calibrage*, qui sont parfois employés, seraient peut-être plus corrects, mais l'usage a adopté le terme équivalent à celui existant en anglais.

1.2.2 Les méthodes classiques

Principe L'idée des méthodes classiques de calibration est simple: nous observons un objet pour lequel les coordonnées 3D de N points de référence $\mathbf{M}_i$ sont connues. Les projections $\mathbf{m}_i$ de ces points sont mesurées dans l'image sous forme de coordonnées pixelliques (u_i, v_i). Les objets utilisés sont en général des mires de calibration formées de motifs répétitifs (cercles ou rectangles) choisis pour définir des points d'intérêt qui peuvent être mesurés avec une très bonne précision. A partir de là, nous voyons qu'il est possible de calculer les paramètres λ_i de la projection à partir des N équations $\mathbf{m}_i = P(\lambda_1 \dots \lambda_n)(\mathbf{M}_i)$, pourvu que nous disposions d'un nombre suffisant N de points. Ceci peut être effectué par des méthodes d'optimisation non-linéaire [35, 143, 2], qui permettent d'utiliser des modèles photogrammétriques sophistiqués prenant en compte la distorsion, ou par un simple calcul linéaire de la matrice de transformation perspective, suivi d'une décomposition de celle-ci [173, 148, 48, 2]. La première approche peut souffrir de problèmes de temps de calcul et d'instabilité, la seconde de problèmes de précision dus à la simplification du modèle, ainsi que de la surdétermination des variables intermédiaires, qui doivent vérifier des contraintes.

A partir de 1986 sont apparues des méthodes fonctionnant en deux étapes, qui permettent d'échapper à ces écueils. A titre d'exemple, nous résumons à présent quelques caractéristiques des deux parmi les plus représentatives. Pour un état de l'art plus détaillé de la calibration en 1989, nous renvoyons le lecteur à [161]. Les travaux que nous décrirons plus bas, en 1.2.4 sont donc tous postérieurs à cette date.

la méthode de Tsai Au moyen d'une contrainte physique (dite d'alignement radial), qui ne dépend que d'un groupe de paramètres, celui-ci peut être calculé facilement et rapidement. La plupart des paramètres extrinsèques sont ainsi obtenus par bijection avec des solutions d'équations linéaires. Ces solutions sont insensibles à la distorsion radiale. Un calcul itératif permet alors de déterminer la composante en z de la translation, la distance focale effective, et les coefficients de distorsion radiale. Dans [159, 160] les coordonnées du point principal étaient supposés connues, alors qu'elles sont également calculées dans [88]. La méthode permet ainsi d'éviter des minimisations non-linéaires impliquant un grand nombre de paramètres, tout en permettant une modélisation exacte qui prend en compte deux paramètres de distorsion radiale. Elle est très rapide, ne faisant intervenir que des opérations linéaires à 5 et 2 inconnues, ainsi que des mires planes, ce qui facilite grandement la construction et l'extraction de points d'intérêt très précis. Cependant plusieurs auteurs [43, 69, 171] ont trouvé que les résultats peuvent être moins précis que ceux obtenus par une méthode linéaire, si la distorsion radiale demeure assez faible, l'explication étant que l'information donnée par la connaissance des points de référence n'est pas utilisée pleinement, puisque seule la composante tangentielle est exploitée.

la méthode de Faugeras et Toscani [43] La matrice de projection perspective $\mathbf{P}$ telle

que

$$\begin{pmatrix} su_i \\ sv_i \\ s \end{pmatrix} = \underbrace{\begin{pmatrix} \mathrm{l}_1^T & l_{14} \\ \mathrm{l}_2^T & l_{24} \\ \mathrm{l}_3^T & l_{34} \end{pmatrix}}_{\mathbf{P}} \begin{pmatrix} \mathrm{M}_i \\ 1 \end{pmatrix}$$

est d'abord estimée par le critère aux moindres carrés linéaire:

$$\sum_i \{(u_i(\mathrm{l}_3\mathrm{M}_i + l_{34}) - (\mathrm{l}_1\mathrm{M}_i + l_{14}))^2 + (v_i(\mathrm{l}_3\mathrm{M}_i + l_{34}) - (\mathrm{l}_2\mathrm{M}_i + l_{24}))^2\} \qquad (1.7)$$

Ceci est la base de toutes les approches linéaires. La contrainte $\|\mathrm{l}_3\| = 1$ assure l'invariance des paramètres intrinsèques, qui peuvent être obtenus sans ambiguïté par décomposition de $\mathbf{P}$ grâce à l'introduction de θ. Les plans constituent une configuration dégénérée. Il est à noter qu'une formulation légèrement différente du modèle due à Grotsky et Tamburino [55] permet l'usage d'une mire plane. Afin d'obtenir une mesure de l'incertitude des résultats de calibration, le filtre de Kalman peut être utilisé lors d'une minimisation non-linéaire portant sur les paramètres. Une phase optionnelle de calibration de la distorsion dans le repère pixel peut aussi être réalisée à partir d'un découpage de l'image en zones, une fois que les coefficients linéaires sont estimés. Cette approche calibre aussi la distorsion tangentielle, cependant le fait que ce calcul soit mené avec des paramètres calculés sans distorsion peut introduire un biais.

D'autres approches en deux étapes, plus récentes, améliorent le principe de celle de Faugeras et Toscani en introduisant un couplage itératif entre estimation linéaire et correction de la distorsion [171] [100].

Performances L'algorithme linéaire de Faugeras et Toscani a également fait l'objet de plusieurs perfectionnements: un découpage optimal en phases moyennées [130], et l'implémentation entièrement automatique de Vaillant [165] qui part des contours extraits et du modèle 3D de la mire bi-plan (dont on trouvera plus bas des images), pour extraire les points d'intérêt avec une grande précision et utiliser une formulation non-linéaire du critère (1.7) en termes de distance image, qui est plus robuste. Récemment, Robert [132] a encore proposé une nouvelle formulation de ce critère, en optimisant directement les coefficients de la matrice de projection perspective à partir des valeurs de gradient dans l'image, et du modèle 3D. Le fait que les résultats de cette dernière méthode soient effectivement beaucoup plus précis montre que même dans le domaine de la calibration ne faisant intervenir qu'un modèle projectif *linéaire*, les limites du modèle n'étaient pas atteintes. Les écarts obtenus sont à présent inférieurs à 1% sur les facteurs d'échelle et à quelques pixels sur le point principal. La valeur de $\theta - \frac{\pi}{2}$ demeure toujours très faible: de l'ordre de 10^{-5} à 10^{-6}. Ces méthodes peuvent donc permettre d'obtenir une très bonne précision dans les cas favorables où la mire occupe la plus grande partie de l'image, et est bien positionnée. Dans le cas de la méthode de Tsai, cela signifie que le plan de la mire forme un angle d'à peu près 30 degrés avec le plan rétinien, alors que la mire en forme de dièdre utilisée actuellement à l'INRIA doit être placée

face à celui-ci, pour que les deux plans qui la composent soient également bien visibles. Il faut noter qu'une erreur de 0.05^o dans l'angle du dièdre (supposé égal à 90^o) dégrade de manière significative la précision de la calibration [57]. De plus, si les images utilisées pour la calibration ne sont pas optimales (mires lointaines, images peu nettes ou peu contrastées, par exemple), les résultats peuvent se dégrader nettement, comme un exemple l'illustrera plus bas. Aussi la calibration demeure un domaine de recherche actif, aucune méthode ne s'étant réellement imposée comme entièrement satisfaisante.

Limitations L'inconvénient majeur de ces méthodes est surtout de *nécessiter une mire de calibration.* Il existe des applications pour lesquelles ces mires ne sont pas disponibles ou pour lesquelles leur emploi est malaisé. Un second inconvénient, qui est étroitement lié au premier, est que ces méthodes ne peuvent en général pas être utilisées en même temps que s'effectue la tâche visuelle des caméras, à moins de maintenir en permanence une mire dans le champ visuel, ce qui est peu réaliste. Or les paramètres des caméras peuvent varier pour plusieurs raisons. Il peut s'agir d'une part de modifications intentionnelles (focale, mise au point), appelées à être de plus en plus fréquentes dans le cadre de la vision active, qui constitue un domaine de recherche en plein essor (et auquel nous avons un peu contribué [93, 94]). L'emploi de la méthode classique suppose que la calibration soit effectuée au début de chaque session de travail. Dans le cas d'un système visuel actif, cela implique de précalibrer le système pour chacune de ces configurations, ce qui n'est pas très simple car un tel système peut avoir en général de 5 à 13 degrés de liberté. D'autre part, même pour les systèmes non actifs, dont les paramètres sont supposés fixes, des changements incontrôlés, dus à des variations thermiques et mécaniques sont également à considérer.

1.2.3 Quels paramètres faut-il calculer ?

Découpler l'estimation des paramètres intrinsèques Les paramètres de calibration, dans les approches classiques, devaient être estimés simultanément. Ceci présentait l'inconvénient que même si l'ensemble des paramètres, considérés comme un tout, donne un minimum, chaque paramètre, pris individuellement, n'est pas nécessairement fiable. Pour illustrer cette idée, supposons que la fonction objectif $C(\lambda_1 \dots \lambda_n)$ soit telle qu'au voisinage du minimum, $\left|\frac{\partial C}{\partial \lambda_i}\right| << \left|\frac{\partial C}{\partial \lambda_j}\right|$. Alors la valeur estimée du paramètre λ_i peut être largement perturbée pour compenser une erreur en le paramètre λ_j. Le nombre total de paramètres étant relativement élevé, ce phénomène est effectivement à craindre. Nous avons remarqué que si les conditions de calibration ne sont pas idéales, alors au cours de plusieurs calibrations obtenues en déplaçant la caméra, les paramètres intrinsèques estimés peuvent avoir une variation anormalement élevée, comme le montre la table 1.1, obtenue avec des images du type de celles présentées figure 1.2, où les mires n'occupent pas la plus grande surface de l'image. Outre l'intérêt supérieur des paramètres intrinsèques par rapport aux paramètres extrinsèques, ceci est une autre raison pour laquelle il est souhaitable de découpler leur estimation de celle des ces derniers.

image	α_u	α_v	u_0	v_0
1	466.07	704.35	281.73	208.00
2	443.63	682.94	253.16	226.87
3	477.74	729.87	229.49	259.02
4	484.97	746.03	264.62	249.06
5	444.75	716.10	224.25	216.64

Tableau 1.1: Exemple de paramètres intrinsèques obtenus à partir de différentes positions de la caméra

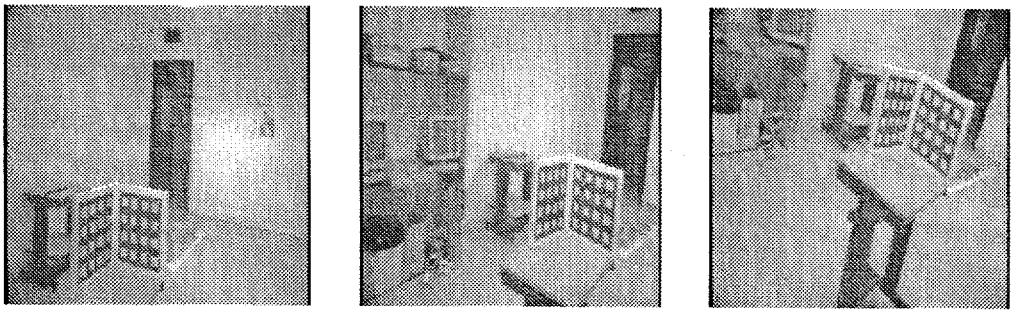


Fig. 1.2: Les images de mire ayant servi à obtenir des paramètres intrinsèques

Calculer tous les paramètres intrinsèques ? Une idée souvent avancée est qu'il n'est pas nécessaire de calculer tous les paramètres intrinsèques, certains pouvant être supposés connus, invariants, ou peu importants. Par exemple l'intersection de l'axe optique avec le plan rétinien est parfois supposée se trouver au centre de l'image par des auteurs comme Tsai [159] et Kanatani [76], qui affirment qu'une variation de la position de ce point relativement importante (une dizaine à une vingtaine de pixels) n'a que très peu d'influence sur la reconstruction tridimensionnelle. Cependant la plupart des auteurs (voir 1.1.2 pour des références) trouvent un écart très significatif et constant entre la position du point principal et le centre de l'image.

Pour des expérimentations de reconstruction, nous renvoyons le lecteur au dernier chapitre de la thèse. Notons tout de suite ici que si une grande précision de calibration est recherchée, ces approches ne nous paraissent pas les meilleures, dans la mesure où des expérimentations de calibration classique menées dans notre laboratoire ont montré que le changement d'une seule caractéristique optique (par exemple la distance focale) affecte l'ensemble des paramètres intrinsèques. Un exemple de tels résultats empruntés à Viéville [166] est donné table 1.2. Il convient de noter qu'il s'agit de valeurs stables, et parfaitement déterminables par la technique de calibration classique opérationnelle. Ces mesures, obtenues avec un objectif Canon $J8 \times 6B48$, montrent que si le rapport des facteurs d'échelle reste sensiblement constant[4], le point principal est translaté linéairement de manière non négligeable.

[4]Le fait que ce rapport demeure constant pour une caméra donnée s'explique bien par le fait qu'il ne

Lors de leurs expérimentations, Lenz et Tsai [88] ont également constaté un déplacement systématique qu'ils assimilent à une spirale, mais Lavest [83] a trouvé que le point principal demeurait stable avec une optique de bonne qualité[5].

Aussi, il semble qu'il y ait au moins trois paramètres intrinsèques significatifs.

focale	u_0	v_0	α_u	α_v	α_u/α_v
06	262.66	248.05	128.63	190.10	0.6766
08	256.33	249.07	126.85	189.60	0.6690
09	253.77	241.16	462.17	660.38	0.6998
10	258.86	251.62	556.26	793.67	0.7009
12	252.85	242.48	668.74	953.09	0.7016
14	251.40	242.41	783.06	1114.55	0.7026
16	252.73	265.28	868.10	1235.35	0.7027
25	244.85	291.51	1427.20	2033.12	0.7020
30	247.81	299.21	1742.43	2482.23	0.7018

Tableau 1.2: Exemple de paramètres intrinsèques obtenus à partir de différentes focales d'une caméra munie d'un zoom

1.2.4 Evolutions de la méthode classique

Les problèmes pratiques liés à l'emploi de mires ont suscité quelques recherches dont le but a été de réduire les exigences de celles-ci en permettant l'utilisation de dispositifs un peu moins contrôlés. Un autre objectif qui a été recherché est d'obtenir le découplage du calcul des paramètres intrinsèques de celui des paramètres extrinsèques, qui est souhaitable, comme nous venons de le voir.

Il faut remarquer que le problème de la calibration des paramètres extrinsèques seuls, une fois que les paramètres intrinsèques sont déterminés, est très proche de celui calcul du mouvement tridimensionnel et a donc suscité de nombreux travaux qu'il n'est pas possible d'énumérer. En revanche il est possible de déterminer une calibration de la géométrie épipolaire d'un système stéréo en n'utilisant que les informations présentes dans les images. Parmi ces méthodes d'autocalibration de paramètres extrinsèques, mentionnons celles de Gennery [50], qui décrit une méthode qui calcule les paramètres itérativement en minimisant les erreurs de contrainte épipolaire[6], ainsi que les approches de Toscani [154], qui utilise un critère plus robuste, Trivedi [158], et Thacker [151].

dépend pas des optiques utilisées avec celle-ci. Pour une discussion des facteurs déterminant ce rapport, ainsi qu'une méthode de calibration de celui-ci, fondée sur l'utilisation d'une méthode physique faisant appel à une mire spéciale, voir [7].

[5]utilisée dans certaines conditions: position non verticale, absence de contraintes mécaniques dues par exemple au couplage avec un moteur.

[6]Qui ne portait pas encore ce nom... Aucun résultat expérimental n'est présenté. Il est mentionné p 253 du Handbook of Artificial Intelligence [22] et p 50 de [117], que cette méthode échoue souvent

Une approche différente de la nôtre, celle de Mohr et de ses collègues [112, 114, 113] est de se passer totalement de l'estimation de la matrice de transformation perspective, et donc des paramètres intrinsèques. Les calculs tridimensionnels usuels tels que reconstruction d'un point de l'espace ou détermination de la géométrie épipolaires sont menés en utilisant des invariants projectifs simples, les birapports, et font intervenir directement les coordonnées tridimensionnelles de points de référence de la scène qui peuvent être choisis arbitrairement. Cette méthode intéressante demeure un prolongement de la méthode classique dans la mesure ou les coordonnées tridimensionnelles d'un certain nombre de points de référence sont nécessaires.

Les autres travaux que nous résumons rapidement visent tous à obtenir au moins une partie des paramètres intrinsèques.

- Caprile et Torre [16] ont été les premiers à exploiter l'information provenant des points de fuite, et à mettre en évidence leur lien avec la distance focale. Les points de fuite d'un plan observé dans l'image, qui sont les points d'intersection des images de droites qui sont parallèles dans l'espace, sont alignés et définissent la ligne de fuite de ce plan. Ils pourraient être obtenus dans les images de scènes comportant un grand nombre de parallèles en détectant celles-ci au moyen de quelques heuristiques, si bien qu'en principe aucune mire n'est requise, mais cette manière de procéder n'est pas du tout fiable. Aussi un cube est utilisé, mais il n'est pas nécessaire de connaître des coordonnées de points de référence, l'observation de trois ensembles de droites mutuellement orthogonales étant suffisant pour déterminer trois des paramètres intrinsèques, f, u_0, et v_0, ainsi que la rotation entre deux caméras. Un segment de référence permet de retrouver la translation. Des travaux similaires sont également présentés dans [168].

- Echigo [34], Kanatani et Onodera [76] ont proposé des méthodes de calibration assez similaires pour une caméra unique. Le premier utilise aussi trois ensembles de lignes parallèles pour obtenir la rotation à partir des trois équations obtenues et des six contraintes d'orthogonalité, puis la distance focale et la translation sont déterminées avec des équations linéaires. La précision semble moyenne. Les seconds utilisent les points de fuite obtenus en détectant deux ensembles coplanaires de droites orthogonales. La translation nécessite de plus la connaissance des coordonnées tridimensionnelles d'un point de référence. Etant donné la spécificité des données requises (qui conduit à des calculs assez simples), la méthode nécessite l'utilisation d'une mire spéciale, qui est néanmoins très simple. Les auteurs ne présentent pas de résultats expérimentaux.

- Beardsley, Murray et Zisserman [9] décrivent une méthode dans laquelle le dispositif consiste en un plan tournant autour d'un axe fixe, dont le positionnement est arbitraire. A partir de chaque image du plan, la ligne de fuite de celui-ci est déterminée. Au cours de sa rotation, les différentes lignes de fuites forment l'enveloppe d'une conique dont la connaissance permet l'obtention de tous les paramètres intrinsèques.

- Lavest, Rives et Dhome [85, 84] montrent que le point principal correspond au projeté du point de fuite d'un faisceau de droites parallèles à l'axe optique, ce qui leur permet

de déterminer les coordonnées du point principal *d'un zoom*[7] par suivi de la trajectoire de points de référence lors de la variation de focale. Les images de ces trajectoires sont des droites, dont l'intersection donne le point principal avec une précision de 1 pixel.

- Daucher, Dhome et Lapreste [26] montrent que l'image d'une sphère est une ellipse dont le grand axe passe par le point principal. L'observation de plusieurs sphères positionnées arbitrairement permet ainsi de déterminer simplement et avec une bonne précision le point principal et le rapport des facteurs d'échelle. La détermination précise de la distance focale est un peu plus délicate car elle prend en compte la correction de la distorsion radiale et un modèle de contour spécifique.

Malgré les progrès constitués par ces recherches, la calibration est très loin de pouvoir être menée sur des images quelconques, *des dispositifs spéciaux demeurant nécessaires.* Il est à noter que dans la plupart de ces nouvelles approches, un modèle projectif linéaire de la caméra est utilisé, en vue d'utiliser pleinement des propriétés de géométrie projective[8]. Celles-ci concernent essentiellement les points et lignes de fuite. Un travail récent, celui de Hartley [61], présente davantage d'analogies avec le notre: il vise à calculer les distances focales, puis le positionnement relatif de deux caméras, à l'aide de méthodes algébriques ne nécessitant que des appariements de points. Ce travail sera discuté plus bas, en introduction du chapitre que nous consacrons au calcul des paramètres intrinsèques.

1.2.5 Le principe de l'autocalibration

Le but principal de la méthode que nous présenterons dans cette thèse est d'effectuer le calcul de *tous* les paramètres intrinsèques d'une caméra *en utilisant uniquement les informations disponibles dans des images quelconques prises par cette caméra.* Aucune grille de calibration, ou autre objet partiellement connu n'est nécessaire. La seule restriction dont nous avons besoin est qu'une partie de la scène soit statique. Il est évident que sans aucune connaissance a priori de la scène observée, il n'est pas possible d'inférer quoi que ce soit sur les caractéristiques de la caméra. Aussi, nous sommes conduits à utiliser plusieurs vues. Notre méthode s'applique donc pour une caméra mobile. Toutefois aucune connaissance sur le mouvement de la caméra n'est nécessaire, la détermination de celui-ci est plutôt un sous-produit naturel de notre méthode. La seule chose que celle-ci exige est que nous soyons capables d'identifier des correspondances ponctuelles entre les images, c'est-à-dire des points dans deux vues différentes qui sont les images d'un même point fixe de l'espace. Par *autocalibration*, nous désignons donc une méthode qui:

- Est automatique.
- Ne nécessite aucun modèle des objets observés

[7] Ils ont travaillé avec deux objectifs pour lesquels le changement de focale est réalisé par un déplacement transversal du bloc optique, ce qui conduit à des positions stables du point principal.

[8] Les derniers travaux constituent une exception, dans la mesure où les primitives utilisées sont partiellement insensibles à la distorsion radiale.

- Ne nécessite aucune connaissance du mouvement de la caméra

- Ne nécessite aucune connaissance a priori des paramètres intrinsèques

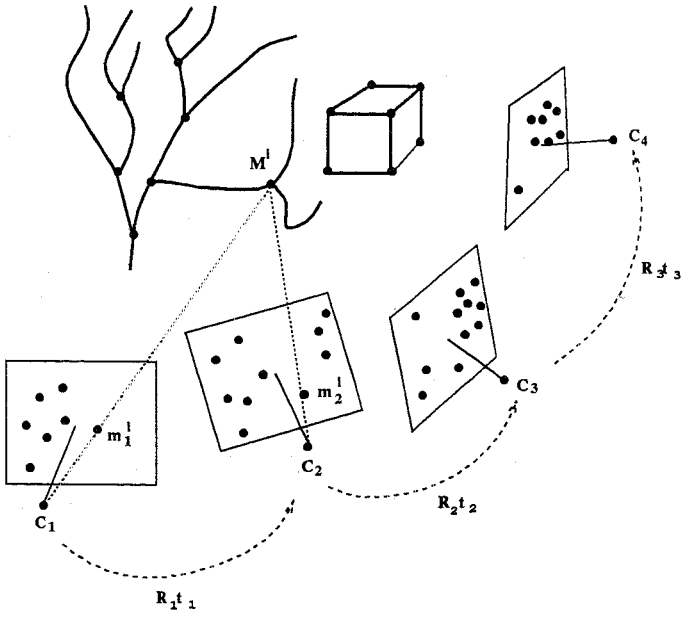


Fig. 1.3: Le principe de l'autocalibration

La figure 1.3 schématise le principe employé pour calibrer une caméra. Le calcul s'effectuera en deux étapes. Dans une première phase, qui est répétée, à partir de paires de points appariés entre des images successives, nous calculons la matrice fondamentale. Le nombre minimal d'appariements est de huit. Dans une seconde phase, à partir de trois matrices fondamentales, obtenues à partir de trois[9] mouvements (dont l'un peut être la composée des deux autres, ce qui fait que seules trois *vues* sont requises), les paramètres intrinsèques de la caméra sont calculés. Il est ensuite possible de calculer les paramètres extrinsèques des positions successives de la caméra, c'est-à-dire son mouvement, dans le repère associé à la position initiale de celle-ci. Si nous voulons, de plus, calibrer un système stéréo formé de deux caméras rigidement liées, nous calculons d'abord les paramètres intrinsèques puis extrinsèques de chaque caméra par la méthode précédente, donc uniquement à partir d'opérations menées indépendamment dans chaque caméra, avant d'obtenir la position relative des deux caméras du système, au moyen de deux des mouvements qui ont déjà été utilisés.

[9]deux mouvements seulement sont nécessaires si nous utilisons le modèle à quatre paramètres avec contrainte de perpendicularité.

1.3 Comment obtenir les appariements ?

1.3.1 Les points d'intérêt

Outre le fait de fournir des coordonnées tridimensionnelles de référence, l'utilisation d'une mire de calibration peut permettre de mesurer les coordonnées bidimensionnelles des images de ces points de référence avec une précision et une rapidité excellentes, dus à la simplicité des points à caractériser et à la connaissance exacte de leurs caractéristiques. Par exemple Tsai [160] annonce des localisations précises à $\frac{1}{30}$ de pixel obtenues à partir de cercles, et ce avec des temps de calcul très rapides (65 ms pour 36 points). Avec le modèle de mire utilisé dans notre équipe, dont les points caractéristiques sont constitués par les coins de carrés, la détection est plus délicate, mais les algorithmes de Vaillant [165] permettent néanmoins d'atteindre une précision de l'ordre de $\frac{1}{10}$ de pixel. Récemment, Peuchot [127] (voir aussi [83] pour des expérimentations) a annoncé une précision de localisation de l'ordre de quelques centièmes de pixels obtenue en utilisant un réseau de droites orthogonales et un modèle de l'intensité lumineuse.

Le premier problème à résoudre en vue d'effectuer une auto-calibration, par l'approche que nous étudions, fondée sur les correspondances de points, est donc de trouver des points caractéristiques dans les images. Comme nous ne voulons utiliser aucune information *a priori* sur les scènes, seules les variations d'intensité de l'image peuvent permettre d'obtenir ces points, dits alors *points d'intérêt généraux*, qui sont principalement de deux types: les coins et les jonctions.

Le problème de la détection des points d'intérêt a suscité un grand nombre de travaux[10] que nous pouvons classer en trois grandes catégories d'approches, dont le principe est de:

1. Partir de contours, et chaîner ceux-ci avant de chercher les points de courbure maximale ([5] [109] [116]), ou effectuer une approximation polygonale en vue d'en déduire les points d'intersection. [131]

2. Appliquer directement des opérateurs dans l'image. Une grande famille d'approches, utilisant des opérateurs de dérivation du second ordre, est fondée sur la détection des maxima locaux et le seuillage du déterminant du hessien H [10], de la courbure gaussienne K [32]

$$\det H = I_{xx}I_{yy} - I_{xy}^2 \qquad K = \frac{\det H}{(1 + I_x^2 + I_y^2)^2}$$

ou de variantes de celle-ci [78, 177], dont il été montré [121, 138] qu'elles présentent certaines équivalences. Une seconde formulation [125], qui a des liens étroits avec la précédente, mais ne fait intervenir que des dérivées premières, est de considérer les maxima du coefficient $\frac{\text{trace}C}{\det C}$, quotient du carré de la norme du gradient par la courbure [59]

[10] Pour une revue détaillée et une bibliographie abondante, voir [31]

(un perfectionnement [58] est: $\det C - 0.04\,\mathrm{trace}^2 C$) où C est la matrice:

$$C = \begin{pmatrix} \hat{I}_x^2 & \hat{I_x I_y} \\ \hat{I_x I_y} & \hat{I}_y^2 \end{pmatrix}$$

$\hat{I}$ désignant l'opération de lissage sur I.

3. Utiliser des modèles analytiques construits à partir de quelques paramètres géométriques (par exemple pour un coin: les coordonnées du sommet, les angles des deux droites, les niveaux de gris à l'intérieur et à l'extérieur du coin), et d'un filtre permettant de modéliser le lissage de l'image au cours du processus d'acquisition[11]. Le filtre retenu préférentiellement est le filtre exponentiel [27] qui possède des propriétés analogues au filtre gaussien, mais est intégrable formellement, à la différence de celui-ci. Ce filtre dépend d'un seul paramètre, la largeur de bande. Cette approche a été utilisée d'un point de vue théorique pour analyser le comportement de détecteurs de coins aussi bien que de jonctions par Deriche et Giraudon [30, 52, 31], et mise en oeuvre par Rohr [133] (filtre gaussien) ainsi que Blaszka et Deriche (filtre exponentiel) pour construire un détecteur opérationnel [13, 28]. L'idée est d'effectuer l'estimation de l'ensemble des paramètres par minimisation aux moindres carrés, sur une fenêtre réduite, de l'écart en chaque point entre le niveau de gris calculé à partir du modèle et celui mesuré dans l'image.

Les méthodes de la première catégorie ne sont pas considérées comme les plus précises, présentant l'inconvénient de dépendre d'une, ou de plusieurs étapes préliminaires. Si les chaînes de pixels extraites dans les différentes images ne sont pas identiques, ce qui se produit souvent, comme la pratique de la stéréo l'a mis en évidence [165], les points d'intérêt extraits ne seront pas exactement correspondants. Les approches de la seconde catégorie ne détectent pas la position exacte du point d'intérêt, l'effet de lissage entraînant une délocalisation qui est souvent supérieure au pixel, comme l'ont montré Deriche et Giraudon. Les approches du troisième type permettent d'obtenir une précision sub-pixellique, cependant leur mise en oeuvre automatique demeure beaucoup plus délicate car il est supposé que les fenêtres de recherche contiennent effectivement la primitive recherchée, et uniquement elle. De plus le calcul est également plus coûteux, puisqu'il s'agit d'effectuer une minimisation non-linéaire.

Nous avons donc actuellement le choix entre d'une part des détecteurs généraux de points d'intérêt relativement rapides et simples à utiliser (il s'agit d'appliquer des opérateurs), mais d'une précision moins bonne que le pixel, et d'autre part des détecteurs nettement plus lents et plus complexes mais qui donnent la précision subpixellique requise pour l'autocalibration. Une méthode pleinement opérationnelle d'extraction de points d'intérêts généraux très précis verra sans doute prochainement le jour, mais il s'agit pour l'instant d'un problème non résolu.

[11] Un modèle à trois paramètres avait été considéré par Guiducci [56], qui en avait tiré des expressions analytiques.

1.3.2 La mise en correspondance

Une fois que les points d'intérêt sont obtenus, nous disposons de deux types de techniques éprouvées pour les mettre en correspondance. La première est assez coûteuse, mais fonctionne à partir de deux vues relativement éloignées. La seconde, plus efficace, suppose que nous disposions d'une assez grande séquence de vues suffisamment proches.

Corrélation Le principe de cette technique de mise en correspondance de points, qui a été très largement utilisée, aussi bien pour des systèmes opérationnels de stéréophotogrammétrie [77, 46] qu'en vision par ordinateur [51], est de comparer les voisinages des points candidats, en mesurant leur similarité par un coefficient de corrélation. Les voisinages sont représentés par deux fenêtres (dites de *corrélation*) identiques, en général carrées et centrées. A partir d'un point $M_1(u_1, v_1)$ de l'image 1, le point $M_2(u_2, v_2)$ est sélectionné parmi ceux de l'image 2 se trouvant dans une *fenêtre de recherche* de taille P, comme étant celui qui conduit à la plus grande valeur de l'expression classique de la corrélation normalisée:

$$C(M_1, M_2) = \frac{\sum_{i=-N}^{N}\sum_{j=-N}^{N}(I_1(u_1+i, v_1+j) - \overline{I_1(u_1, v_1)})\cdot(I_2(u_2+i, v_2+j) - \overline{I_2(u_2, v_2)})}{\sqrt{\sum\sum(I_1(u_1+i, v_1+j) - \overline{I_1(u_1, v_1)})^2 \cdot \sum\sum(I_2(u_2+i, v_2+j) - \overline{I_2(u_2, v_2)})^2}}$$

où $I(u, v)$ est l'intensité en (u, v), et $\overline{I(u, v)}$ la moyenne des intensités sur la fenêtre de corrélation $[u - N, u + N] \times [v - N, v + N]$ de ce point.

Dans une première phase, la corrélation s'effectue sur toute l'image, mais en ne testant que les points d'intérêt entre eux, à l'aide de fenêtres de recherche centrées en (x_1, y_1), et de fenêtres de corrélation toutes deux assez grandes. Plusieurs contraintes additionnelles permettent d'éliminer les faux appariements:

- un seuillage sur le score $C(M_1, M_2)$
- un seuillage sur l'angle que fait le gradient en M_1 avec le gradient en M_2
- le principe du retour: la même technique de corrélation est appliquée au meilleur point M_2 trouvé dans la seconde image, elle doit permettre de retrouver le point M_1 initial

Les seuils sont d'abord choisis sévèrement et conduisent à des appariements en nombre faible, mais qui sont fiables. Une estimée de la matrice fondamentale est alors calculée à partir de ces appariements. Elle permet une nouvelle recherche par corrélation plus efficace. Chaque point d'intérêt $M_1(u_1, v_1)$ de la première image est testé seulement avec chacun des points d'intérêt de l'image 2 contenu dans une bande centrée sur l'épipolaire de $M_1(u_1, v_1)$. L'espace de recherche étant cette fois beaucoup plus réduit et mieux localisé, il est possible d'utiliser des seuils moins rigoureux, ce qui permet d'obtenir des correspondances supplémentaires [134]. Il est même possible d'utiliser un algorithme de stéréovision comme celui de Robert [132] qui est formulé uniquement en termes de matrices fondamentales, et non de matrices de projection.

Il faut noter que l'opération de corrélation elle-même fournit des informations suffisamment riches pour permettre de trouver des correspondances de points qui n'ont pas été extraits au préalable comme points d'intérêt.

Suivi Le principe de cette technique est de suivre les points d'intérêt dans une séquence d'images prises au cours d'un mouvement de la caméra. Une approche par prédiction et appariement, développée avec succès par Deriche et Faugeras [29] pour le suivi de segments de droite, est également adaptée au suivi de points. Ceux-ci sont représentés par une liste de paramètres et de leurs dérivées temporelles jusqu'à l'ordre deux. Pour chaque image, et pour chaque point détecté,

1. Un modèle cinématique (accélération constante ou vitesse constante) est assigné à chaque paramètre.

2. Le modèle est utilisé pour prédire la valeur de chaque paramètre dans l'image suivante, ainsi que l'incertitude associée, par un filtre de Kalman dynamique. Il est important de tenir compte du fait que le modèle n'est correct que localement, soit en introduisant un bruit dans le modèle lui-même ($\alpha - \beta$ tracker), soit en éliminant les vieilles mesures.

3. L'incertitude est utilisée pour déterminer une zone de recherche plausible autour de la position prédite.

4. A l'intérieur de celle-ci, la correspondance est effectuée à l'aide d'une fonction de similarité fondée sur la *distance de Mahalanobis* entre les attributs

5. Si un appariement satisfaisant est trouvé, il est utilisé comme nouvelle mesure pour remettre à jour le modèle cinématique.

A l'instant initial, aucune information n'étant disponible sur la cinématique des attributs présents dans l'image, nous devons supposer des vitesses et accélérations nulles avec une grande incertitude. La pratique a montré qu'alors les faux appariements disparaissent en général au bout de trois itérations. Cette approche s'est révélée très efficace sur des longues séquences d'images réelles. Son inconvénient est de nécessiter un nombre assez élevé de vues, qui doivent toutes être traitées pour en extraire les points d'intérêt. Toutes ces vues ne seront pas utilisables pour la suite car, comme nous le verrons plus tard, les mouvements de faible amplitude de la caméra ne conduisent pas à des estimations suffisamment précises de la matrice fondamentale.

1.4 Contributions au problème de la calibration

L'origine de nos travaux L'idée consistant à lier les paramètres intrinsèques par les équations de Kruppa en observant des points dans l'environnement et à remarquer que deux ou trois mouvements suffisent *en théorie* à calculer ces paramètres, est due à Maybank et

Faugeras [107]. Dans les sous-sections 3.1.1 et 7.2.2, nous avons repris, sous une forme un peu différente, l'exposition de ces auteurs. Notre travail a consisté à explorer l'applicabilité de leur résultat, qui a été obtenu sous forme de théorème de géométrie algébrique, résumé en 8.1. Nous avons constaté qu'afin de l'utiliser il faut changer totalement de démonstration, la méthode suggérée dans leur preuve n'étant pas stable. Nous sommes donc partis de cette idée, et l'avons ensuite considérablement approfondie et affinée afin de pouvoir l'appliquer réellement. Nous avons dû pour cela introduire un cadre et une formulation qui sont totalement nouveaux, et ont fait apparaître d'autres problèmes très importants, que nous avons étudiés en détail, aussi bien d'un point de vue théorique qu'expérimental.

Les notions de géométrie épipolaire (2.1.1,2.1.4,2.2.1), algorithme des huit points (4.1.1), homographie entre plans (5.1.1), calcul d'incertitude (6.1.1,6.1.2) surface critique (6.2.1), conique absolue (7.2.1), méthode des continuations (8.2.1), décomposition de la matrice essentielle et calcul du mouvement en coordonnées normalisées (9.1.1) sont déjà connues. Nous les avons néanmoins exposées sous une forme parfois originale. Mis à part ces sous-sections, ainsi que ce chapitre d'introduction, le matériel du reste du manuscrit est nouveau. Nous détaillons à présent ces contributions personnelles, qui ont fait l'objet de quelques publications, dont nous donnons les références.

La matrice fondamentale D'un point de vue théorique, nous apportons une clarification certaine en résumant sous une forme concise et opératoire un certain nombre de notions liées à la géométrie épipolaire, en introduisant:

- la notion de matrice fondamentale, son lien avec la matrice essentielle et avec les paramètres de la transformation épipolaire [40, 92].

Une grande partie de la thèse est ensuite consacrée au problème de la détermination robuste de cette matrice. Nous proposons:

- la contrainte d'ordre épipolaire,
- plusieurs algorithmes de détermination des épipoles fondés sur l'invariance du birapport par la transformation épipolaire,
- plusieurs méthodes de détermination directe de la matrice fondamentale, leur analyse et leur comparaison [92],
- dans le cadre d'une configuration spatiale particulière constituée de plans, la mise en évidence de l'instabilité, ainsi que des algorithmes spécifiques [98],
- plusieurs algorithmes de calcul d'homographies et de transformations quadratiques,
- une étude théorique et expérimentale de l'influence des mouvements sur la stabilité, faisant intervenir notamment calculs de covariance et surface critique.

L'autocalibration D'un point de vue théorique, nous mettons en évidence, outre quelques propriétés des équations de Kruppa:

- l'équivalence des différentes formulations de la contrainte de rigidité du mouvement, et leur important lien avec la matrice fondamentale,
- les liens algébriques qui existent entre la conique absolue, les paramètres intrinsèques réels, et les coefficients obtenus en résolvant les équations de Kruppa, ainsi que les contraintes qui leur sont liées.

Nous proposons ensuite:

- une méthode semi-analytique [40] et des méthodes itératives [95] de résolution des équations de Kruppa,
- une étude comparative sur la détermination du mouvement tridimensionnel d'une caméra, tenant en particulier compte de l'influence d'erreurs sur les paramètres intrinsèques,
- une méthode directe de calcul simultané des paramètres intrinsèques et du mouvement,
- un algorithme de calcul des paramètres extrinsèques d'un système de vision stéréoscopique à partir de mouvements calculés de manière monoculaire [96, 97].

Chapitre 2

Relations géométriques et matrice fondamentale

La *matrice fondamentale* est un concept-clé lorsque des images non calibrées, issues de l'utilisation de deux points de vue, sont utilisées. Elle contient l'ensemble de l'information géométrique reliant les deux vues. Nous commençons par définir précisément cette matrice, et montrons clairement sa relation avec la géométrie épipolaire et avec la *matrice essentielle* introduite antérieurement par Longuet-Higgins. Nous montrons de plusieurs manières que la matrice fondamentale, définie à un facteur d'échelle près, doit être de rang deux. Nous présentons ensuite une propriété importante de la géométrie épipolaire, l'existence d'une homographie entre les faisceaux de droites épipolaires. Conjointement aux coordonnées des épipoles, la connaissance de cette homographie détermine totalement la géométrie épipolaire. Au moyen de plusieurs paramétrisations, nous explicitons la relation bijective entre les sept paramètres de la transformation épipolaire et la matrice fondamentale, ce qui met en évidence le fait que cette matrice constitue une expression naturelle des contraintes géométriques dues à l'utilisation de deux points de vue. Le cas du calcul approché est abordé, puis nous présentons une nouvelle contrainte sur les positions possibles des épipoles, obtenue à partir d'une relation d'ordre sur les droites épipolaires d'une image.

2.1 Matrice fondamentale et contraintes géométriques

Nous présentons deux contraintes géométriques liées à l'utilisation de deux points de vue. La première est la contrainte épipolaire, utilisée principalement en stéréovision. La seconde, la contrainte de Longuet-Higgins, a été introduite pour l'étude du mouvement. A l'aide du concept de *matrice fondamentale* que nous introduisons, nous montrons que ces contraintes sont équivalentes. Pour finir, nous explicitons le lien entre mouvement et géométrie épipolaire.

2.1.1 La contrainte épipolaire

Considérons deux caméras observant la même portion de scène. Ces deux caméras peuvent être deux caméras physiquement distinctes, comme dans le cas de la stéréovision, ou une seule caméra observant la scène à partir de deux positions différentes obtenues par déplacement. Par paire de points correspondants entre deux images, nous désignons les paires comprenant un point de chacune des images, qui sont les projections d'un même point de l'espace par les caméras.

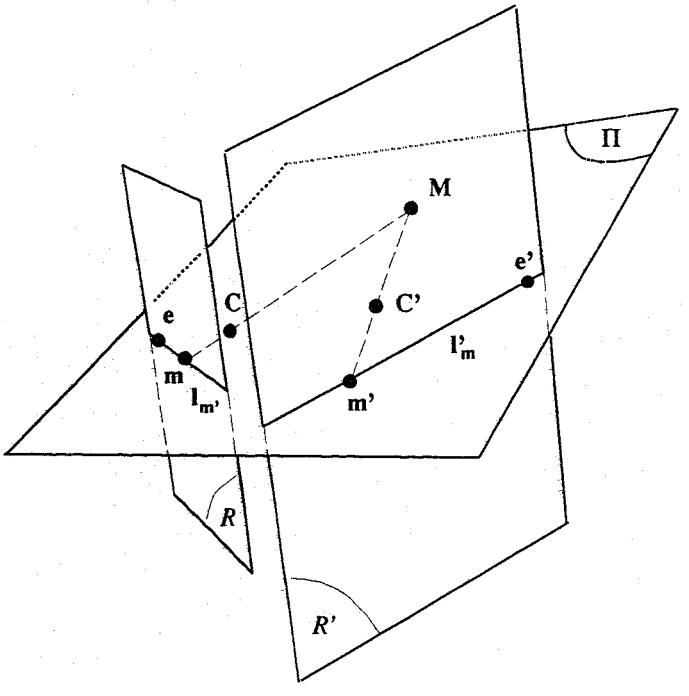


Fig. 2.1: La contrainte épipolaire

Faisons une remarque à l'aide de la figure 2.1. Nous pouvons y voir qu'étant donné $\mathbf{m}$ dans le plan image $\mathcal{R}$, l'ensemble des points physiques $\mathbf{M}$ qui ont pu produire $\mathbf{m}$ sont sur la demi-droite $\mathbf{Cm}$. Aussi, tous les correspondants possibles $\mathbf{m}'$ de $\mathbf{m}$ dans le plan image $\mathcal{R}'$ sont situés sur l'image donnée par la seconde caméra de cette demi-droite, qui est une

demi-droite l'_m passant par le point $\mathbf{e}'$, intersection de la droite $\mathbf{CC}'$ et du plan $\mathcal{R}'$. Le point $\mathbf{e}'$ est appelé épipole de la seconde caméra par rapport à la première caméra. La droite l'_m est appelée droite épipolaire du point $\mathbf{m}$ dans le plan image $\mathcal{R}'$ de la seconde caméra. De manière symétrique, on définit l'épipole $\mathbf{e}$ de la première caméra par rapport à la seconde caméra comme intersection de la droite $\mathbf{CC}'$ et du plan $\mathcal{R}$. Il est à noter que les droites épipolaires $l_{m'}$ et l'_m sont les traces sur les plans rétiniens $\mathcal{R}$ et $\mathcal{R}'$ d'un plan Π, appelé plan épipolaire de $\mathbf{M}$, contenant les centres optiques $\mathbf{C}$ et $\mathbf{C}'$, les épipoles $\mathbf{e}$ et $\mathbf{e}'$, et le point $\mathbf{M}$. La *contrainte épipolaire* est que, étant donné un point $\mathbf{m}$ dans le plan image $\mathcal{R}$, ses correspondants possibles dans le plan $\mathcal{R}'$ se situent sur la droite épipolaire l'_m.

2.1.2 La matrice fondamentale

Nous nous plaçons à présent en coordonnées rétiniennes[1]. La seule hypothèse que nous faisons est l'existence d'une relation linéaire projective entre les repères rétiniens de chaque caméra et un repère de l'espace. La relation entre un point $\mathbf{q}$ de la première image et sa droite épipolaire $\mathbf{l}'_q$ se décompose alors en deux relations linéaires projectives: une relation entre $\mathbf{q}$ et le rayon optique $\mathbf{Cm}$ dont il est l'image et une relation entre la droite $\mathbf{Cm}$ et sa projection $\mathbf{l}'_q$. Donc la relation entre $\mathbf{q}$ et $\mathbf{l}'_q$ est linéaire projective. Une manière plus précise de parvenir à ce résultat est d'utiliser les matrices de projection 3×4 $\mathbf{P} = [\hat{\mathbf{P}}\ \mathbf{p}]$ et $\mathbf{P}' = [\hat{\mathbf{P}}'\ \mathbf{p}']$, décomposées chacune en matrice 3×3 et vecteur 3×1. Le centre optique de la première caméra est $\mathbf{C} = -\hat{\mathbf{P}}^{-1}\mathbf{p}$, le point à l'infini de la droite $\mathbf{Cm}$ est $\hat{\mathbf{P}}^{-1}\mathbf{q}$. La représentation projective $\mathbf{l}'_q$ de la droite qui nous intéresse est obtenue à partir de la projection de ces deux points dans le second plan rétinien: $\mathbf{l}'_q = -\mathbf{P}'[\hat{\mathbf{P}}^{-1}\mathbf{p}\ ,\ 1] \wedge \hat{\mathbf{P}}'\hat{\mathbf{P}}^{-1}\mathbf{q}$. Cette dernière relation est linéaire en $\mathbf{q}$.

Nous appelons la matrice 3×3 décrivant cette relation la *matrice fondamentale*, ou F-matrice. Elle joue un rôle central dans tous les problèmes abordés dans notre thèse. S'agissant d'une quantité projective, il est clair que $\mathbf{F}$ *n'est définie qu'à une facteur d'échelle près.* Au sein du groupe projectif, les collinéations, ou homographies forment un sous-groupe. Elles transforment un point en un point, ou une droite en une droite. Les transformations duales des collinéations, les corrélations, transforment un point en une droite, ou une droite en un point. La matrice fondamentale est formellement une matrice de corrélation, cependant contrairement à celles-ci, elle n'est pas inversible. En effet la correspondance entre points de la première image et droites de la seconde image n'est pas bijective: l'épipole $\mathbf{e}$ admet pour droite épipolaire toute droite $\mathbf{l}'_q$ passant par $\mathbf{e}'$. L'ensemble de ces droites, étant un faisceau de droites du plan, est de dimension un, aussi nous en déduisons que *le rang de* $\mathbf{F}$ *est deux.* Le seul cas dégénéré est celui où $\mathbf{F}$ est nulle. Il se produit lorsque le mouvement entre les deux caméras est une rotation autour d'un axe passant par le centre optique. Les centres optiques de la première et de la seconde caméra sont alors confondus et il n'y a pas de contrainte épipolaire.

[1] Rappelons que par convention, les notations $\mathbf{p}$ et $\mathbf{q}$ désignent les coordonnées rétiniennes, par opposition aux notations $\mathbf{m}$ qui désignent les coordonnées normalisées.

Ce sont les deux seules propriétés caractéristiques des matrices fondamentales. Une manière de le voir est d'anticiper sur la section suivante dans laquelle nous montrons qu'il y a une correspondance bijective entre matrices fondamentales et transformations épipolaires. Une matrice fondamentale est définie par la donnée de sept paramètres au maximum. Or une transformation épipolaire est définie par sept paramètres, qui sont indépendants, puisqu'il est toujours possible de choisir deux épipoles et une homographie pour la définir. Les sept paramètres d'une matrice fondamentale doivent donc être également indépendants dans le cas général.

L'importance de la matrice fondamentale, qui a été négligée dans la littérature, vient du fait qu'elle contient l'ensemble des informations géométriques reliant deux images, qui sont accessibles uniquement à partir de ces images. Le fait qu'il ne soit pas possible d'obtenir d'autres informations à partir de correspondances ponctuelles peut se comprendre facilement si nous interprétons la position des points le long les épipolaires en termes de disparité [132]. Comme deux vues non calibrées ne déterminent pas la structure tridimensionnelle (euclidienne) d'une scène observée, il n'est pas possible d'obtenir d'information de profondeur, et donc de disparité. Nous ne pouvons donc contraindre les points correspondants qu'à se situer sur les droites épipolaires, la position sur celle-ci demeurant indéterminée. Néanmoins l'information contenue dans la matrice fondamentale est importante. Par exemple, si nous possédons deux vues d'une scène et que nous nous donnons les deux matrices fondamentales reliant chacune de ces vues à une troisième vue, alors nous pouvons prédire la position dans la troisième vue de tout point observé simultanément dans les deux vues initiales. Ce fait, déjà bien connu, puisqu'il constitue la base des algorithmes de stéréovision trinoculaire [99, 132], a été récemment formalisé en termes d'invariants [8].

Utilisons la matrice fondamentale pour exprimer la contrainte épipolaire, qui est que $\mathbf{q}'$, le point correspondant à $\mathbf{q}$, appartienne à $\mathbf{l}'_q = \mathbf{F}\mathbf{q}$. Il vient:

$$\mathbf{q}'^T\mathbf{F}\mathbf{q} = 0 \tag{2.1}$$

Notons que cette équation peut se réécrire:

$$\mathbf{q}^T\mathbf{F}^T\mathbf{q}' = 0$$

ce qui montre qu'en échangeant le rôle des deux images, $\mathbf{F}$ est changée en sa transposée.

Nous avons présenté une justification *purement géométrique* de l'équation fondamentale (2.1). Il est également possible d'en obtenir une justification *purement algébrique*[2]. Soit $\mathbf{P}$ et $\mathbf{P}'$, les deux matrices de projection, et $\mathbf{q}$ et $\mathbf{q}'$, les deux projections d'un point tridimensionnel $\mathbf{M}$. L'une des coordonnées projectives de $\mathbf{q}$, ainsi que de $\mathbf{q}'$, est non-nulle. Sans perte de généralité, nous pouvons supposer qu'il s'agit de la troisième coordonnée, c'est-à-dire que nous sommes dans le cas de points projetés à distance finie. Les équations de projection:

$$\mathbf{q} = \begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \mathbf{P}\mathbf{M} \qquad \mathbf{q}' = \begin{pmatrix} s'u' \\ s'v' \\ s' \end{pmatrix} = \mathbf{P}\mathbf{M}$$

[2]Celles existant dans la littérature se présentent sous forme hybride.

forment six équations scalaires qui peuvent alors s'écrire sous la forme:

$$\begin{pmatrix} \begin{bmatrix} & \mathbf{P} & \\ & & \end{bmatrix} & \begin{matrix} -u & 0 \\ -v & 0 \\ -1 & 0 \end{matrix} \\ \begin{bmatrix} & \mathbf{P}' & \\ & & \end{bmatrix} & \begin{matrix} 0 & -u \\ 0 & -v \\ 0 & -1 \end{matrix} \end{pmatrix} \begin{pmatrix} \begin{bmatrix} & \mathbf{M} & \\ & & \end{bmatrix} \\ \begin{matrix} s \\ s' \end{matrix} \end{pmatrix} = 0$$

L'existence d'une solution non-nulle entraîne la nullité du déterminant de la matrice 6×6 apparaissant au premier membre. Il est aisé de voir que ce déterminant s'écrit sous la forme (2.1).

2.1.3 Lien avec la contrainte de Longuet-Higgins

Considérons à présent le cas où nous sommes en présence d'une caméra unique effectuant un déplacement entre deux positions. Ce déplacement peut être décomposé d'une infinité de manières comme la composée d'une rotation d'axe fixé et d'une translation. Nous choisissons comme convention de prendre une rotation de matrice $\mathbf{R}$ autour d'un axe passant par le centre optique final $\mathbf{C}'$ (donc $\mathbf{R}$ envoie les vecteurs de base initiaux sur les finaux) suivi de la translation $\mathbf{t} = \mathbf{C}'\mathbf{C}$, exprimée dans le repère final[3]. Nous avons alors:

$$\mathbf{M}_{/final} = \mathbf{R}\mathbf{M}_{/initial} + \mathbf{t} \tag{2.2}$$

Plaçons-nous dans les repères normalisés. Le point $\mathbf{M}$ de l'espace a pour images $\mathbf{m}$ et $\mathbf{m}'$ dans les deux positions respectives. Comme le montre la figure 2.2, il est clair que $\mathbf{m}$ et $\mathbf{m}'$ sont en correspondance si, et seulement si les trois vecteurs $\mathbf{Cm}$, $\mathbf{C}'\mathbf{m}'$, et $\mathbf{t} = \mathbf{C}'\mathbf{C}$ sont coplanaires.

Un vecteur $\mathbf{x}$ exprimé dans le repère de la caméra au premier instant a pour coordonnées $\mathbf{Rx}$ dans le repère de la caméra au second instant, donc $\mathbf{Cm}_{/final} = \mathbf{RCm}_{/initial}$ La contrainte de coplanarité peut donc s'écrire, dans le repère final de la caméra:

$$\mathbf{C}'\mathbf{m}' \cdot (\mathbf{t} \wedge \mathbf{R}\,\mathbf{Cm}) = 0 \tag{2.3}$$

Une manière directe de retrouver cette relation très importante dite de Longuet-Higgins [90] est de multiplier la relation (2.2) par $\mathbf{C}'\mathbf{m}' \cdot (\mathbf{t} \wedge \mathbf{R})$.

Définissons la matrice antisymétrique $\mathbf{T}$ telle que $\mathbf{Tx} = \mathbf{t} \wedge \mathbf{x}$ pour tout vecteur $\mathbf{x}$:

$$\mathbf{T} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

[3]La convention inverse est également rencontrée dans la littérature, c'est-à-dire: $\mathbf{M}_{/initial} = \mathbf{R}'\mathbf{M}_{/final} + \mathbf{t}'$. Il suffit d'échanger les points des deux images, et le déplacement par son inverse pour retrouver nos conventions

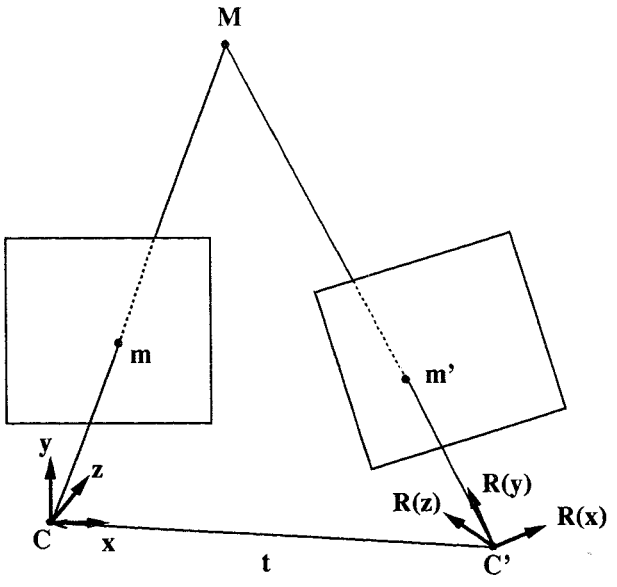


Fig. 2.2: La contrainte de Longuet-Higgins

En posant:

$$\mathbf{E} = \mathbf{TR} \tag{2.4}$$

l'équation (2.3) peut se réécrire:

$$\mathbf{C'm'}^T \mathbf{E}\, \mathbf{Cm} = 0 \tag{2.5}$$

La matrice $\mathbf{E}$ est appelée *matrice essentielle* et joue un rôle important dans les problèmes de détermination du mouvement. La matrice $\mathbf{E}$ étant le produit d'une matrice antisymétrique par une matrice de rotation, son rang est toujours de deux, sauf dans le cas où elle est nulle. Comme elle est définie par l'équation (2.5), elle n'est déterminée qu'à une constante multiplicative près. Une autre propriété des matrices essentielles mise en évidence par Faugeras et Huang est qu'elles doivent satisfaire les trois propriétés équivalentes:

$$\begin{gathered} \text{les deux valeurs singulières non nulles de } \mathbf{E} \text{ sont égales} \\ (\mathbf{l}_1^2 + \mathbf{l}_2^2 + \mathbf{l}_3^2)^2 = 4(\|\mathbf{l}_1 \wedge \mathbf{l}_2\|^2 + \|\mathbf{l}_1 \wedge \mathbf{l}_2\|^2 + \|\mathbf{l}_1 \wedge \mathbf{l}_2\|^2) \;,\quad \mathbf{l}_i \text{ vecteurs-ligne de } \mathbf{E} \\ \text{trace}^2(\mathbf{EE}^T) = 2\text{trace}((\mathbf{EE}^T)^2) \end{gathered} \tag{2.6}$$

Cette dernière condition jointe à celle sur le rang est nécessaire et suffisante pour assurer une décomposition de $\mathbf{E}$ sous la forme (2.4).

Revenons à l'équation de Longuet-Higgins et considérons cette fois les coordonnées *rétiniennes* $\mathbf{q}$, $\mathbf{q}'$, respectivement de $\mathbf{m}$ et $\mathbf{m}'$, qui sont celles auxquelles nous avons accès en pratique. Comme nous avons $\mathbf{q} = \mathbf{Am}$ and $\mathbf{q}' = \mathbf{A}'\mathbf{m}'$, $\mathbf{A}$ et $\mathbf{A}'$ étant les matrices des paramètres intrinsèques de la première et de la seconde caméra, respectivement, l'équation de Longuet-Higgins peut se réécrire:

$$\mathbf{q'}^T \mathbf{F} \mathbf{q} = 0 \tag{2.7}$$

où nous avons posé:

$$\mathbf{F} = \mathbf{A}'^{-1T}\mathbf{E}\mathbf{A}^{-1} \tag{2.8}$$

Les équations (2.7) et (2.1) étant identiques, nous avons ainsi montré que la contrainte de Longuet-Higgins est équivalente à la contrainte épipolaire, si nous prenons en compte la différence éventuelle des caméras. Nous avons de plus obtenu la relation très importante (2.8), qui permet de retrouver les deux propriétés de la matrice fondamentale: les matrices des paramètres intrinsèques $\mathbf{A}$ et $\mathbf{A}'$ étant inversibles, le rang de $\mathbf{F}$ est celui de $\mathbf{E}$, donc deux, et $\mathbf{F}$, comme $\mathbf{E}$, est définie à un facteur d'échelle près. Il faut noter que dans la littérature, la matrice fondamentale est parfois considérée comme une forme de matrice essentielle dans les repères rétiniens. Cela n'est pas correct, puisqu'elle ne vérifie pas la condition (2.6) que nous venons de mentionner.

Le fait que la matrice fondamentale ne dépende que de sept paramètres indépendants montre qu'il n'est pas possible de calculer simultanément à partir de deux vues les paramètres intrinsèques *et* le mouvement, en l'absence d'autres informations. En effet, il y a cinq paramètres indépendants pour le mouvement (trois pour la rotation, deux pour la direction[4] de la translation) et cinq paramètres intrinsèques dans le cas général. On peut donc obtenir au mieux deux équations reliant paramètres intrinsèques et coefficients de la matrice fondamentale. Ces équations sont les relations de Kruppa qui seront présentées ultérieurement. A partir du simple décompte précédent, nous voyons qu'il faut trois mouvements pour déterminer les paramètres intrinsèques dans le cas général, et deux mouvements dans le cas du modèle simplifié à quatre paramètres.

2.1.4 Lien entre la géométrie épipolaire et le mouvement

Il est aussi intéressant de remarquer que l'équation de Longuet-Higgins suppose un déplacement rigide entre les repères des deux caméras; ce ne semble pas être le cas de la contrainte épipolaire écrite en repères rétiniens, dont nous avons vu qu'elle ne suppose que l'existence d'une relation linéaire projective entre les repères rétiniens respectifs des deux caméras et le repère du monde. Autorisons-nous à considérer deux caméras différentes. Soit $\mathbf{P}$, et $\mathbf{P}'$, les matrices de projection. Nous avons vu que $\mathbf{P}$ (resp. $\mathbf{P}'$) peut se décomposer en une matrice de déplacement $\mathbf{D}$ (resp. $\mathbf{D}'$) et une matrice de paramètres intrinsèques $\mathbf{A}$ (resp. $\mathbf{A}'$), sauf dans un cas dégénéré qui ne peut pas se produire physiquement. Il existe donc bien un déplacement rigide entre le repère *normalisé* de la première caméra et celui *normalisé* de la seconde, sa matrice est donnée par $\mathbf{D}^{-1}\mathbf{D}'$.

Nous montrons pour finir que si les paramètres intrinsèques sont connus, la donnée de la géométrie épipolaire et celle du mouvement sont équivalentes en donnant explicitant les relations. Leur intérêt est d'une part de permettre de mener une étude théorique des problèmes liés aux choix des mouvements, d'autre part de calculer le mouvement une fois que les paramètres intrinsèques ont été déterminés.

[4] Il est connu que la norme de la translation ne peut être déterminée à partir des images, en raison de l'ambiguïté échelle-vitesse.

L'épipole $\mathbf{e}'$ dans la seconde image est obtenu en écrivant à l'aide de la relation (2.7) que les droites épipolaires de la seconde image, correspondant à tous les points de la première image, passent toutes par $\mathbf{e}'$:

$$\forall \mathbf{q} \qquad \mathbf{e}'^T \mathbf{F} \mathbf{q} = 0$$

Nous déduisons de cette relation, $\mathbf{e}'^T \mathbf{F} = 0$, et donc $\mathbf{F}^T \mathbf{e}' = 0$. Cette dernière relation se développe, en utilisant (2.8) en:

$$-\mathbf{A}^{-1T} \mathbf{R}^T \mathbf{T} \mathbf{A}'^{-1} \mathbf{e}' = 0$$

Dans cette relation, les matrices $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{R}$ sont toutes inversibles, alors que la matrice $\mathbf{T}$ est de rang deux et vérifie $\mathbf{T}\mathbf{t} = 0$. Cette relation entraîne donc $\mathbf{A}'^{-1}\mathbf{e}' = \mathbf{t}$, ce qui donne $\mathbf{e}'$. De manière similaire, l'épipole e dans la première image est obtenu en écrivant que les droites épipolaires de la première image, correspondant à tous les points de la seconde image, passent toutes par $\mathbf{e}$:

$$\forall \mathbf{q}' \qquad \mathbf{e}^T \mathbf{F}^T \mathbf{q}' = 0$$

Nous en déduisons $\mathbf{F}\mathbf{e} = 0$, qui s'écrit aussi

$$\mathbf{A}'^{-1T} \mathbf{T} \mathbf{R} \mathbf{A}^{-1} \mathbf{e} = 0$$

d'où $\mathbf{R}\mathbf{A}^{-1}\mathbf{e} = \mathbf{t}$. Nous obtenons finalement les deux épipoles:

$$\mathbf{e} = \mathbf{A}\mathbf{R}^T\mathbf{t} \qquad \mathbf{e}' = \mathbf{A}'\mathbf{t} \tag{2.9}$$

Ces dernières relations montrent que si nous nous plaçons en coordonnées normalisées, dans le cas où il n'y a pas de rotation, les épipoles ne sont autres que les foyers d'expansion[5] bien connus dans le cadre de la détermination et de l'interprétation du flot optique.

2.2 La transformation épipolaire

Nous présentons dans cette section une propriété importante de la géométrie épipolaire, l'existence d'une homographie entre les faisceaux de droites épipolaires. Conjointement aux coordonnées des épipoles, la connaissance de cette homographie détermine totalement la géométrie épipolaire. Nous montrons comment la transformation épipolaire peut être déterminée à partir de la matrice fondamentale, et vice-versa, ce qui met en évidence le fait que cette matrice constitue une expression naturelle des contraintes géométriques dues à l'utilisation de deux points de vue.

2.2.1 Une remarque géométrique

Considérons les ensembles de droites épipolaires dans la première et dans la seconde image. Passant par $\mathbf{e}$ (respectivement $\mathbf{e}'$), ils forment un faisceau de droites. Soit $\mathbf{M}$ un point de

[5] FOE: focus of expansion

l'espace, $\mathbf{m}$ et $\mathbf{m}'$ ses deux images. $\mathbf{M}$, $\mathbf{C}$ et $\mathbf{C}'$ définissent le plan épipolaire Π. Ces plans forment aussi un faisceau de plans, passant par $\mathbf{CC}'$. Le plan Π se projette en une droite épipolaire $l = \mathbf{em}$ dans la première image et en une droite épipolaire $l' = \mathbf{e}'\mathbf{m}'$ dans la seconde image. La correspondance $\Pi \overline{\wedge_1} l$ est une homographie entre le faisceau de droites épipolaires de la première image et le faisceau de plans épipolaires. En effet, si nous considérons quatre plans épipolaires coupant le plan image pour donner quatre droites épipolaires, par définition, le birapport des quatre plans est égal à celui des quatre droites. De même, il y a une seconde homographie $\Pi \overline{\wedge_2} l'$ entre le faisceau de droites épipolaires de la seconde image et le faisceau de plans épipolaires. Les deux faisceaux de droites épipolaires sont donc en correspondance homographique par la relation $l \overline{\wedge_1}^{-1} \overline{\wedge_2} l'$. Cette relation demeure une homographie en coordonnées rétiniennes, car le changement de repère rétinien est linéaire projectif. Nous appelons *transformation épipolaire* la relation entre les droites épipolaires.

La transformation épipolaire est donc définie par les coordonnées projectives (e_1, e_2, e_3) et (e'_1, e'_2, e'_3) des épipoles $\mathbf{e}$ et $\mathbf{e}'$, et les quatre paramètres a,b,c,d, de l'homographie entre les faisceaux d'épipolaires. S'agissant de quantités projectives, nous sommes en présence de sept valeurs indépendantes, deux pour chaque épipole, et trois pour l'homographie. C'est exactement le nombre de paramètres de la matrice fondamentale, et nous allons montrer comment décomposer celle-ci pour obtenir ces valeurs.

2.2.2 Le cas des épipoles finis

Nous commençons par un cas particulier qui a une grande importance dans la pratique. Nous paramétrons les faisceaux de droites épipolaires par le paramètre projectif obtenu par intersection avec la droite à l'infini $x_3 = 0$. Ceci suppose qu'aucun des épipoles ne se trouve à l'infini. Une droite l passant par l'épipole $\mathbf{e} = (e_1, e_2, e_3)^T$ et le point $\mathbf{q} = (q_1, q_2, q_3)^T$ a pour représentation projective $\mathbf{l} = \mathbf{e} \wedge \mathbf{q}$. Elle intersecte la droite à l'infini au point $\mathbf{l} \wedge (0, 0, 1)^T$ que nous pouvons écrire $(1, \tau, 0)^T$ avec

$$\tau = l_2/l_1 = \frac{e_3 q_1 - e_1 q_3}{e_3 q_2 - e_2 q_3} \tag{2.10}$$

Notons que ce paramétrage correspond au vecteur directeur de la droite $\mathbf{eq}$, et que donc dans le cas d'un épipole fini elle est bien définie, alors que dans le cas d'un épipole situé à l'infini elle n'a pas de sens, puisque les droites épipolaires sont parallèles dans au moins une des images. Les mêmes relations existent dans la seconde image pour la droite épipolaire correspondante l', et nous obtenons, avec des notations évidentes:

$$\tau' = l'_2/l'_1 = \frac{e'_3 q'_1 - e'_1 q'_3}{e'_3 q'_2 - e'_2 q'_3} \tag{2.11}$$

L'homographie entre l et l' peut alors s'écrire:

$$\tau' = \frac{a\tau + b}{c\tau + d} \tag{2.12}$$

Le cas exact En identifiant terme à terme, d'une part l'équation obtenue en substituant (2.10) et (2.11) dans (2.12) et d'autre part l'équation (2.1), nous obtenons les relations:

$$\begin{aligned} F_{11} &= be_3e'_3 \\ F_{12} &= ae_3e'_3 \\ F_{13} &= -ae_2e'_3 - be_1e'_3 \\ F_{21} &= -de'_3e_3 \\ F_{22} &= -ce'_3e_3 \\ F_{23} &= ce'_3e_2 + de'_3e_1 \\ F_{31} &= de'_2e_3 - be_3e'_1 \\ F_{32} &= ce'_2e_3 - ae_3e'_1 \\ F_{33} &= -ce'_2e_2 - de'_2e_1 + ae_2e'_1 + be_1e'_1 \end{aligned} \tag{2.13}$$

On peut vérifier à partir de ces relations d'une part que la matrice $\mathbf{F}$ n'est définie qu'a un facteur multiplicatif près (elle ne fait intervenir que des quantités projectives), d'autre part qu'elle est de rang au plus 2: $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$, les vecteurs colonne de $\mathbf{F}$, vérifient $e_1\mathbf{c}_1+e_2\mathbf{c}_2+e_3\mathbf{c}_3 = 0$. On montre facilement aussi que si e_3 et e'_3 ne sont pas nuls, alors le rang de $\mathbf{F}$ est exactement deux. En résolvant les équations (2.13), les éléments de la transformation épipolaire sont obtenus comme suit:

$$\begin{aligned} a &= F_{12} \\ b &= F_{11} \\ c &= -F_{22} \\ d &= -F_{21} \\ e_1 &= \frac{F_{23}F_{12} - F_{22}F_{13}}{F_{22}F_{11} - F_{21}F_{12}}e_3 \\ e_2 &= \frac{F_{13}F_{21} - F_{11}F_{23}}{F_{22}F_{11} - F_{21}F_{12}}e_3 \\ e'_1 &= \frac{F_{32}F_{21} - F_{22}F_{31}}{F_{22}F_{11} - F_{21}F_{12}}e'_3 \\ e'_2 &= \frac{F_{31}F_{12} - F_{11}F_{32}}{F_{22}F_{11} - F_{21}F_{12}}e'_3 \end{aligned} \tag{2.14}$$

On peut noter que F_{33} n'apparaît pas dans ces dernières relations. En effet, des neuf équations (2.13) seulement huit sont indépendantes, puisque la matrice $\mathbf{F}$ est singulière. Le déterminant de l'homographie épipolaire est $F_{22}F_{11} - F_{21}F_{12}$. Dans le cas des épipoles finis, il est non nul.

L'interprétation des équations (2.14) est simple: les coordonnées de $\mathbf{e}$ (resp. $\mathbf{e}'$) sont les vecteurs du noyau de $\mathbf{F}$ (resp. $\mathbf{F}^T$). En exprimant τ' en fonction de τ à partir de la relation $\mathbf{y}'_\infty\mathbf{F}\mathbf{y}_\infty = 0$ qui traduit juste le fait que les points à l'infini $\mathbf{y}_\infty = (1,\tau,0)^T$ et $\mathbf{y}'_\infty = (1,\tau',0)^T$, de droites correspondantes se correspondent, nous trouvons la bonne relation homographique. La situation est schématisée par la figure 2.3.

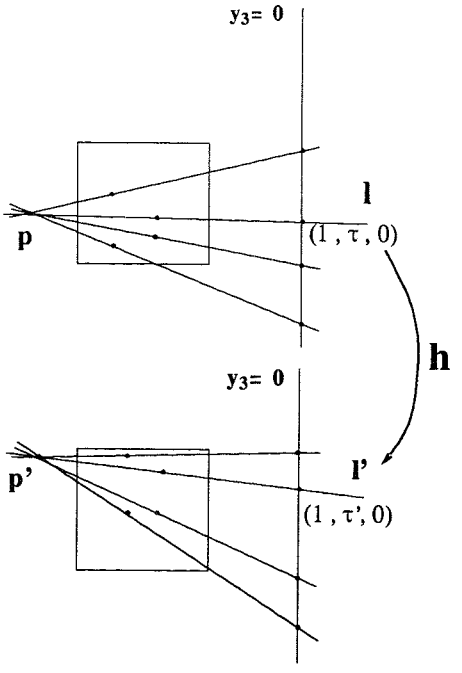


Fig. 2.3: La transformation épipolaire

Calcul approché Lorsque qu'une matrice fondamentale obtenue numériquement n'est pas exactement de rang 2, les relations énoncées précédemment ne sont pas vérifiées. Il n'y a pas de solution exacte à l'équation $\mathbf{Fe} = 0$. Nous pouvons dans ce cas calculer l'épipoles e en résolvant le classique problème de minimisation sous contrainte:

$$\min_{\mathbf{e}} \|\mathbf{Fe}\|^2 \quad \text{avec} \quad \|\mathbf{e}\|^2 = 1 \tag{2.15}$$

dont la solution est le vecteur propre unitaire de la matrice définie positive $\mathbf{F}^T\mathbf{F}$ correspondant à la plus petite valeur propre. L'épipole e′ est obtenu de la même manière en remplaçant $\mathbf{F}$ par $\mathbf{F}^T$. Une fois que les deux épipoles ont été obtenus, l'homographie peut être calculée à partir des coordonnées des correspondances au moyen des équations (2.10), (2.10), et (2.12). Celles-ci conduisent en effet à l'équation:

$$\begin{aligned}(e_y e'_x - e_y u'_i + v_i u'_i - v_i e'_x)a + (u_i u'_i - u_i e'_x - e_x u'_i + e_x e'_x)b \\ +(e_y v'_i + v_i e'_y - e_y e'_y - v_i v'_i)c + (e_x v'_i + u_i e'_y - u_i v'_i - e_x e'_y)d \;=\; 0\end{aligned} \tag{2.16}$$

où (u_i, v_i) (resp. (u'_i, v'_i)) sont les coordonnées d'un appariement dans la première (resp. seconde) image. Les équations (2.16) étant linéaires et homogènes, il nous suffit de trois appariements pour déterminer à un facteur d'echelle près les coefficients a, b, c et d. En pratique, nous possédons davantage de données et utilisons un moindres carrés linéaire.

Il est important de noter que la phase difficile est la détermination de la position des épipoles. Une fois que ceux-ci sont obtenus, le calcul de l'homographie, et donc de la matrice

fondamentale, est stable et n'entraîne que peu d'erreurs supplémentaires, même avec des correspondances bruitées, ce qu'illustre la figure 2.4, obtenue en moyennant 100 essais effectués selon une procédure décrite en 4.4: avec un pixel de bruit nous obtenons une erreur relative moyenne sur les quatre coefficients de l'homographie qui n'est que de 11%. Pour chaque essai, nous sommes partis des épipoles exacts et de 50 appariements bruités, et nous avons calculé l'homographie par la méthode linéaire décrite précédemment Ces résultats suggèrent donc que les quantités importantes à déterminer en vue de calculer la transformation épipolaire sont les épipoles. Aussi, le prochain chapitre sera consacré à une méthode dont le but est uniquement de calculer ceux-ci. Il s'agit, nous le verrons, d'un problème numériquement difficile.

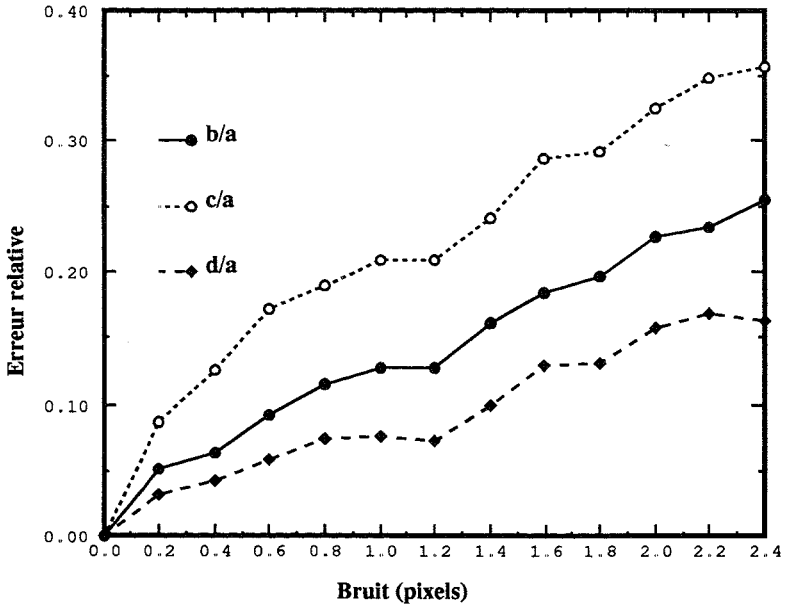


Fig. 2.4: Statistique: erreur relative sur les coefficients de l'homographie épipolaire

2.2.3 Le cas général

Examinons à présent le cas où au moins l'un des épipoles est à l'infini. Il n'est évidemment plus possible de paramétrer le faisceau d'épipolaires correspondant par son intersection avec la droite à l'infini, car cette intersection est toujours égale à l'épipole. Si les formules (2.13) sont appliquées, une matrice $\mathbf{F}$ de rang 1 est obtenue, ce qui est absurde. Réciproquement, les formules (2.14) ne déterminent pas les épipoles.

Nous généralisons les calculs précédents en paramétrant les faisceaux de droites épipolaires par intersection avec la droite de référence $d_r = \mathbf{p}_1\mathbf{p}_2$, $\mathbf{p}_1$ et $\mathbf{p}_2$ étant deux points fixés arbitrairement. Ce paramétrage ne sera dégénéré que si l'un des deux épipoles appartient à d_r. Les calculs présentés précédemment constituaient donc un cas particulier de ceux menés ici, dans lequel $\mathbf{p}_1 = (0,1,0)^T$ et $\mathbf{p}_2 = (1,0,0)^T$. Une droite l passant par l'épipole

$\mathbf{e} = (e_1, e_2, e_3)^T$ et le point $\mathbf{q} = (q_1, q_2, q_3)^T$ a pour représentation projective $\mathbf{l} = \mathbf{e} \wedge \mathbf{q}$. Elle intersecte la droite d_τ, dont une représentation projective est $\mathbf{p}_1 \wedge \mathbf{p}_2$, au point $\mathbf{l} \wedge (\mathbf{p}_1 \wedge \mathbf{p}_2)$, qui peut aussi s'écrire $\lambda \mathbf{p}_1 + \mu \mathbf{p}_2$, avec $\lambda = \mathbf{l} \cdot \mathbf{p}_2$ et $\mu = -\mathbf{l} \cdot \mathbf{p}_1$. Un paramètre projectif τ pour l est donc obtenu par: $\tau = \frac{\lambda}{\mu}$. Il vient:

$$\tau = -\frac{(q_2 e_3 - q_3 e_2) q_{21} + (q_3 e_1 - q_1 e_3) q_{22} + (q_1 e_2 - q_2 e_1) q_{23}}{(q_2 e_3 - q_3 e_2) q_{11} + (q_3 e_1 - q_1 e_3) q_{12} + (q_1 e_2 - q_2 e_1) q_{13}} \tag{2.17}$$

Nous écrivons les mêmes relations dans la seconde image en remplaçant $\mathbf{e}$ et $\mathbf{q}$ respectivement par $\mathbf{e}'$ et $\mathbf{q}'$, et nous obtenons de même le paramètre projectif τ' par une formule analogue. L'homographie entre l et l' peut alors s'écrire:

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

Elle dépend bien sûr cette fois du choix des points $\mathbf{p}_1$ et $\mathbf{p}_2$. En substituant les valeurs (2.17) de τ et τ' dans cette formule, on obtient une expression plus compliquée que précédemment, mais qu'il est également possible d'identifier à l'équation (2.1), ce qui conduit aux relations donnant $\mathbf{F}$ en fonction des épipoles, de l'homographie, et des deux points de référence, qui sont de la forme:

$$\begin{aligned} F_{ij} = {} & (\mathbf{p}_1 \wedge \mathbf{e}')_i (\mathbf{p}_2 \wedge \mathbf{e})_j a - (\mathbf{p}_1 \wedge \mathbf{e}')_i (\mathbf{p}_1 \wedge \mathbf{e})_j b \\ & + (\mathbf{p}_2 \wedge \mathbf{e}')_i (\mathbf{p}_2 \wedge \mathbf{e})_j c - (\mathbf{p}_2 \wedge \mathbf{e}')_i (\mathbf{p}_1 \wedge \mathbf{e})_j d \end{aligned} \tag{2.18}$$

Donnons également l'expression de $\mathbf{F}^*$, la matrice adjointe de $\mathbf{F}$, qui est simple et joue un rôle important pour le calcul de la transformation épipolaire:

$$F^*_{ij} = (bc - ad)(\mathbf{e}.(\mathbf{p}_1 \wedge \mathbf{p}_2))(\mathbf{e}' \cdot (\mathbf{p}_1 \wedge \mathbf{p}_2)) e_i e'_j \tag{2.19}$$

Le premier facteur est le déterminant de l'homographie épipolaire. Sauf dans le cas dégénéré où la matrice fondamentale est nulle, l'existence de la géométrie épipolaire assure sa non-nullité. Le second (respectivement troisième) facteur s'annule si, et seulement si l'épipole $\mathbf{e}$ (respectivement $\mathbf{e}'$) appartient à la droite d_τ choisie pour paramétrer l'homographie.

A partir de la relation (2.19), il est possible de déterminer simplement les épipoles $\mathbf{e}$ et $\mathbf{e}'$. Comme la matrice F est de rang deux, l'un au moins des coefficients de $\mathbf{F}^*$ est non nul. Notons son indice (i, j). Les coordonnées projectives de $\mathbf{e}$ sont alors le $j$ième vecteur colonne, celles de $\mathbf{e}'$ sont le $i$ième vecteur ligne. Le fait que l'élément F^*_{ij} soit non nul garantit que nous obtenons bien des points projectifs, puisque les coordonnées trouvées ne sont pas toutes nulles. Notons que le cas des épipoles finis correspond à $F^*_{33} \neq 0$. Une fois que les épipoles sont calculés à l'aide du cofacteur (i, j), il suffit de résoudre le système linéaire en les a, b, c, d obtenu à partir des quatre équations (2.18), chacune étant formée à partir de l'un des quatre coefficients de $\mathbf{F}$ qui ont servi à calculer ce cofacteur. Des calculs effectués en MAPLE, que nous ne détaillons pas, permettent de montrer que le déterminant de ce système est non nul, et l'homographie ainsi obtenue a également un déterminant non nul, sous réserve qu'aucun des deux épipoles ne se trouve sur la droite d_τ.

2.3 La contrainte d'ordre épipolaire

Nous montrons dans cette section que la connaissance des épipoles détermine une relation d'ordre sur les droites épipolaires de l'image. Etant donné des correspondances de points, cette relation d'ordre détermine à son tour des contraintes fortes sur les positions possibles des épipoles.

Montrons d'abord que chaque paire de droites épipolaires partitionne les plans rétiniens en deux demi-plans qui se correspondent. Soit l une droite épipolaire du plan rétinien $\mathcal{R}$ de la première image, et l' la droite épipolaire correspondante dans le plan rétinien $\mathcal{R}'$ de la seconde image. Le plan $\mathcal{R}$ est délimité par la droite l en deux demi-plans $\mathcal{R}_a$ et $\mathcal{R}_b$, et de même dans la seconde image avec des primes. Aucune hypothèse n'est faite sur $\mathcal{R}_a$, $\mathcal{R}_b$, $\mathcal{R}'_a$, $\mathcal{R}'_b$. Soit aussi m_1 et m'_1 deux points correspondants, situés respectivement dans les demi plans $\mathcal{R}_a$ et $\mathcal{R}'_a$ Nous allons montrer que si m_2 est un point de $\mathcal{R}_a$, alors son correspondant m'_2 se trouve nécessairement dans $\mathcal{R}'_a$. Considérons le plan épipolaire Π dont l et l' sont les traces sur $\mathcal{R}$ et $\mathcal{R}'$. Il délimite deux demi-espaces $\mathcal{E}_c$ et $\mathcal{E}_d$, et contient les centres optiques C et C' des deux caméras. Soit M_1 le point de l'espace se projetant en m_1 et m'_1, et M_2 le point de l'espace se projetant en m_2 et m'_2. Les points m_1 et m_2 sont situés dans le même demi-espace. Les droites Cm_1 et Cm_2 coupent le plan Π en C. Les points M_1 et M_2 appartiennent à ces droites. Or nous savons qu'il doit y avoir sur ces droites l'ordre Cm_1M_1 et Cm_2M_2 ou l'ordre m_1CM_1 et m_2CM_2, puisque les caméras ne peuvent regarder qu'un demi-espace au plus. Dans les deux cas, nous en déduisons que M_1 et M_2 appartiennent à un même demi-espace $\mathcal{E}_c$. Le même argument montre alors que M_1 et M_2 se projettent dans le même demi-plan dans la seconde image. La figure 2.5 illustre cette propriété.

Nous pouvons déduire de la remarque précédente le fait suivant, illustré par la figure 2.6: si dans la première image nous observons une configuration de trois demi-droites épipolaires l_1, l_2, l_3 ayant pour origine l'épipole e et passant par les points m_1, m_2, et m_3, respectivement, telles que:

$$0 \leq \langle l_1, l_2 \rangle \leq \langle l_1, l_3 \rangle \leq \pi \tag{2.20}$$

et si l'_1, l'_2, l'_3 sont les demi-droites épipolaires ayant pour origine l'épipole e' et passant par les points correspondants m'_1, m'_2, m'_3 dans la seconde image, nous devons observer une configuration de demi-droites épipolaires telles que:

$$0 \leq \langle l'_1, l'_2 \rangle \leq \langle l'_1, l'_3 \rangle \leq \pi \tag{2.21}$$

les orientations des plans $\mathcal{R}$ et $\mathcal{R}'$ étant fixées de manière arbitraire, nous considérons des angles non orientés. Étant donné les trois demi-droites épipolaires l_1, l_2, l_3 vérifiant la contrainte (2.20), les deux types de configurations de demi-droites épipolaires correspondantes l'_1, l'_2, l'_3 pouvant violer la contrainte (2.21) sont celles telles que:

$$0 \leq \langle l'_1, l'_3 \rangle \leq \langle l'_1, l'_2 \rangle \leq \pi$$

ou

$$0 \leq \langle l'_1, l'_2 \rangle \leq \pi \leq \langle l'_1, l'_3 \rangle$$

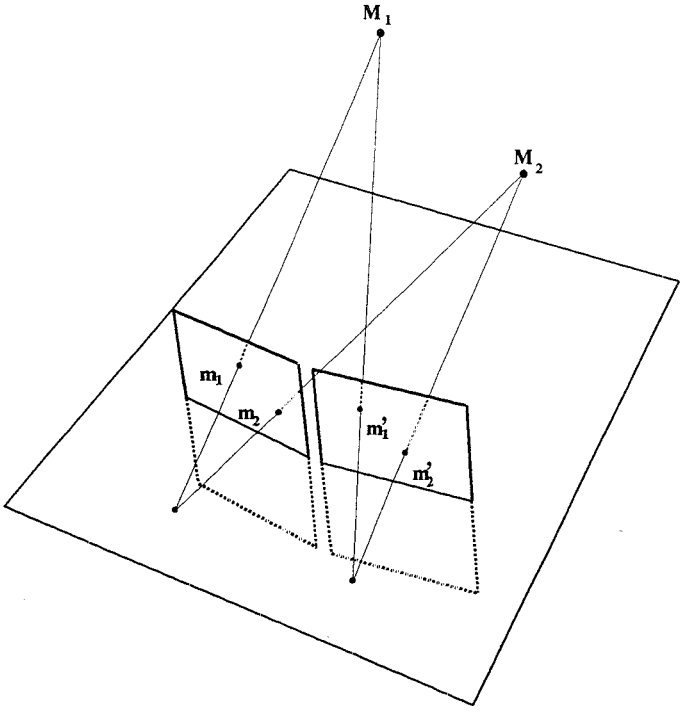


Fig. 2.5: Chaque paire de droites épipolaires partitionne les plans rétiniens en deux demi-plans qui se correspondent

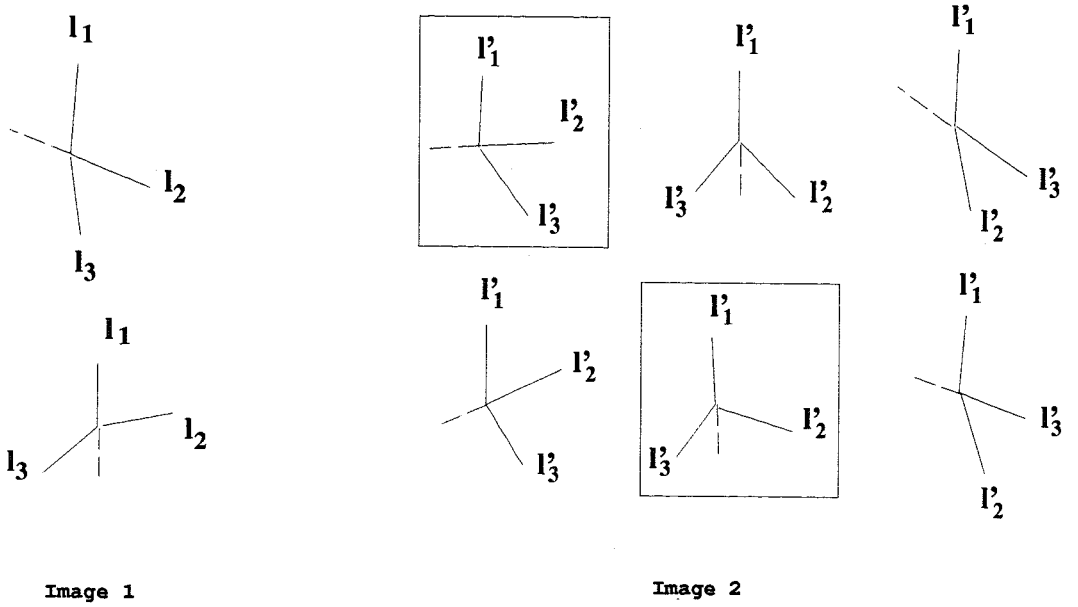


Fig. 2.6: Configurations d'épipolaires permises et configurations interdites

Dans la première de ces configurations, les points de la demi-droite l'_1 et ceux de la demi-droite l'_2 se trouvent dans des demi-plans séparés par la droite l'_3. Dans la seconde de ces configurations, les points de la demi-droite l'_2 et ceux de la demi-droite l'_3 se trouvent dans des demi-plans séparés par la droite l'_1. Or la droite l_2 sépare les points de la demi-droite l_1 et ceux de la demi-droite l_3, d'après ce que nous venions de voir, la demi-droite l'_1 et la demi-droite l'_3 sont séparées par la droite l'_2, donc les demi-droites l'_1 et l'_2 (resp. l'_3 et l'_2) sont dans un même demi-plan par rapport à la droite l'_3 (resp. l'_1). Aussi, les deux configurations précédemment mentionnées ne peuvent se produire, ce qui démontre que nous avons bien (2.21). Des considérations similaires permettent aussi de vérifier que si, avec les mêmes notations que précédemment,

$$0 \leq \langle l_1, l_2 \rangle \leq \pi \leq \langle l_1, l_3 \rangle$$

alors nous devons aussi avoir:

$$0 \leq \langle l'_1, l'_2 \rangle \leq \pi \leq \langle l'_1, l'_3 \rangle$$

Donnons une interprétation géométrique simple de ces contraintes: comme le montre la figure 2.7, trois correspondances (m_1, m'_1), (m_2, m'_2), et (m_3, m'_3) limitent la distribution des paires d'épipoles. Les deux épipoles doivent de situer dans une zone du plan délimitée par les mêmes droites. Il y a quatre zones qui constituent une partition des plans. Une approche mettant en jeu des techniques de géométrie algorithmique pourrait peut-être, en utilisant des combinaisons de triangles, donner une localisation des paires d'épipoles possibles, cependant la combinatoire semble très élevée, ce qui nous conduit à examiner une autre interprétation de la contrainte d'ordre épipolaire.

Une autre manière de voir les choses est de considérer le faisceau de plans épipolaires passant par la droite $\mathbf{CC}'$, et de définir une orientation de cet axe, ce qui nous permet de parler d'angles orientés de plans épipolaires. L'utilisation d'angles orientés permet de donner une formalisation plus précise de la contrainte d'ordre que les notions purement projectives. En effet, si nous considérons un plan P (resp. une droite d) ne contenant pas $\mathbf{CC}'$ (resp. $\mathbf{e}$), nous pouvons définir un ordre sur les paramètres projectifs obtenus par intersection des plans épipolaires (resp. droites épipolaires) avec P (resp. d). Cependant cet ordre ne fait intervenir que les angles non orientés. Nous pouvons également définir une relation $\succ_E$ sur les plans épipolaires par:

$$\Pi_1 \succ_E \Pi_2 \quad \Longleftrightarrow \quad \langle \Pi_1, \Pi_2 \rangle \geq 0$$

Le fait que cette relation soit effectivement un ordre provient du fait que nous utilisons une caméra physique pour laquelle seuls les points situés devant (c'est-à-dire de troisième coordonnée positive dans le repère normalisé) peuvent donner une image. Aussi la partie de l'espace observée est incluse dans un demi espace, ce qui assure l'antisymétrie de la relation $\succ_E$. L'ordre $\succ_E$ de l'espace induit deux ordres $\succ$ et $\succ'$ sur les droites épipolaires des plans $\mathcal{R}$ et $\mathcal{R}'$, respectivement par:

$$\begin{aligned} l_1 \succ l_2 \quad &\Longleftrightarrow \quad \langle l_1, l_2 \rangle \geq 0 \\ l'_1 \succ' l'_2 \quad &\Longleftrightarrow \quad \langle l'_1, l'_2 \rangle \geq 0 \end{aligned} \tag{2.22}$$

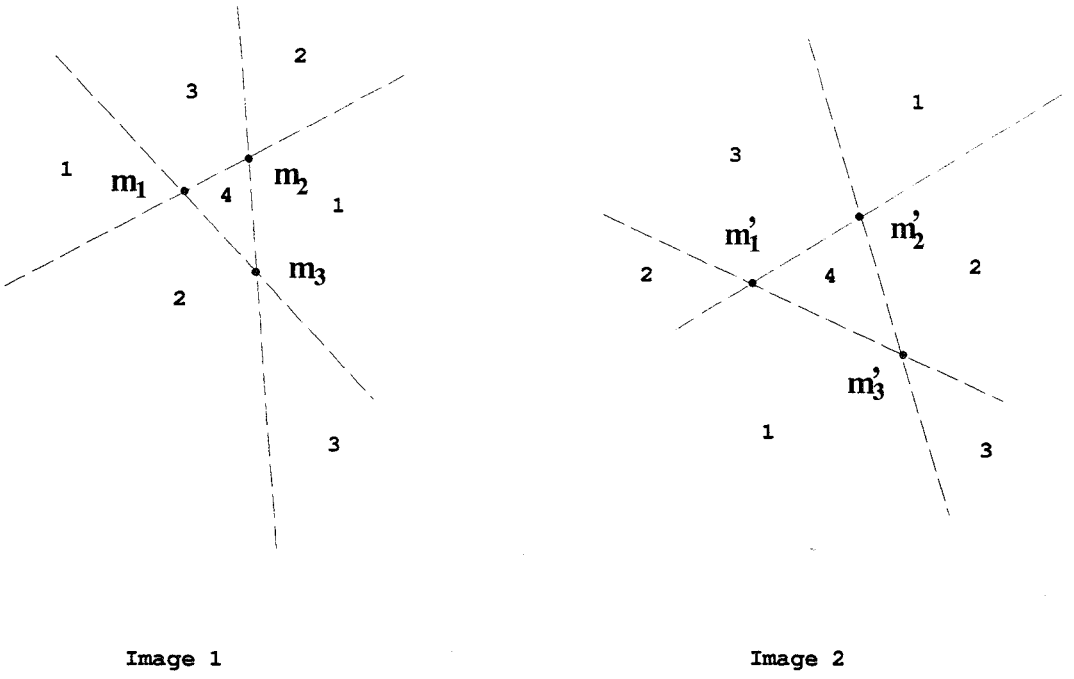


Fig. 2.7: Zones permises pour les paires d'épipoles

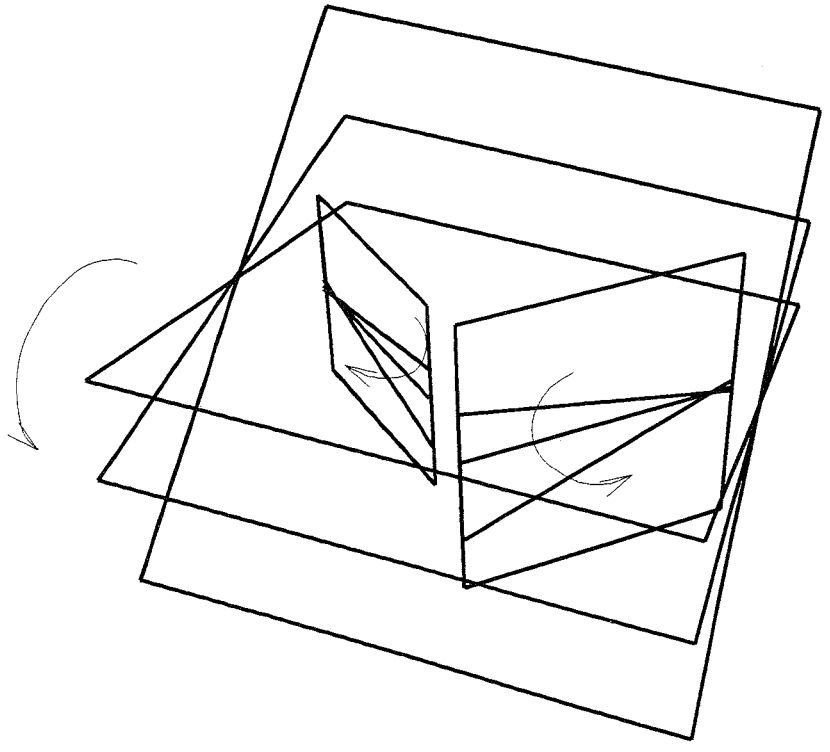


Fig. 2.8: Relations d'ordre sur les rétines induites par l'ordre sur les plans épipolaires

La propriété importante, illustrée par la figure 2.8 est que pour une certaine orientation des deux plans $\mathcal{R}$ et $\mathcal{R}'$, les deux ordres $\succ$ et $\succ'$ sont compatibles avec la relation de correspondance entre les deux images, c'est-à-dire que si l_1, l_2, l_3 sont des demi-droites épipolaires de la première image et si l'_1, l'_2, l'_3 sont les demi-droites épipolaires correspondantes dans la seconde image alors:

$$l_1 \succ l_2 \succ l_3 \Longrightarrow l'_1 \succ' l'_2 \succ' l'_3 \tag{2.23}$$

Il n'est pas possible de déterminer a priori (c'est-à-dire sans connaissance des épipoles) l'orientation adéquate des plans $\mathcal{R}$ et $\mathcal{R}'$ pour que nous ayons (2.23) plutôt que:

$$l_1 \succ l_2 \succ l_3 \Longrightarrow l'_3 \succ' l'_2 \succ' l'_1 \tag{2.24}$$

En effet celle-ci dépend de la position respective des centres optiques $\mathbf{C}$ et $\mathbf{C}'$ et des épipoles $\mathbf{e}$ et $\mathbf{e}'$ le long de l'axe $\mathbf{CC}'$. L'ordre épipolaire constitue néanmoins une contrainte forte permettant de déterminer si une paire de points de $\mathcal{R} \times \mathcal{R}'$ constitue une paire d'épipoles compatibles avec un ensemble de correspondances, au moyen d'une vérification très simple présentée ci-dessous.

vérification de l'ordre épipolaire

- Choisir une paire de points correspondants (m_0, m'_0)
- Pour chaque correspondance, (m_i, m'_i), calculer les angles $\alpha_i = \langle \mathbf{em}_0, \mathbf{em}_i \rangle$ et $\alpha'_i = \langle \mathbf{e}'\mathbf{m}'_0, \mathbf{e}'\mathbf{m}'_i \rangle$
- Effectuer un tri des α_i par ordre croissant, et des α'_i, par ordre croissant et décroissant.
- Vérifier que l'ordre des α_i est identique à celui d'une des deux listes obtenues pour les α'_i.

Chapitre 3

Calcul des épipoles à partir de la transformation épipolaire

Nous examinons une approche du problème de la détermination de la géométrie épipolaire en tant que correspondance homographique entre faisceaux de droites épipolaires. Cette approche semble *a priori* très attrayante car elle nous permet de calculer explicitement les épipoles, qui sont les quantités qui nous intéressent le plus. Le calcul est très découplé, puisque nous ne déterminons que les coordonnées de chacun des deux épipoles à la fois.

Nous présentons d'abord la théorie de la méthode de Sturm, puis nous expliquons les approches successives que nous avons dérivées pour calculer les épipoles. La première consiste à effectuer des statistiques sur les résultats donnés par cette méthode. Nous avons en particulier utilisé une variante de l'algorithme RANSAC. La seconde utilise directement les cubiques planes qui interviennent dans la méthode de Sturm. Leur intersection approchée peut être obtenue avec une technique algébrique, ou avec une technique graphique. La troisième, quant à elle, ne fait intervenir que les birapports qui ont permis de construire ces cubiques. Dans ce dernier cas, il est important de prendre en compte l'incertitude et l'ordre des points.

Comme on pourra le constater, les premières familles de méthodes font intervenir des solutions explicites, la seconde famille nécessite des calculs itératifs à complexité fixe, alors que la dernière famille doit faire appel à des techniques de minimisation non-linéaire. Ceci est la conséquence de l'abandon progressif d'opérations algébriques qui sont certes très élégantes, mais surtout très sensibles au bruit, ce qu'illustrent les résultats expérimentaux présentés tout au long de ce chapitre.

3.1 La méthode de Sturm

La méthode de Sturm est fondée sur la transformation épipolaire et permet de déterminer les épipoles à partir de sept correspondances de points entre deux images. Le nombre sept est naturel, puisqu'il s'agit du nombre de paramètres de la transformation épipolaire. Il s'agit là d'un vieux problème posé par Chasles [17][1]. en ces termes:

> *On se donne dans le même plan deux systèmes de sept points chacun et qui se correspondent. Faire passer par chacun de ce systèmes un faisceau de sept rayons, de telle sorte que les faisceaux soient homographiques .*

Le problème a été résolu par Hesse [62][1] puis par Sturm [147][1], de manière purement géométrique.

3.1.1 Théorie

Nous exposons la formulation algébrique de la méthode de Sturm due à Maybank et Faugeras. Soit $\mathbf{q}_i \leftrightarrow \mathbf{q}_i'$, $\mathbf{q}_i$, un ensemble de correspondances. Nous avons vu qu'alors $\mathbf{x}$ et $\mathbf{x}'$ sont des épipoles compatibles avec ces correspondances si, et seulement si, il existe une homographie $\overline{\wedge}$ telle que:

$$\mathbf{x}\mathbf{q}_i \,\overline{\wedge}\, \mathbf{x}'\mathbf{q}_i' \qquad\qquad 1 \leq i \leq n \tag{3.1}$$

L'existence d'une telle homographie est équivalente à l'égalité des birapports $\{\mathbf{x}\mathbf{q}_i, \mathbf{x}\mathbf{q}_j, \mathbf{x}\mathbf{q}_k, \mathbf{x}\mathbf{q}_l\}$ et $\{\mathbf{x}'\mathbf{q}_i', \mathbf{x}'\mathbf{q}_j', \mathbf{x}'\mathbf{q}_k', \mathbf{x}'\mathbf{q}_l'\}$, quels que soient les quatre indices différents i, j, k, l.

Nous commençons dans un premier temps par montrer que si $n = 6$, alors $\mathbf{x}$ ne peut se situer que sur une cubique plane. Paramétrons le faisceau de droites issu de $\mathbf{x}$ en considérant les points obtenus par l'intersection de ces droites avec la droite $\mathbf{q}_1\mathbf{q}_2$. Soit $\mathbf{x}\mathbf{q}$ une droite passant par $\mathbf{x}$. Alors $\mathbf{x}\mathbf{q}$ et $\mathbf{q}_1\mathbf{q}_2$ s'intersectent au point $\mathbf{x}$ donné par:

$$\begin{aligned} \mathbf{x} &= (\mathbf{x} \wedge \mathbf{q}) \wedge (\mathbf{q}_1 \wedge \mathbf{q}_2) \\ &= [(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_2]\mathbf{q}_1 - [(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_1]\mathbf{q}_2 \end{aligned}$$

Nous pouvons paramétrer les points de $\mathbf{x}\mathbf{q}$ par les coordonnées inhomogènes θ_q, définies par:

$$\theta_q = \frac{(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_2}{(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_1} = \frac{(\mathbf{q} \wedge \mathbf{q}_2).\mathbf{x}}{(\mathbf{q} \wedge \mathbf{q}_1).\mathbf{x}}$$

Remarquons que $\theta_{q_1} = \infty$, $\theta_{q_2} = 0$. Soit θ_1, θ_2, θ_3, θ_4 les coordonnées inhomogènes des quatre droites $\mathbf{x}\mathbf{q}_i$, $1 \leq i \leq 4$. Le birapport τ de ces quatre droites est défini par:

$$\tau = \left(\frac{\theta_1 - \theta_3}{\theta_2 - \theta_3}\right) / \left(\frac{\theta_1 - \theta_4}{\theta_2 - \theta_4}\right)$$

[1] Cité dans [42]

D'où:

$$\tau = \frac{\theta_q}{\theta_{q_3}} = \frac{(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_2}{(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_1} \frac{(\mathbf{x} \wedge \mathbf{q}_3).\mathbf{q}_1}{(\mathbf{x} \wedge \mathbf{q}_3).\mathbf{q}_2}$$

En écrivant l'égalité des birapports des droites $\mathbf{x}\mathbf{q}_i$, $1 \leq i \leq 4$ et des droites $\mathbf{x}'\mathbf{q}'_i$, $1 \leq i \leq 4$, il vient:

$$\frac{(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_2}{(\mathbf{x} \wedge \mathbf{q}).\mathbf{q}_1} \frac{(\mathbf{x} \wedge \mathbf{q}_3).\mathbf{q}_1}{(\mathbf{x} \wedge \mathbf{q}_3).\mathbf{q}_2} = \frac{(\mathbf{x}' \wedge \mathbf{q}').\mathbf{q}'_2}{(\mathbf{x}' \wedge \mathbf{q}').\mathbf{q}'_1} \frac{(\mathbf{x}' \wedge \mathbf{q}'_3).\mathbf{q}'_1}{(\mathbf{x}' \wedge \mathbf{q}'_3).\mathbf{q}'_2} \tag{3.2}$$

En utilisant six correspondances et en posant successivement $\mathbf{q} = \mathbf{q}_4, \mathbf{q}_5, \mathbf{q}_6$, trois équations de la forme (3.2) sont obtenues . Pour simplifier les calculs ultérieurs, un changement de coordonnées projectives est effectué dans chacune des images en prenant pour base canonique dans la première image les points $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4$, et dans la seconde image leurs correspondants $\mathbf{q}'_1, \mathbf{q}'_2, \mathbf{q}'_3, \mathbf{q}'_4$:

$$\begin{array}{llll} \mathbf{q}_1 & = & (1,0,0)^\top & \mathbf{q}'_1 = (1,0,0)^\top \\ \mathbf{q}_2 & = & (0,1,0)^\top & \mathbf{q}'_2 = (0,1,0)^\top \\ \mathbf{q}_3 & = & (0,0,1)^\top & \mathbf{q}'_3 = (0,0,1)^\top \\ \mathbf{q}_4 & = & (1,1,1)^\top & \mathbf{q}'_4 = (1,1,1)^\top \end{array}$$

Avec ce choix de coordonnées, les équations (3.2) se réduisent à:

$$\frac{(x_3q_1 - x_1q_3)x_2}{(x_3q_2 - x_2q_3)x_1} = \frac{(x'_3q'_1 - x'_1q'_3)x'_2}{(x'_3q'_2 - x'_2q'_3)x'_1} \tag{3.3}$$

Définissons la transformation quadratique fondamentale Φ_0, ainsi que $\mathbf{u}$ et $\mathbf{u}'$, par:

$$\begin{array}{l} \Phi_0(\mathbf{x}) = (x_2x_3, x_3x_1, x_1x_2)^\top \\ \mathbf{u} = \Phi_0(\mathbf{x}) \quad \mathbf{u}' = \Phi'_0(\mathbf{x}') \end{array} \tag{3.4}$$

En appliquant Φ_0 à (3.3), il vient:

$$\frac{u_1q_1 - u_3q_3}{u_2q_2 - u_3q_3} = \frac{u'_1q'_1 - u'_3q'_3}{u'_2q'_2 - u'_3q'_3} \tag{3.5}$$

En définissant le vecteur $\mathbf{a}_q(\mathbf{u})$ par:

$$\mathbf{a}_q(\mathbf{u}) = ((u_3q_3 - u_2q_2)q'_1, (u_1q_1 - u_3q_3)q'_2, (u_2q_2 - u_1q_1)q'_3)^\top \tag{3.6}$$

il vient, à partir de (3.5):

$$\mathbf{a}_q(\mathbf{u}).\mathbf{u}' = 0 \tag{3.7}$$

Nous obtenons donc trois équations de la forme (3.7), à partir des correspondances $\mathbf{q}_4 \leftrightarrow \mathbf{q}'_4$, $\mathbf{q}_5 \leftrightarrow \mathbf{q}'_5$, $\mathbf{q}_6 \leftrightarrow \mathbf{q}'_6$. Notons les trois vecteurs $\mathbf{a}_q(\mathbf{u})$ ainsi obtenus $\mathbf{a}_4(\mathbf{u})$, $\mathbf{a}_5(\mathbf{u})$, $\mathbf{a}_6(\mathbf{u})$. Pour que les trois équations (3.7) donnent une solution pour $\mathbf{u}'$, la condition de coplanarité doit être vérifiée:

$$f_u(\mathbf{u}) = (\mathbf{a}_4(\mathbf{u}) \wedge \mathbf{a}_5(\mathbf{u})).\mathbf{a}_6(\mathbf{u}) = 0 \tag{3.8}$$

L'équation (3.8) définit une contrainte polynômiale du troisième degré. En remplaçant $\mathbf{u}$ par sa valeur en fonction de $\mathbf{x}$, nous trouvons:

$$f_u(\mathbf{u}) = x_1 x_2 x_3 f(\mathbf{x}) \tag{3.9}$$

où $f(\mathbf{x})$, définie par l'équation précédente, est une contrainte polynômiale du troisième degré sur $\mathbf{x}$,

Montrons que l'on peut trouver six points ne dépendant pas de $\mathbf{q}_6$ ni de $\mathbf{q}_6'$, par lesquels $f(\mathbf{x})$ passe. Une simple inspection des formules permet de vérifier que $f(\mathbf{x})$ passe par les trois points $\mathbf{q}_i$, $1 \leq i \leq 3$. Il est aussi aisé de voir que $f(\mathbf{x})$ passe par les points $\mathbf{q}_i$, $4 \leq i \leq 6$, puisque l'on a $\mathbf{a}_i(\Phi_0(\mathbf{q}_i)) = 0$ pour $4 \leq i \leq 6$. Les cinq points $\mathbf{q}_i$, $1 \leq i \leq 5$ conviennent donc. Il est également simple de voir que $f(\mathbf{x})$ passe par les points $\mathbf{b}_i$ $1 \leq i \leq 3$ satisfaisant:

$$\begin{aligned} \mathbf{a}_4(\Phi_0(\mathbf{b}_1)) &= \lambda_1 \mathbf{a}_5(\Phi_0(\mathbf{b}_1)) \quad \lambda_1 \neq 0 \\ \mathbf{a}_5(\Phi_0(\mathbf{b}_2)) &= \lambda_2 \mathbf{a}_6(\Phi_0(\mathbf{b}_2)) \quad \lambda_2 \neq 0 \\ \mathbf{a}_6(\Phi_0(\mathbf{b}_3)) &= \lambda_3 \mathbf{a}_4(\Phi_0(\mathbf{b}_3)) \quad \lambda_3 \neq 0 \end{aligned} \tag{3.10}$$

(3.11)

Le point $\mathbf{b} = \mathbf{b}_1$ ne dépend pas de $\mathbf{q}_6$ ni de $\mathbf{q}_6'$.

Soit $\mathbf{q}_7 \leftrightarrow \mathbf{q}_7'$ une septième correspondance. En remplaçant $\mathbf{q}_6$ et $\mathbf{q}_6'$ par $\mathbf{q}_7$ et $\mathbf{q}_7'$, respectivement, nous avons une seconde contrainte polynômiale du troisième degré sur $\mathbf{x}$, $g(\mathbf{x}) = 0$. En général, les courbes f, g ont neuf points d'intersection. Les épipoles $\mathbf{x}$ compatibles avec les sept correspondances sont parmi ces neuf intersections. Six de ces intersections sont les points fixes mentionnés précédemment, qui en général, ne sont pas des épipoles. On peut en conclure qu'il y a au plus $3 = 9 - 6$ épipoles compatibles avec sept correspondances.

Voyons à présent pourquoi il y a en général exactement trois épipoles compatibles avec sept correspondances. La transformation épipolaire peut être décrite par sept paramètres, deux pour chaque épipole et trois pour l'homographie. Comme chaque correspondance entraîne une contrainte unique, il y a au moins une transformation épipolaire compatible avec les correspondances $\mathbf{q}_i \leftrightarrow \mathbf{q}_i'$, $1 \leq i \leq 7$. En utilisant cette transformation, des points de l'espace $\mathbf{C}$, $\mathbf{C}'$ et $\mathbf{M}_i$ peuvent être construits pour que $\mathbf{M}_i$ se projette sur $\mathbf{q}_i$ dans une image obtenue par une camera dont le centre optique est $\mathbf{C}$, et que $\mathbf{M}_i$ se projette en $\mathbf{q}_i'$ dans une image obtenue par une camera dont le centre optique est $\mathbf{C}'$. Soit ψ la quadrique passant par les neuf points $\mathbf{C}$, $\mathbf{C}'$, $\mathbf{M}_i$, $1 \leq i \leq 7$. Elle définit une transformation quadratique Ψ de la première vers la seconde image, faisant correspondre à chaque point $\mathbf{q}_i$ de la première image le point $\mathbf{q}_i'$ qui est l'image par la seconde camera de l'intersection de la droite $\mathbf{Cq}$ avec la quadrique ψ. Cette transformation vérifie: $\Psi(\mathbf{q}_i) = \mathbf{q}_i'$, $1 \leq i \leq 7$. Les trois points fondamentaux de Ψ sont des épipoles possibles, comme nous le verrons plus précisément en 6.2.

L'analyse du problème de la détermination des épipoles par la méthode de Sturm nous permet donc de voir qu'avec sept correspondances, nous n'avons pas, en général, d'unicité de la solution. Nous en concluons: *un minimum de huit correspondances est nécessaire pour déterminer la transformation épipolaire de manière unique.*

3.1.2 Un exemple

Nous illustrons la méthode de Sturm par un exemple. Le mouvement dans le repère de la caméra est une translation de vecteur $(-7, 1, 7)^T$. La figure 3.1 montre la partie du plan rétinien de la première caméra contenant l'image 512×512, ainsi qu'un agrandissement à l'intérieur de l'image, qui permet de discerner les points particuliers. Sur ces figures, sont représentés en rouge les 7 points mis en correspondance, dont on voit bien que les 5 premiers sont communs aux deux cubiques. Le point **b**, marqué en violet, a ici pour coordonnées $(177.07, 238.16)^T$, il est également commun aux deux courbes. Les trois autres points d'intersection sont marqués en vert, ce sont les épipoles possibles, en $\mathbf{x}_1 = (-393.98, 390.46)^T$, $\mathbf{x}_2 = (313.46, 187.95)^T$ et $\mathbf{x}_3 = (170.78, 229.78)^T$. En effectuant des calculs similaires dans la seconde image, on trouve également trois épipoles possibles, en $\mathbf{x}'_1 = (-393.98, 390.46)^T$, $\mathbf{x}'_2 = (311.06, 189.86)^T$, et $\mathbf{x}'_3 = (169.04, 229.29)^T$. L'utilisation de la contrainte d'ordre épipolaire permet de montrer que les paires d'épipoles correspondantes sont $(\mathbf{x}_1, \mathbf{x}'_1)$, $(\mathbf{x}_2, \mathbf{x}'_2)$ et $(\mathbf{x}_3, \mathbf{x}'_3)$, donnant, après calcul de l'homographie par moindres carrés et utilisation des relations (2.13) du chapitre précédent, les matrices fondamentales: $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{F}_3$. La matrice des paramètres intrinsèques est connue dans cet exemple et permet d'obtenir les matrices essentielles par la relation (2.8):

$$\mathbf{E}_1 = \begin{bmatrix} 0 & -604082.7619 & 86296.5315 \\ 604082.7619 & 0 & 604036.5208 \\ -86296.53150 & -604036.5208 & 0 \end{bmatrix}$$

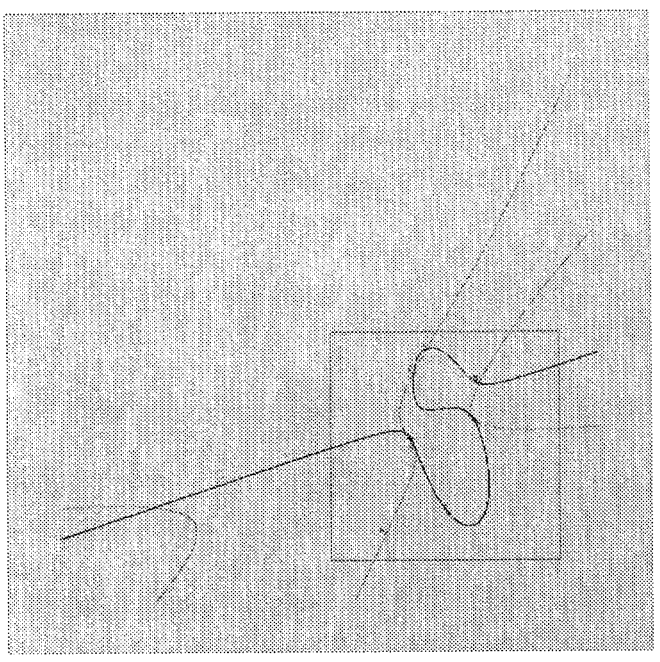
$$\mathbf{E}_2 = \begin{bmatrix} -20.09918795 & -3309.356723 & -235.2887031 \\ 3306.849964 & 6.953227105 & -347.499852 \\ 232.5708935 & 336.3395256 & -.3467171 \end{bmatrix}$$

$$\mathbf{E}_3 = \begin{bmatrix} -22.51376542 & -4369.697851 & -122.428461 \\ 4367.306838 & 7.679878501 & 514.060282 \\ 119.2715224 & -525.7943815 & -.37947841 \end{bmatrix}$$

Nous voyons que $\mathbf{E}_1$ correspond effectivement au mouvement effectué, à un facteur d'échelle près. En revanche, $\mathbf{E}_2$ et $\mathbf{E}_3$ ne sont pas des matrices essentielles, ce que nous constatons en vérifiant la contrainte additionnelle (2.6), et induisent une transformation linéaire projective qui n'est pas un déplacement. La bonne paire d'épipoles est donc ici $(\mathbf{x}_1, \mathbf{x}'_1)$. Notons que dans le cadre des méthodes que nous développons, il n'est pas possible d'effectuer une discrimination en procédant ainsi, puisque les paramètres intrinsèques ne sont pas supposés connus.

3.1.3 Deux mouvements typiques

L'étude expérimentale de toute cette partie sera illustrée par deux mouvements particuliers qui sont typiques, dont les caractéristiques sont présentées table 3.1, et le champ de déplacement illustré par la figure 3.1.3. Dans le cas du premier mouvement, les épipoles se trouvent dans l'image, dans celui du second, ils en sont éloignés. Nous verrons que ces deux



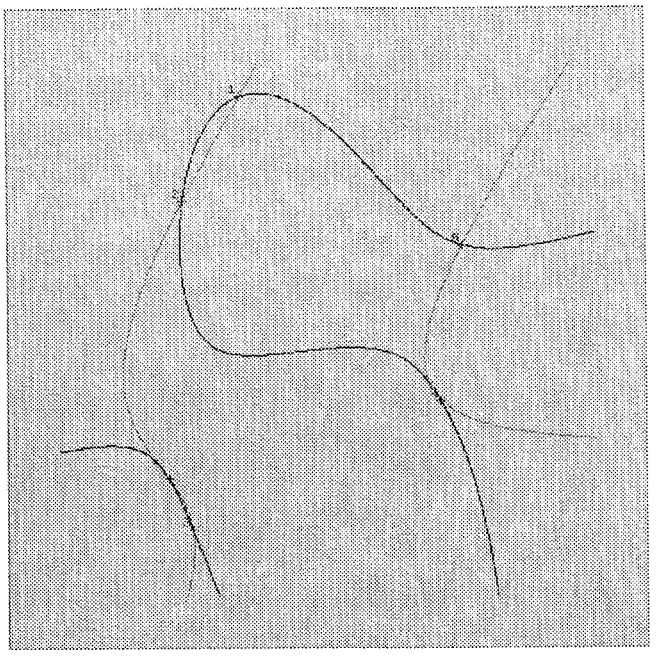


Fig. 3.1: La méthode de Sturm: un exemple montrant les cubiques et points remarquables

configurations conduisent à des comportements très différents pour les algorithmes. L'image que nous utiliserons pour illustrer nos propos sera toujours celle prise par la caméra avant déplacement. Du point de vue des problèmes de stabilité qui nous préoccupent, la seconde image, c'est-à-dire celle prise par la caméra après déplacement, se révèle équivalente. Les

Tableau 3.1: Les deux mouvements utilisés pour illustrer la méthode de Sturm

mouvement 1					
rotation			translation		
0	.05	0	0	0	400
mouvement 2					
rotation			translation		
0.1	0	0	50	20	20

appariements sont obtenus à partir d'un nuage de points de l'espace dont les coordonnées sont tirées aléatoirement. Nous calculons la projection de chacun de ces points sur le plan rétinien et ne gardons que ceux qui tombent dans le champ de vision de la caméra. Puis nous bruitons les points du plan ainsi obtenu par une distribution gaussienne d'écart-type σ. Par la nature des opérations impliquées (éliminations algébriques), la méthode de Sturm se

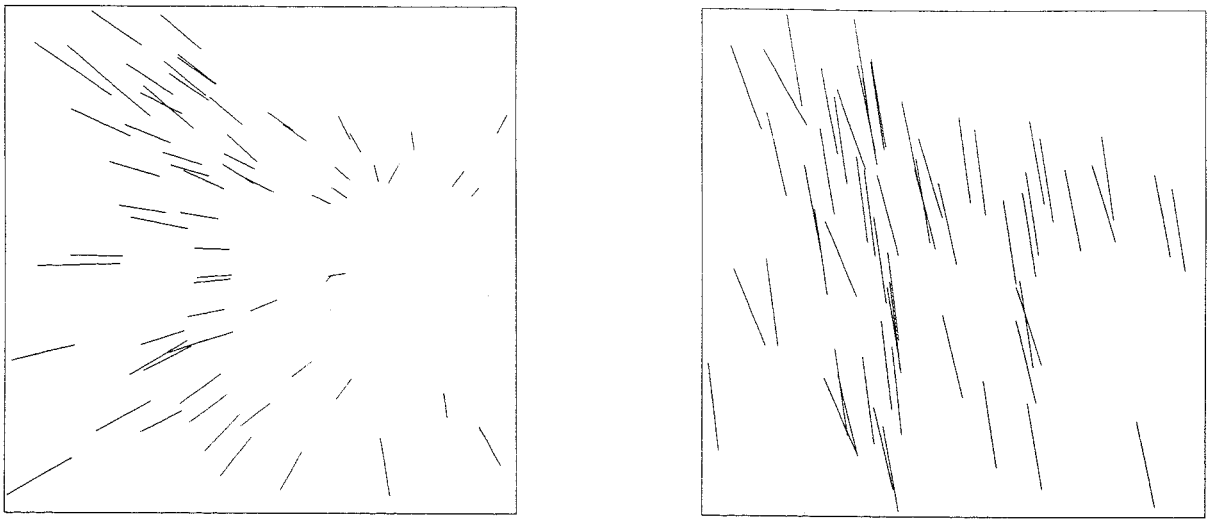


Fig. 3.2: Champ des déplacements avec les mouvements 1 (gauche) et 2 (droite)

prête peu au calcul d'erreur, cependant on peut avoir une idée de sa stabilité en considérant les modifications des cubiques en présence de bruit, sur des exemples. La figure 3.1.3 montre la superposition de cubiques obtenues pour le premier mouvement à partir des même sept points, mais à différents niveaux de bruit (exprimés en pixels): $\sigma = 0$, $\sigma = 0.01$, $\sigma = 0.1$, $\sigma = 0.5$, $\sigma = 1$, $\sigma = 2$. Le carré représente la rétine. La figure 3.1.3 montre les courbes analogues obtenues pour le second mouvement. On peut noter que les courbes obtenues avec des points bruités s'écartent davantage de l'épipole exact (figuré par la croix) dans le cas où celui-ci est éloigné de l'image. Nous pouvons essayer d'expliquer ce phénomène. Les cubiques intervenant dans la méthode de Sturm sont entièrement déterminées par les neuf points suivants: $\mathbf{q}_i$, $1 \leq i \leq 6$, $\mathbf{b}_i$, $1 \leq i \leq 3$. Bruiter les correspondances revient à perturber ces points qui définissent la cubique. On comprend intuitivement que plus on s'éloigne de ces points, plus la perturbation de la cubique est importante. Une autre remarque est le fait que des épipoles très éloignés se situent sur des branches asymptotiques de cubiques, pour lesquelles les directions sont très voisines. L'intersection de deux courbes presque tangentes est très sensible au bruit.

La méthode de Sturm présente l'avantage de permettre une résolution explicite du problème, non seulement de manière exacte, ce qui est l'objet d'une première série d'approches, mais aussi d'une manière approchée, par minimisation aux moindres carrés d'un critère d'erreur, ce qui est l'objet d'une seconde approche. Nous verrons néanmoins que les résultats ne sont pas suffisamment robustes, et que l'utilisation d'une méthode itérative se révèle en fin de compte préférable.

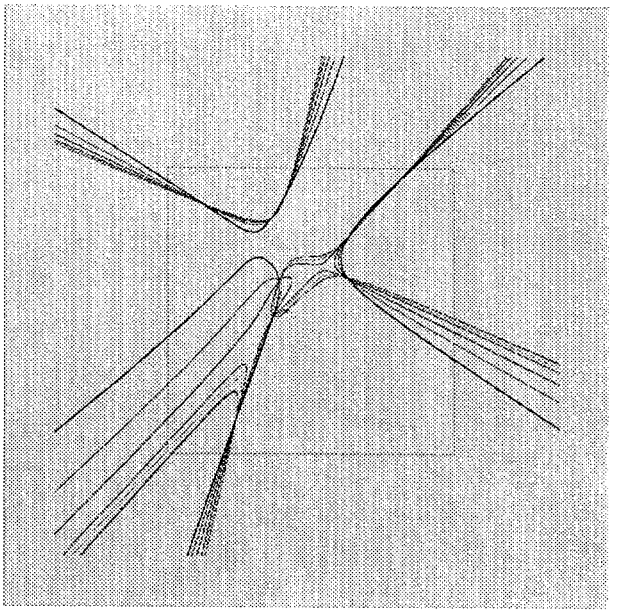


Fig. 3.3: Superposition de cubiques obtenues à différents niveaux de bruit avec le mouvement 1

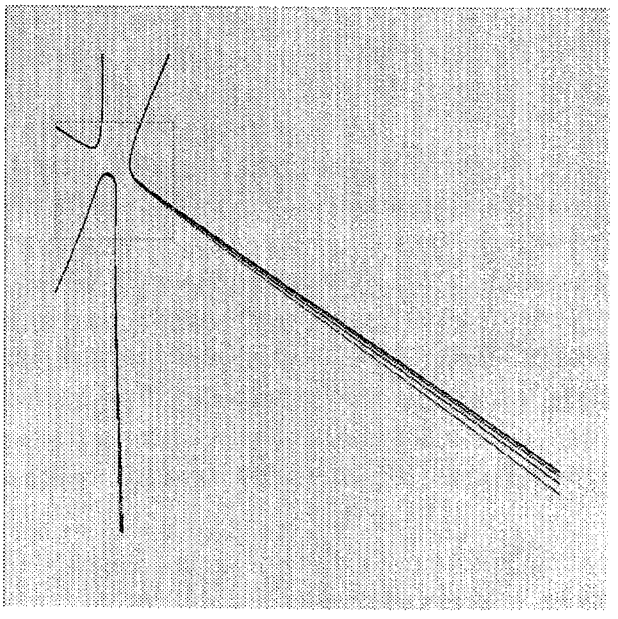


Fig. 3.4: Superposition de cubiques obtenues à différents niveaux de bruit avec le mouvement 2

3.2 Utilisation de la méthode de Sturm

3.2.1 L'algorithme de Sturm: quelques détails

Nous donnons quelques précisions pour montrer comment la méthode décrite précédemment est susceptible d'être implémentée uniquement avec des opérations exactes. Nous avons d'abord effectué cette implémentation au moyen du système de calcul formel MAPLE, puis en C.

Nous débutons par le calcul de la matrice de changement de repère:

$$\mathbf{Q} = [\rho_1 \mathbf{p}_1 \ \rho_2 \mathbf{p}_2 \ \rho_3 \mathbf{p}_3] \qquad \text{avec} \quad \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3]^{-1} \mathbf{p}_4 \tag{3.12}$$

Nous obtenons les nouvelles coordonnées $\mathbf{q}_i$ en appliquant la matrice $\mathbf{Q}^{-1}$ aux points initiaux $\mathbf{p}_i$. Nous procédons de la même manière également dans la seconde image pour obtenir les points $\mathbf{q}'_i$.

Les cubiques f et g sont calculées avec les formules (3.4) (3.6) (3.8) (3.9), présentées précédemment. Il reste à obtenir leur intersection. Soit $h(x_1, x_2)$ le résultant de f et g par rapport à x_3. Comme $\mathbf{x}$ est un point projectif, nous pouvons réduire le nombre d'inconnues en examinant les cas $x_2 = 0$ et $x_2 \neq 0$. Dans le premier cas, les solutions obtenues sont les points $(1,0,0)^T$ et $(0,0,1)^T$, qui correspondent à $\mathbf{q}_1$ et $\mathbf{q}_3$ dont nous savons qu'ils sont des points fixes de f et g. Dans le second cas, nous pouvons poser $x_2 = 1$, et considérer le résultant $h_1(x_1) = h(x_1, 1)$. Celui-ci conduit d'une part à la solution $(0,1,0)^T$ obtenue pour $x_1 = 0$, d'autre part à six valeurs de x_1 non nulles. Or nous savons que parmi ces six valeurs se trouvent 1, $(\mathbf{q}_5)_1/(\mathbf{q}_5)_2$, et b_1/b_2 puisque $\mathbf{q}_4$, $\mathbf{q}_5$, et $\mathbf{b}$ (que nous calculons en résolvant l'équation (3.10)) sont connus comme points fixes de f et g. Nous pouvons donc réduire le degré de h_1 par division par les trois monômes obtenus à partir de ces trois valeurs. Nous obtenons ainsi une équation du troisième degré que nous résolvons par les formules de Cardan. Dans le cas général, il en résulte soit une, soit trois solutions réelles. Pour chacune de ces valeurs de x_1, nous effectuons la substitution dans f et g, et nous obtenons deux équations du second degré en x_3, dont nous prenons la racine commune, déterminant ainsi un ou trois points projectifs du plan, $\mathbf{x}_i$. Nous obtenons alors les épipoles exprimés en coordonnées rétiniennes en appliquant la matrice $\mathbf{Q}$ définie par (3.12) aux points $\mathbf{x}_i$.

3.2.2 Sélection de l'épipole: la méthode exhaustive

Nous obtenons ainsi soit un seul épipole réel, ce qui est le cas le plus favorable, soit trois épipoles réels. Comme nous l'avons vu, ces trois épipoles sont tous compatibles avec les sept correspondances, aussi pour pouvoir sélectionner le bon épipole, nous devons utiliser plus que sept correspondances. Nous pouvons alors appliquer la méthode de Sturm à chaque sous-ensemble de sept correspondances et obtenir trois candidats. L'idée est que les mauvais candidats sont alors compatibles seulement avec le sous-ensemble de correspondances qui a

servi à les déterminer, alors que le bon candidat sera compatible avec tous les sous-ensembles de correspondances. En utilisant au minimum huit points, une procédure de choix exhaustif très simple mais relativement coûteuse, dont la description est donnée ci-dessous, permet donc en principe de sélectionner le bon épipole.

méthode de Sturm avec choix exhaustif

- Appliquer la méthode de Sturm n fois
- Pour chacune des 3^n combinaisons obtenues en prenant un point dans chaque triplet, calculer le diamètre de la boule contenant les n points sélectionnés
- Retenir la configuration correspondant au diamètre minimal
- Effectuer une moyenne sur les n points ainsi obtenus

Nous avons essayé plusieurs opérateurs pour effectuer le calcul de la boule ainsi que pour effectuer la moyenne finale, cependant aucun d'entre eux n'a donné de résultat significativement meilleur.

Le comportement de l'algorithme de Sturm avec procédure de choix exhaustif est illustré par les tables 3.2 et 3.3. Elles montrent les triplets d'épipoles obtenus en utilisant huit sous-ensembles d'appariements obtenus respectivement à partir des deux mouvements présentés plus haut. L'épipole sélectionné à chaque fois par l'algorithme figure dans la première colonne.

Les résultats peuvent être très instables, même pour des bruits très faibles. L'un des problèmes rencontrés est que si des ensembles de correspondances de cardinalité faible sont pris, les mêmes points sont souvent réutilisés et conduisent à des faux candidats qui ont des positions assez proches, ce qui rend le choix du bon épipole difficile. D'un autre côté, il est trop coûteux d'essayer tous les sous-ensembles de sept points. Cette remarque suggère l'utilisation d'un algorithme avec tirage aléatoire.

3.2.3 Statistique robuste: utilisation d'une procédure de consensus aléatoire

Nous avons tenté d'améliorer le comportement de l'algorithme, sans changer la procédure de calcul par la méthode de Sturm. Pour cela, nous avons remplacé la procédure de choix exhaustif par une procédure de consensus sur des tirages aléatoires (RANSAC), introduite dans le domaine du traitement d'images par Fischler et Bolles [45], dont le mode de fonctionnement est adapté à notre méthodologie. En effet, dans notre cas, nous disposons d'un nombre élevé de points, mais n'utilisons que des sous-ensembles de sept points pour déterminer les paramètres que nous recherchons (les coordonnées de l'épipole). La méthode RANSAC diffère des méthodes usuelles en ce qu'au lieu d'utiliser le plus de données possible pour obtenir une solution initiale en calculant les paramètres inconnus par une méthode de moyennage,

bruit	214.63 254.57	295.94 195.33	196.73 267.02
0.01 pixel	214.65 257.63	283.98 349.00	335.38 406.11
	214.19 254.46	212.76 364.87	215.66 219.93
point moyen	214.09 255.73	160.12 434.10	197.74 402.63
214.21 255.52	213.95 255.54	210.49 275.51	87.63 364.35
écart type	214.06 255.20		
0.768	213.77 255.92	34.05 240.87	323.53 237.11
	214.38 255.11	162.60 296.62	271.24 201.11
bruit	217.40 246.72	294.41 192.39	190.14 264.99
0.1 pixel	219.73 274.79	279.38 350.68	329.88 405.74
	212.37 365.30		
point moyen	214.38 256.50	161.00 435.87	198.60 405.10
214.39 272.31	209.33 274.80	88.19 363.39	212.94 254.61
écart type	214.02 251.19		
25.41	211.00 258.42	32.27 246.94	323.46 236.95
	216.89 250.72	164.53 290.62	271.92 200.29
bruit	168.55 251.34	212.95 226.10	297.03 174.50
0.5 pixel	298.45 399.59		
	210.50 367.51		
point moyen	215.97 259.68	165.65 445.99	203.33 418.14
209.83 290.65	204.56 272.20	90.72 358.99	208.91 250.75
écart type	212.90 234.28		
56.06	195.11 272.08	24.92 277.29	323.12 236.14
	172.59 268.51	275.04 197.31	223.67 237.51
bruit	201.00 208.02	145.58 233.57	307.37 152.37
1 pixel	239.37 381.65		
	207.70 370.94		
point moyen	218.74 263.29	173.70 466.48	212.60 441.50
202.57 281.99	199.35 269.86	93.98 353.31	204.71 246.64
écart type	209.71 214.52		
57.00	163.98 298.59	11.28 324.60	322.65 234.95
	180.72 249.05	279.15 194.81	227.63 228.00

Tableau 3.2: Résultats de l'algorithme de Sturm avec choix exhaustif, pour le mouvement 1

bruit	2061.37 1429.59	185.43 89.48	321.25 206.51
0.01 pixel	1723.92 1229.74	276.00 342.41	363.99 399.44
	2021.32 1401.84	219.87 360.90	382.37 449.56
point moyen	2190.57 1485.96	1265.45 -123.90	2445.93 1450.01
2010.77 1395.10	2014.10 1399.35	190.24 282.83	124.82 288.30
écart type	2032.64 1407.84		
83.34	2009.49 1393.98	242.31 153.54	324.30 238.72
	2032.75 1412.50	977.63 1361.55	-61.92 541.83
bruit	2364.24 1646.16	184.80 91.54	321.73 208.99
0.1 pixel	790.92 684.22	277.31 342.92	364.16 403.48
	1917.58 1335.27	219.92 360.81	374.40 444.31
point moyen	1772.02 -396.31	8466.96 4576.25	9682.96 4972.67
1821.20 1085.56	1857.21 1316.17	190.07 282.78	124.92 288.21
écart type	2022.09 1390.05		
603.46	1819.15 1269.88	243.77 154.26	324.50 239.16
	2026.35 1439.05	1039.95 1426.21	-68.44 545.90
bruit	322.29 211.65	184.10 93.68	2879.72 2013.39
0.2 pixel	278.66 343.53	368.18 410.75	536.14 530.28
	219.98 360.69	1817.51 1270.81	366.24 438.93
point moyen	-2991.66 -1004.45	-3569.01 -1554.60	3573.79 -1365.63
26.65 295.05	125.03 288.11	1717.15 1242.44	189.89 282.72
écart type	2010.12 1370.60		
873.44	324.71 239.66	245.36 155.07	1653.81 1162.14
	-75.91 550.58	1118.38 1505.55	2025.66 1474.49

Tableau 3.3: Résultats de l'algorithme de Sturm avec choix exhaustif, pour le mouvement 2

elle utilise des ensembles initiaux de données S, de taille minimale pour permettre le calcul d'un jeu de paramètres M, puis élargit ces ensembles en incorporant des données consistantes (fait intervenir un seuil de tolérance t_1, et un test avec chacun des points restants) avec les paramètres calculés à partir de l'ensemble initial. Les N_1 ensembles initiaux sont obtenus par tirage aléatoire et conduisent à des paramètres acceptés si l'ensemble de consensus S^* a un cardinal supérieur à un certain seuil, t_2. L'important avantage de RANSAC sur les méthodes de moyennage est sa robustesse vis-à-vis des points de données dont l'erreur est grande, provenant par exemple de faux appariements.

Dans le cas de notre problème, une donnée est un sous-ensemble de sept points, et le modèle est constitué par un point 2D et un rayon d'erreur. Il est en effet nécessaire de tenir compte du fait que suivant la position de l'épipole, la déviation des courbes bruitées est plus ou moins importante, ce qu'illustrent tous les graphes que nous présentons. Le nombre minimal de données permettant d'instancier le modèle est donc de deux sous-ensembles de sept points, que nous prenons disjoints pour obtenir une bonne distinction des faux épipoles. L'adaptation de l'algorithme RANSAC à notre problème est présentée ci-dessous, elle diffère en quelques points de l'algorithme original:

- Au lieu d'effectuer un simple tirage aléatoire, nous effectuons une sélection des données servant à instancier le modèle initial. Nous introduisons pour cela une distance d_1 qui permet de tenir compte du fait que les écarts augmentent lorsque l'on s'éloigne de l'image. Dans la version qui a servi a obtenir les résultats présentés plus bas, nous avons pris le maximum des erreurs relatives sur les deux coordonnées. Deux paramètres supplémentaires sont introduits: un seuil de tolérance t_0 pour le choix du modèle initial, un nombre d'essais maximal N_2.

- Au lieu de tester toutes les autres données, nous nous restreignons à une partie de N_3 sous-ensembles obtenus par tirage aléatoire. En effet, nos données, qui sont constituées par les C_n^7 sous-ensembles de sept points sont en nombre trop élevé pour pouvoir être toutes testées, de plus cela favoriserait certains faux épipoles.

Augmenter la valeur des paramètres d'essai N_1, N_2, N_3 augmente le nombre de calculs effectués, mais aussi les chances d'obtenir un résultat à paramètres de seuil fixés. Augmenter la valeur des paramètres de seuil t_0, t_1, t_2 diminue la sélectivité de la procédure: on augmente la possibilité d'obtenir un résultat, mais on diminue sa certitude d'être correct. On peut aussi noter qu'il y a un couplage entre les paramètres N_2 et t_0, N_3 et t_2, et, plus indirectement N_1 et t_1.

méthode de Sturm avec consensus aléatoire

Répéter jusqu'à l'obtention d'un épipole satisfaisant (N_1 essais au maximum):

- *Initialiser le modèle [Point-diamètre].* Répéter jusqu'à obtenir un modèle satisfaisant (N_2 essais au maximum):
  - tirer aléatoirement deux sous-ensembles disjoints de 7 correspondances, S_1 et S_1'.
  - calculer les triplets d'épipoles $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ et $E' = (\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3')$ à partir de S_1 et S_1'.
  - prendre la paire d'épipoles $(\mathbf{e}_i, \mathbf{e}_i') \in E \times E'$ donnant la plus petite distance $d_1 = d_1(\mathbf{e}_i, \mathbf{e}_i')$.
  - si $d_1 > t_0$, prendre $\mathbf{e} = (\mathbf{e}_i + \mathbf{e}_i')/2$, et $r = d_2(\mathbf{e}_i, \mathbf{e})$, sinon recommencer.
- *Chercher un ensemble de consensus pour* $(\mathbf{e}, d)$. Répéter pour obtenir un nombre significatif N_3 de tirages de 7 correspondances:
  - tirer aléatoirement un sous-ensemble de 7 correspondances S'
  - calculer le triplet d'épipoles $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ à partir de S'.
  - si l'un des $\mathbf{e}_i$ est dans la boule de centre $\mathbf{e}$ et de rayon $r \times t_1$, ajouter 1 au nombre de points de consensus N^* et remettre à jour le modèle: $\mathbf{e} = N^*\mathbf{e} + \mathbf{e}_i$, $r = \sup(r, d_2(\mathbf{e}, \mathbf{e}_i))$
- Si $N_3 - N^* > t_2$, donner $\mathbf{e}$ en sortie, sinon recommencer, et sortir finalement en échec avec l'épipole ayant donné le plus grand ensemble de consensus.

Cette approche est satisfaisante dans le cas d'un épipole dans l'image, ou qui en est proche, mais elle reste également très sensible au bruit dès que l'épipole en est très éloigné, comme le montrent des résultats présentés tables 3.4 et 3.5. Les expérimentations ont été effectués avec 58 appariements, les paramètres étant $N_1 = 100$, $N_2 = 2\times$, $N_3 = 2 \times 58$, $t_0 = 0.05$, $t_1 = 1.2$, $t_2 = 58$. Les raisons de la dégradation des résultats pour le mouvement 2 proviennent du fait que lors de la phase d'initialisation du modèle, si la distance des vrais épipoles perturbés devient comparable à la taille de l'image, ce sont deux points de l'image qui seront sélectionnés pour instancier le modèle. Nous avons remarqué qu'en l'absence du critère de sélection par l'erreur relative, ce phénomène se produit à des niveaux de bruit encore plus faibles. En diminuant le paramètre t_0 on peut favoriser un peu les épipoles éloignés, mais on n'améliore pas considérablement les résultats.

Le problème, à la base, vient des résultats instable obtenus par l'utilisation de la méthode de Sturm, qui sert de noyau au programme RANSAC. La méthode de Sturm en effet ne tient pas compte explicitement du bruit, et fait intervenir des opérations algébriques rela-

bruit	épipole trouvé		rayon d'erreur	% consensus	essais
0	214.06	255.64	0.000021	0.676	1
0.01	214.09	255.62	1.55	0.846	1
0.1	214.40	254.72	10.09	0.882	3
0.2	214.45	255.19	7.65	0.544	2
0.5	215.60	252.42	16.12	0.603	43
1	217.54	258.32	26.00	0.507	81
2	215.23	236.34	11.87	0.199	100/échec

Tableau 3.4: Résultats de l'algorithme de Sturm avec consensus aléatoire, pour le mouvement 1

bruit	épipole trouvé		rayon d'erreur	% consensus	essais
0	2033.80	1409.83	0.07	0.966	1
0.01	2031.63	1408.18	194.54	0.810	2
0.1	1797.28	1264.76	370.93	0.569	22
0.2	1666.87	1179.85	535.60	0.397	100/échec
0.5	160.90	459.66	11.72	0.086	100/échec
1	166.13	415.36	12.34	0.138	100/échec

Tableau 3.5: Résultats de l'algorithme de Sturm avec consensus aléatoire, pour le mouvement 2

tivement complexes qui y sont très sensibles. Nous devons donc nous tourner à présent vers des approches prenant en compte le bruit et la redondance des données de manière explicite.

3.3 Résolution approximative: la méthode des cubiques

3.3.1 Minimisation d'un critère algébrique

Nous avons vu que dans la méthode de Sturm deux cubiques f et g sont déterminées, sur lesquelles doit se situer l'épipole. Une approche permettant d'intégrer la donnée d'un grand nombre de points est de calculer plusieurs de ces cubiques, puis d'essayer de trouver un point d'intersection qui leur soit toutes communes. En raison de la présence de bruit dans les données, nous allons rechercher ce point d'intersection de manière approximative, au sens d'un moindres carrés. Plus précisément, si nous pouvons calculer n cubiques $f_i(\mathbf{x}_c)$ et $g_i(\mathbf{x}_c)$ dans un repère commun, nous cherchons le point $\mathbf{x}_c$ qui minimise la somme:

$$S(\mathbf{x}_c) = \sum_{i=1}^{n} (f_i(\mathbf{x}_c)^2 + g_i(\mathbf{x}_c)^2) \tag{3.13}$$

Dans cette expression, $f_i(\mathbf{x}_c)$ et $g_i(\mathbf{x}_c)$ désignent les cubiques écrites dans le repère rétinien original de la caméra, qui sont obtenues à partir des cubiques f et g définies par les formules (3.4) (3.6) (3.8) (3.9) en effectuant un changement de repère inverse:

$$f_i(\mathbf{x}_c) = f(\mathbf{Q}^{-1}\mathbf{x}) \qquad g_i(\mathbf{x}_c) = g(\mathbf{Q}^{-1}\mathbf{x})$$

la matrice $\mathbf{Q}$ étant donnée pour chaque sous ensemble de sept points par l'équation (3.12).

Le polynôme S est homogène, du sixième degré. En fixant l'une des variables à 1, nous obtenons un polynôme du sixième degré à deux variables, $S(x, y)$. Montrons à présent qu'il est possible de trouver son minimum de manière directe. On peut d'ailleurs remarquer que cette directe méthode de minimisation d'un polynôme à deux indéterminées est applicable quel que soit le degré de celui-ci. Nous calculons $S_x(x, y)$ et $S_y(x, y)$, les dérivées partielles de S par rapport à x et y, qui doivent s'annuler en tout extremum de S. Le résultant de $S_x(x, y)$ et $S_y(x, y)$ par rapport à x, $R(y)$, est un polynôme de degré 25 à une indéterminée dont il est possible d'obtenir de manière précise toutes les racines y_i par une méthode itérative. Pour chacune de ces 25 racines, nous résolvons les deux équations du cinquième degré $S_x(x, y_i) = 0$ et $S_y(x, y_i) = 0$ et nous obtenons x_i comme racine commune de ces deux équations par simple comparaison. Il ne nous reste plus qu'à comparer les 25 valeurs $S(x_i, y_i)$ pour obtenir le minimum, tout en gardant en vue la possibilité de ne pas exclure d'autres minima locaux.

Nous avons implémenté cette approche uniquement à l'aide du système de calcul formel MAPLE. En effet, nous avons constaté expérimentalement que les opérations d'élimination et de calcul de racines de polynômes de degré élevé nécessitent de mener les calculs avec une précision très élevée, de l'ordre de 30 à 40 digits. Avec une précision de 10 digits, de très mauvais résultats sont obtenus, même dans le cas non-bruité. Quelques résultats, obtenus en

utilisant dix cubiques, sont présentés dans les tables (3.6) et (3.7). Ils sont assez décevants. Comme d'autres exemples nous ont permis de le constater, le comportement typique de cette méthode de calcul dépend de la présence où non de l'épipole dans l'image. Dans le cas où l'épipole en est assez éloigné, il se retrouve avec une certaine précision parmi les minima de $S(x, y)$, jusqu'à un certain niveau de bruit à partir duquel il disparaît complètement, mais il ne correspond pas en général au minimum global, celui-ci étant obtenu dans l'image. En revanche, dans le cas où il est dans l'image, il est localisé avec une précision qui est plutôt satisfaisante, et correspond bien au minimum global. La principale cause du problème est que le critère (3.13) ne correspond pas à la quantité que nous souhaitons effectivement minimiser, qui est la somme des distances du point $\mathbf{x}_c$ aux cubiques $f_i(\mathbf{x}_c)$ et $g_i(\mathbf{x}_c)$. Le critère (3.13) tend à privilégier les points de l'image, car ceux-ci donnent, à distance égale, des résidus plus faibles. Cette observation explique bien le comportement de la méthode, tel qu'il a été illustré dans les tables (3.6) et (3.7).

bruit	minima trouvés		valeur du critère
0	**214.06**	**255.64**	10^{-13}
	174.92	323.41	.00027
	188.90	293.18	.00033
0.01	**214.27**	**255.52**	10^{-6}
	175.08	322.85	.00027
	188.62	293.36	.00033
0.1	**215.58**	**254.54**	.000041
	176.93	316.59	.00029
	186.13	295.20	.00031
0.2	**211.63**	**256.32**	.00018
0.5	**184.22**	**280.70**	.00030
1	**178.20**	**278.28**	.00049
	300.12	240.88	.02949
	265.16	213.82	.03170
	253.73	215.18	.03182
	278.54	218.25	.03191
2	**175.63**	**314.82**	.00596
	149.35	376.04	.00703
	165.20	361.87	.03853
	228.23	117.79	.2526
	194.48	113.32	.2927
	10.68	25.23	.5132
	68.18	58.68	.6871

Tableau 3.6: Résultats de l'algorithme algébrique d'intersection, pour le mouvement 1

Il est malheureusement très difficile d'obtenir une expression analytique de la distance, ce qui nous a conduit à rechercher une solution graphique que nous allons présenter plus bas. Auparavant, nous présentons deux séries de graphes pour montrer comment se comporte l'intersection des cubiques en présence de bruit. La première série est obtenue avec le

bruit	minima trouvés		valeur du critère
0	**2033.79**	**1409.82**	10^{-6}
	163.74	343.79	.00021
	206.55	206.59	.00984
	196.97	237.34	.01200
	300.34	256.43	.02244
	267.39	214.68	.04562
	1404.3	974.87	20.21
0.01	163.73	343.81	.00021
	206.60	206.59	.00984
	196.98	237.37	.01201
	300.22	256.39	.02236
	267.19	215.01	.04461
	2041.6	**1415.9**	.15637
	1410.8	979.81	20.61
0.1	163.66	343.91	.00022
	207.09	206.65	.00980
	197.02	237.64	.01208
	299.07	255.97	.02165
	265.84	217.85	.03695
	2008.0	**1398.8**	17.70
	1528.4	1065.9	26.73
0.2	163.59	344.02	.00023
	207.68	206.75	.00975
	197.07	237.94	.01216
	297.53	255.31	.02083
	265.07	220.81	.03075
0.5	163.40	344.30	.00026
	209.82	207.24	.00961
	197.21	238.74	.01241
	290.94	251.80	.01829
	265.95	228.83	.02029
1	163.10	344.62	.00033
	215.15	209.04	.00934
	197.42	239.82	.01289

Tableau 3.7: Résultats de l'algorithme algébrique d'intersection, pour le mouvement 2

mouvement 1. La figure 3.5, représente l'image avec les quatre même cubiques obtenues avec des niveaux de bruit respectifs de 0.1, 0.5, et 1 pixel. On constate que si la situation est peu claire, avec un nombre important d'intersections parasites, le point de rencontre des courbes demeure à peu près au centre de l'image.

La seconde série est obtenue avec le mouvement 2. Une vue d'ensemble du plan rétinien avec cinq cubiques est donnée aux niveaux de bruit de 0.1 et 1 pixel, dans la figure 3.6. Un détail au voisinage de l'épipole, dans une fenêtre de dimension 900×900, est donné pour les niveaux de bruit 0.1, 0.5, et 1 pixel, dans le bas de cette figure. Sur ces graphes, on peut constater que la zone d'intersection est nettement plus délocalisée que dans le cas du mouvement précédent.

3.3.2 Une approche graphique

La méthode graphique que nous exposons est également générale et permet de trouver des intersections approchées de n courbes polynômiales. L'idée est de partitionner une partie du plan en cellules, et de vérifier le passage de chaque courbe dans chaque cellule au moyen d'un échantillonnage sur les deux axes suffisamment fin. L'algorithme de base, qui est particulièrement simple, est présenté dans l'encadré plus bas. Cet algorithme exploite une idée similaire aux transformations de Hough. Il est fondé sur l'idée qu'étant donné une courbe polynômiale plane, en fixant l'une des deux variables, les valeurs possibles de l'autre variable sont obtenues simplement en résolvant une équation polynômiale. L'utilisation d'un échantillonnage suivant l'axe des abscisses suffisamment fin par rapport aux tailles des cellules permet d'obtenir, pour chaque colonne de cellules, au moins une cellule marquée par branche de courbe. Il reste le cas des zones où la pente de la branche de courbe est très forte, qui est traité grâce à un échantillonnage le long de l'axe des ordonnées.

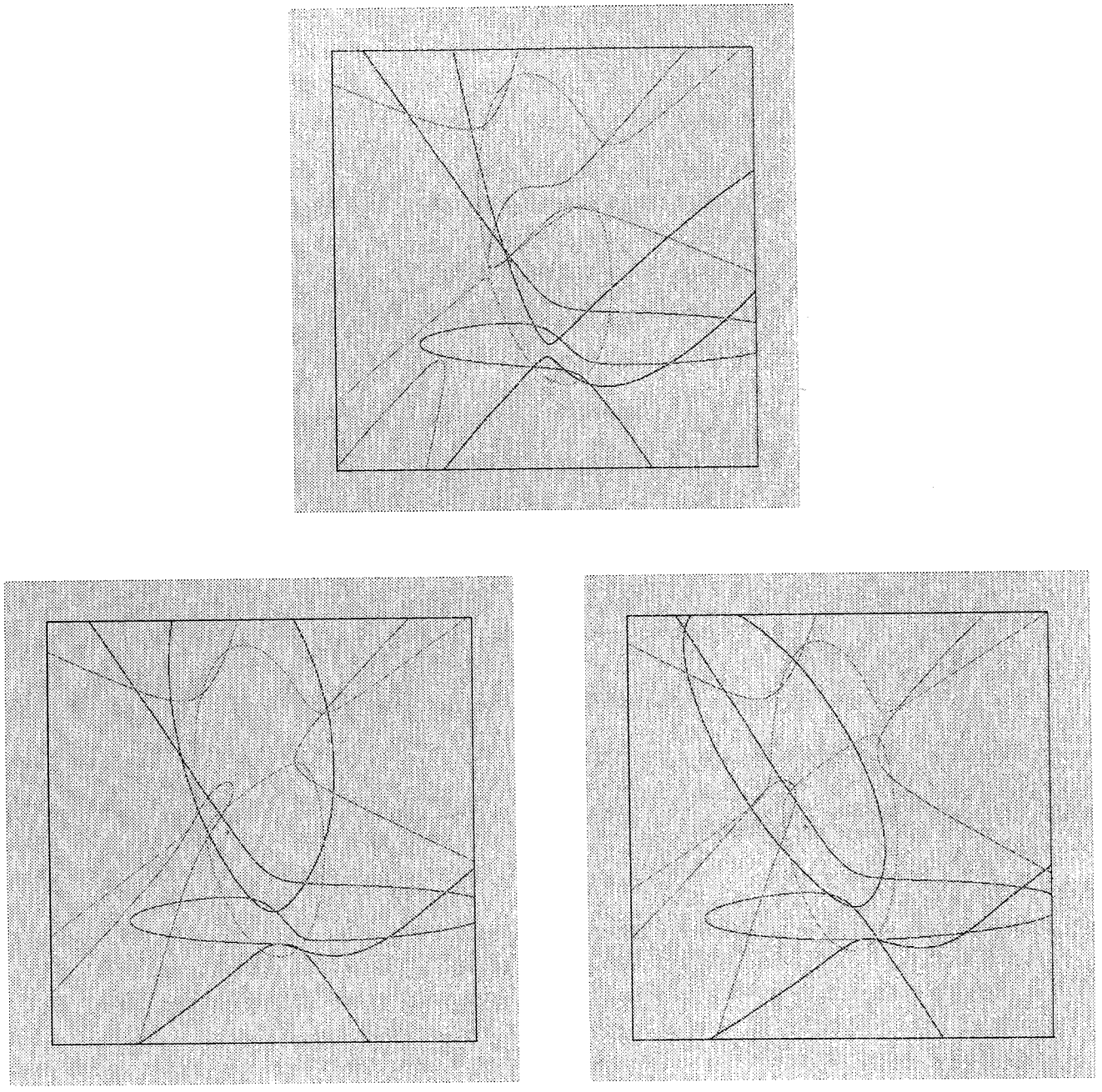


Fig. 3.5: Quatre cubiques obtenues avec le mouvement 1, à 0.1, 0.5 et 1 pixel de bruit

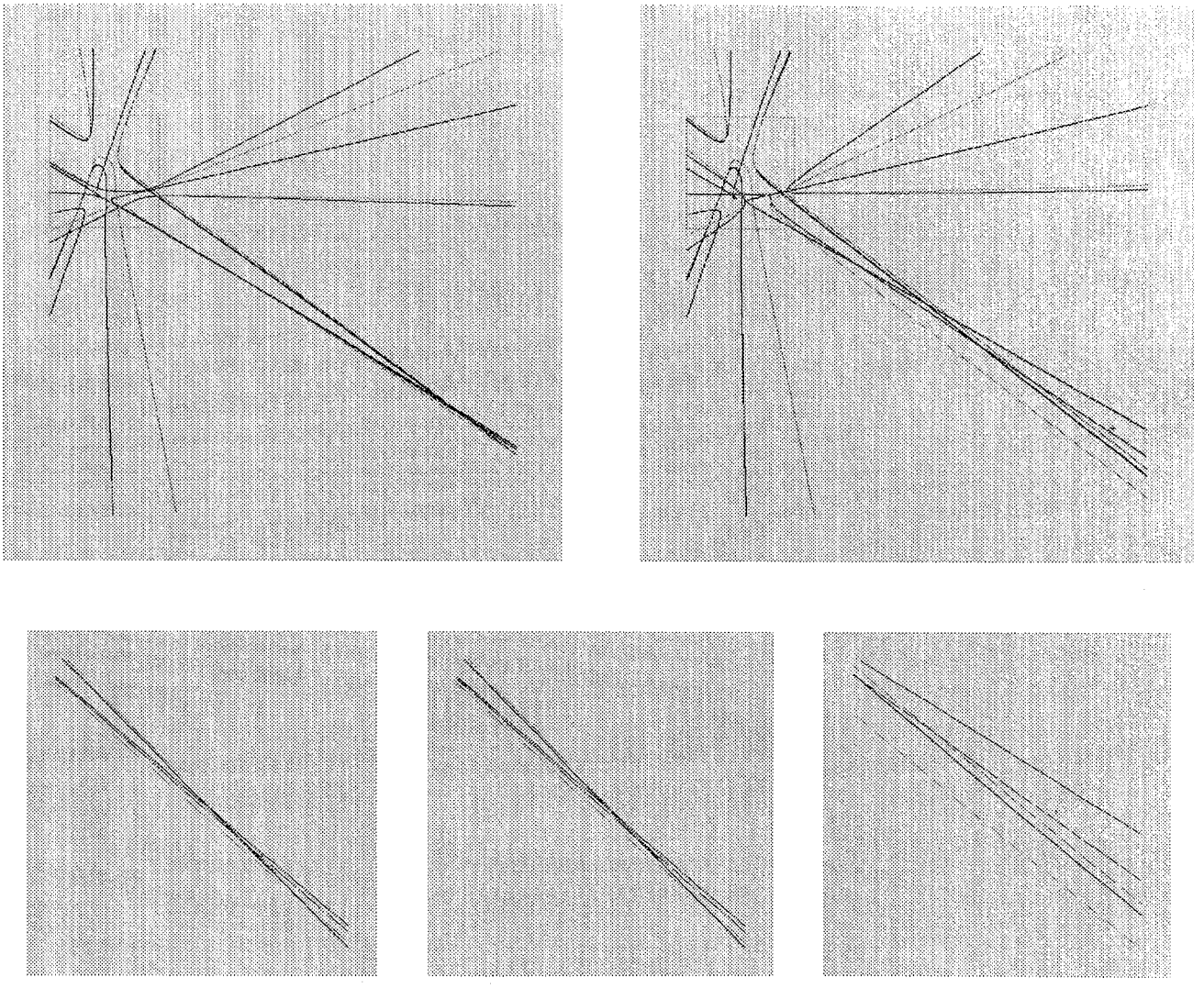


Fig. 3.6: Cinq cubiques obtenues avec le mouvement 2, à 0.1, (0.5) et 1 pixel de bruit

intersection approximative

- Diviser la zone de recherche $[x_{min}, x_{max}] \times [y_{min}, y_{max}]$ en cellules rectangulaires $dim_x \times dim_y$
- Pour chaque sous-ensemble i de 7 appariements, calculer les deux cubiques $c_{2i} = f$ et $c_{2i+1} = g$
- Pour chaque cubique c_i
  - Pour x variant de x_{min} à x_{max}
    - résoudre l'équation $c_i(x, y) = 0$ avec $y_{min} \leq y \leq y_{max}$
    - marquer toutes les cellules contenant au moins un point satisfaisant l'équation précédente avec l'indice i
  - Pour y variant de y_{min} à y_{max}
    - résoudre l'équation $c_i(x, y) = 0$ avec $x_{min} \leq x \leq x_{max}$
    - marquer toutes les cellules contenant au moins un point satisfaisant l'équation précédente avec l'indice i
- Effectuer un classement des cellules en fonction du nombre de cubiques rencontrées
- Conserver les N premières cellules.

Les paramètres d'entrée de cet algorithme sont la taille de la zone de recherche $[x_{min}, x_{max}] \times [y_{min}, y_{max}]$, la taille des cellules $dim_x \times dim_y$ et le nombre N de cellules à retenir.

Le choix optimal de la taille des cellules et du nombre de cellules retenues est assez délicat. Il est fonction du niveau de bruit estimé. En présence d'un bruit très faible, il est possible d'utiliser des cellules petites (de l'ordre de quelques pixels), et de ne retenir que celle rencontrant toutes les cubiques. Il n'est pas possible de procéder ainsi en présence d'un bruit assez important, car en prenant des cellules trop petites aucune intersection ne serait obtenue. Il est donc nécessaire de définir des cellules assez grandes, mais alors le problème va provenir des intersections parasites. En particulier, si la taille de la cellule s'approche de celle de l'image, comme toutes les cubiques passent par définition dans l'image, l'algorithme détectera toujours des cellules d'intersection dans l'image. De toutes manières il n'est plus possible dans ces conditions de localiser l'épipole avec certitude. Le meilleur choix de la taille des cellules est dans ce cas celui des valeurs pour lesquelles l'épipole exact se trouvera dans une cellule de rang le plus faible possible possible. Il y a une taille optimale car si, d'une part une taille trop petite est choisie, la cellule ne rencontrera pas un nombre de courbes suffisamment grand pour la distinguer d'autres cellules, d'autre part si une taille trop grande est choisie, la localisation ne sera pas bonne et il pourra y avoir de nombreuses autres cellules qui rencontreront autant de cubiques.

Les tables 3.3.2 et 3.3.2 présentent les résultats de localisation obtenus à différents niveaux de bruit et avec différentes tailles de cellules carrées. Les mêmes dix cubiques que celles de l'exemple précédent ont été utilisées, la recherche a été limitée à la zone $[-10000, 10000] \times [-10000, 10000]$. Dans chaque case d'un tableau est indiqué le nombre de cubiques passant dans la cellule contenant l'épipole exact, puis le rang de cette cellule, compte tenu des ex-aequos. On constate en particulier qu'à partir d'un certain niveau de bruit, il n'est plus toujours possible, quelle que soit la taille des cellules, de localiser la cellule contenant l'épipole exact puisque celle-ci n'est plus classée en premier rang, ou ne l'est plus de manière unique. On peut aussi noter la différence de comportement de l'algorithme selon que l'épipole se trouve dans l'image ou qu'il en est éloigné, le second cas donnant lieu à des nettement moins bonnes localisations, qui s'expliquent si on revient aux graphes donnés plus haut. Cette différence de comportement, que nous avons déjà signalée lors de la présentation de l'algorithme d'intersection algébrique, est consistante avec les résultats obtenus précédemment par cet algorithme. A titre d'exemple, la figure 3.7 illustre les 14 premières cellules de taille 150×150 pixels trouvées pour le mouvement 2, avec un bruit de 0.5 pixel.

	taille des cellules carrées															
bruit	10		20		50		100		150		200		300		500	
0	10	1	10	1	10	1-2	10	1	10	1-2	10	1	10	1-2	10	1
0.01	10	1	10	1	10	1-2	10	1	10	1-2	10	1	10	1-2	10	1
0.1	8	1	9	1	9	2	10	1	10	1-2	10	1	10	1-2	10	1
0.2	3	17-143	7	1	9	1-2	10	1	10	1-2	10	1	10	1-2	10	1
0.5	3	18-111	4	7-34	5	10-22	9	1-2	10	1	10	1	10	1-2	10	1
1	1	748-51337	3	34-109	5	12-26	10	1	10	1	10	1	10	1-2	10	1
2	0	58876-	2	146-717	3	58-110	6	10-16	8	4-8	10	1	10	1-2	10	1-2

Tableau 3.8: Résultats de l'algorithme graphique d'intersection, pour le mouvement 1

	taille des cellules carrées													
bruit	20		50		100		150		200		300		500	
0	10	1-5	10	1-7	10	1-10	10	1-11	10	1-15	10	1-13	10	1-
0.01	8	7-18	10	1-7	10	1-11	10	1-12	10	1-15	10	1-13	10	1-
0.1	7	3-13	8	2-13	8	9-26	10	1-12	10	1-13	10	1-17	10	1-
0.2	3	205-455	6	22-48	8	4-19	8	10-24	8	13-24	10	1-13	10	1-
0.5	1	3335-34163	4	59-186	7	14-30	8	10-20	8	8-17	8	11-20	10	1-
1	0	34732-	2	359-1971	4	62-131	6	21-38	7	14-22	8	8-17	6	24-
2	0	36527-	1	1192-13779	2	188-965	3	66-169	3	70-163	6	15-22	4	38-

Tableau 3.9: Résultats de l'algorithme graphique d'intersection, pour le mouvement 2

Nous faisons à présent une remarque similaire à celle que nous avions faite quant à la méthode algébrique d'intersection des cubiques: les quantités intervenant dans la méthode

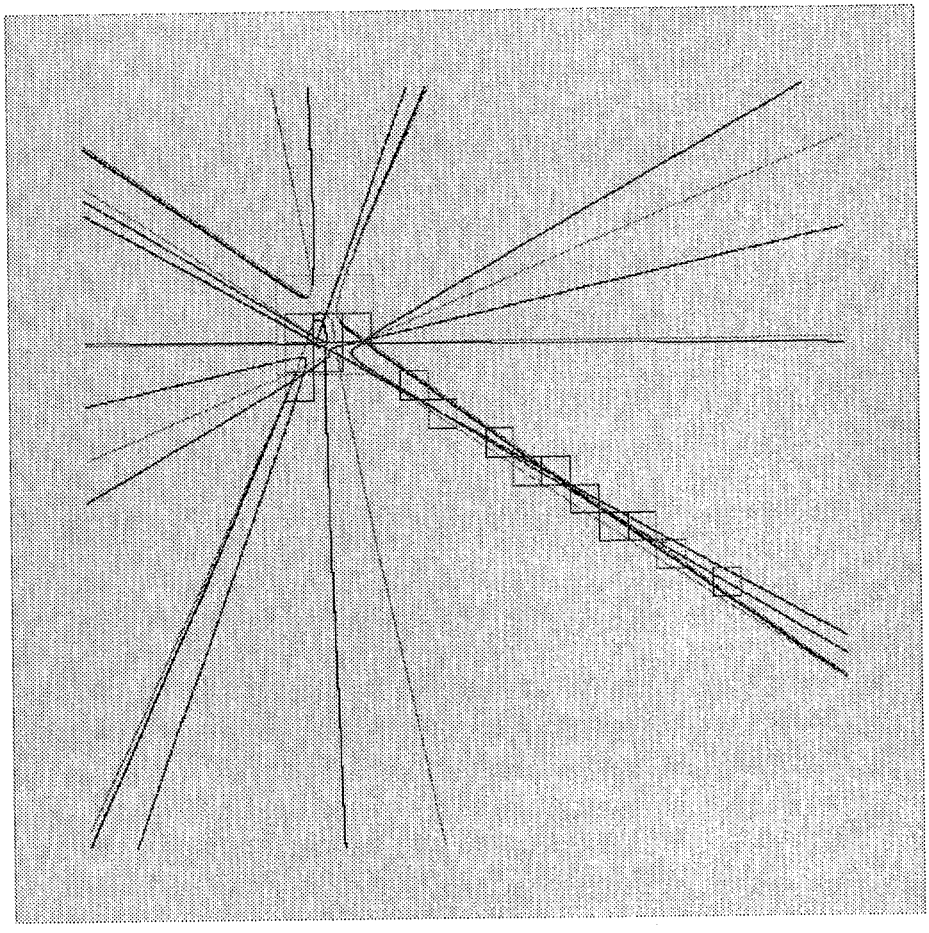


Fig. 3.7: Un exemple de cellules trouvées par la méthode des cubiques

de Sturm, qui ont une signification en terme de mesures dans l'image sont les birapports de droites épipolaires. Or en transformant les égalités (3.2), nous multiplions les deux membres par les dénominateurs, ce qui revient, lorsque nous cherchons à minimiser des quantités pour résoudre les équations finalement obtenues, de manière approchée, à pondérer celles-ci par des quantités qui sont fonction des coordonnées des points. Il est connu que minimiser un critère du type $\sum(\frac{a_i}{b_i} - \frac{a'_i}{b'_i})^2$ n'est pas du tout équivalent à minimiser $\sum(a_i b'_i - a'_i b_i)^2$ car dans le second cas on pondère de manière non intentionnelle par $b_i b'_i$. Pour savoir si il s'agit effectivement de la raison pour laquelle toutes les méthodes présentées jusqu'à présent ne peuvent pas localiser des épipoles éloignés, en présence de bruit, nous allons donc renoncer aux cubiques et revenir aux birapports. Mais auparavant, nous expliquons l'usage qu'il est malgré tout possible de faire de la méthode des cubiques pour la résolution de notre problème.

3.3.3 L'usage de la méthode des cubiques pour la recherche de candidats

Les méthodes de calcul de la matrice fondamentale que nous étudierons ultérieurement feront intervenir des minimisations non-linéaires. Il est connu que l'un des problèmes à résoudre pour utiliser ces algorithmes avec succès est le choix de bonnes initialisations, puisque généralement il y plusieurs minima locaux. Or nous souhaitons trouver le meilleur minimum de notre critère. Si on ne peut pas disposer d'estimées a priori suffisamment fiables, une manière de procéder est d'effectuer une tesselation de l'espace de recherche, puis d'exécuter la minimisation en initialisant avec chacun des points ainsi obtenu. Cette approche s'avère cependant complètement impraticable si la dimension de l'espace de recherche est élevée. Une alternative est alors d'essayer de sélectionner un certain nombre fixe de candidats possibles. en utilisant diverses heuristiques. Nous proposons une telle heuristique, au moyen de la méthodes des cubiques, de la contrainte d'ordre épipolaire, et de l'homographie épipolaire.

Construction d'une liste de paires d'épipoles

- Appliquer la méthode des cubiques dans l'image 1 et retenir les N_1 premières cellules dans $\mathcal{E}_1$.
- Appliquer la méthode des cubiques dans l'image 2 et retenir les N_1 premières cellules dans $\mathcal{E}_2$.
- Pour chaque élément de $\mathcal{E}_1 \times \mathcal{E}_2$,
  - vérifier la contrainte d'ordre épipolaire, et rejeter l'élément si elle n'est pas satisfaite.
  - calculer l'homographie épipolaire aux moindre carrés, et le résidu ϵ_{ij} obtenu.
- Trier les éléments de $\mathcal{E}_1 \times \mathcal{E}_2$ suivant les valeurs ϵ_{ij}, et conserver les N_2 premiers.

L'avantage de cette approche est de tirer parti du fait que l'espace de recherche pour le calcul des épipoles selon la méthode des cubiques est seulement de dimension deux, et de pouvoir utiliser des contraintes fortes pour restreindre la liste finale.

3.4 Une approche itérative: la méthode des birapports

3.4.1 Principe

L'approche proposée dans cette section est de partir directement de birapports de droites épipolaires, et de minimiser directement leurs différences par des méthodes d'optimisation non-linéaire. Le principe est de partir d'un ensemble de correspondances $\mathbf{q}_i \leftrightarrow \mathbf{q}'_i$, $\mathbf{q}_i$, $i \in \{1, \ldots, n\}$, et de considérer des sous-ensembles de quatre correspondances $j = \{j_1, j_2, j_3, j_4\}$, j_1, j_2, j_3, j_4 étant des des indices de $\{1, \ldots, n\}$, tous différents. Pour chacun de ces sous-ensembles j, nous calculons le birapport des droites épipolaires de la première image passant par les points $\mathbf{q}_{j_1}$, $\mathbf{q}_{j_2}$, $\mathbf{q}_{j_3}$ et $\mathbf{q}_{j_4}$:

$$\lambda_j = \{\mathbf{eq}_{j_1}, \mathbf{eq}_{j_2}, \mathbf{eq}_{j_3}, \mathbf{eq}_{j_4}\}$$

où $\mathbf{e}$ est l'épipole dans la première image. Nous obtenons de même le birapport λ'_j des droites épipolaires de la seconde image passant par les points $\mathbf{q}'_{j_1}$, $\mathbf{q}'_{j_2}$, $\mathbf{q}'_{j_3}$ et $\mathbf{q}'_{j_4}$. En écrivant l'égalité des birapports λ_j et λ'_j, nous obtenons une équation contraignant la position des épipoles $\mathbf{e}$ et $\mathbf{e}'$. Nous avons quatre inconnues, les coordonnées de ces deux épipoles, donc il nous faut en général quatre sous-ensembles de correspondances. Nous avons vu avec la méthode de Sturm que si nous nous autorisons à réutiliser des points dans plusieurs birapports, nous tombons sur des solutions multiples qui peuvent être délicates à discriminer. Si nous disposons d'un

plus grand nombre de points, nous pouvons rechercher les épipoles comme les points $\mathbf{e}$ et $\mathbf{e}'$ qui minimisent le critère:

$$C(\mathbf{e}, \mathbf{e}') = \sum_j (\lambda_j - \lambda'_j)^2 \tag{3.14}$$

Nous pouvons interpréter les birapports comme des angles de droites, ce qui rend clair le fait que lorsque les épipoles sont très éloignés, le calcul sera moins précis: les différences d'angles ne sont que très peu modifiées par un large déplacement d'un épipole. En revanche, si par exemple l'épipole se trouve dans l'image, nous voyons facilement qu'un déplacement de celui-ci de quelques dizaines de pixels induirait une variation non négligeable dans les angles de droites formées avec les points appariés. Nous retrouvons ainsi les conclusions tirées des descriptions précédentes de l'algorithme.

3.4.2 Prise en compte de l'incertitude

Toutes les configurations de quatre points ne donnent pas des birapports également stables. Par exemple, le fait de confondre deux points conduit à une singularité du birapport, aussi on peut s'attendre à ce que sa variance soit élevée lorsque la distance entre deux points est comparable à la variance de la position des points.

Nous commençons par calculer les variances de λ_j et λ'_j. Elles sont données par:

$$\sigma^2_{\lambda_j} = \begin{bmatrix} \frac{\partial \lambda_j^T}{\partial \mathbf{q}_{j_1}} & \frac{\partial \lambda_j^T}{\partial \mathbf{q}_{j_2}} & \frac{\partial \lambda_j^T}{\partial \mathbf{q}_{j_3}} & \frac{\partial \lambda_j^T}{\partial \mathbf{q}_{j_4}} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{j_1} & & & \\ & \mathbf{\Lambda}_{j_2} & & \\ & & \mathbf{\Lambda}_{j_3} & \\ & & & \mathbf{\Lambda}_{j_4} \end{bmatrix} \begin{bmatrix} \frac{\partial \lambda_j}{\partial \mathbf{q}_{j_1}} \\ \frac{\partial \lambda_j}{\partial \mathbf{q}_{j_2}} \\ \frac{\partial \lambda_j}{\partial \mathbf{q}_{j_3}} \\ \frac{\partial \lambda_j}{\partial \mathbf{q}_{j_4}} \end{bmatrix} \tag{3.15}$$

où les Λ_{j_k} $1 \leq k \leq 4$ sont les matrices de covariance des points $\mathbf{q}_{j_k}$, qui sont mutuellement indépendants. Nous faisons de plus l'hypothèse classique que la variance des points est isotrope et uniforme, c'est-à-dire:

$$\mathbf{\Lambda}_{j_1} = \ldots = \mathbf{\Lambda}_{j_4} = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix}$$

L'équation (3.15) se simplifie alors en:

$$\sigma^2_{\lambda_j} = \sigma^2 \|\nabla \lambda_j\|^2$$

où $\nabla \lambda_j$ désigne de gradient de λ_j par rapport au vecteur de dimension huit $(\mathbf{q}_{j_1}, \ldots, \mathbf{q}_{j_4})^T$, constitué par les coordonnées des points $\mathbf{q}_{j_k}$. On voit donc que la variance du birapport est proportionnelle á la variance des points mesurés. En revanche il n'est pas du tout correct de dire comme dans [21] que la variance du birapport est proportionnelle au birapport lui-même. Tout au plus, lorsque le représentant[2] est choisi entre 2 et $+\infty$, nous avons vérifié que

[2] voir les explications plus bas

la variance croissait sensiblement comme le birapport, tendant vers $+\infty$ quand deux points tendent l'un vers l'autre.

L'idée est à présent de pondérer chaque terme du critère (3.14) par la variance de $\lambda_j - \lambda'_j$, de sorte que la contribution des birapports au critère soit inversement proportionnelle à leur variance, ce qu permet de réduire l'influence des configurations instables de points. Comme λ_j et λ'_j sont indépendants, la variance de $\lambda_j - \lambda'_j$ est $\sigma^2_{\lambda_j} + \sigma^2_{\lambda'_j}$, ce qui nous conduit au nouveau critère:

$$C_1(\mathrm{e},\mathrm{e}') = \sum_j \frac{(\lambda_j - \lambda'_j)^2}{\sigma^2_{\lambda_j} + \sigma^2_{\lambda'_j}} \tag{3.16}$$

3.4.3 L'ordre dans les birapports

Le birapport dépend de l'ordre dans lequel les points sont numérotés. Il est connu que les 24 possibilités de numérotation donnent naissance à six valeurs distinctes. Un des problèmes que nous rencontrons lors de l'utilisation de la méthode des birapports est que la numérotation que nous déterminons en choisissant des sous-ensembles de quatre points est complètement arbitraire et va nous entraîner, pour chaque sous-ensemble, à calculer l'une des six valeurs plutôt qu'une autre, sans raison. Les six différentes valeurs et leurs relations sont:

$$\begin{aligned}
\{x_1, x_2, x_3, x_4\} &= \lambda \\
\{x_1, x_3, x_2, x_4\} &= 1 - \lambda \\
\{x_2, x_1, x_3, x_4\} &= \frac{1}{\lambda} \\
\{x_2, x_3, x_1, x_4\} &= 1 - \frac{1}{\lambda} \\
\{x_3, x_1, x_2, x_4\} &= \frac{1}{1-\lambda} \\
\{x_3, x_2, x_1, x_4\} &= 1 - \frac{1}{1-\lambda}
\end{aligned}$$

Il s'agit de l'action du groupe des permutations sur quatre symboles. Les invariants de cette action consistent en les symétries et on a donc:

$$\{x_1, x_2, x_3, x_4\} = \{x_2, x_1, x_4, x_3\} = \{x_3, x_4, x_1, x_2\} = \{x_4, x_3, x_2, x_1\}$$

Une première approche est d'utiliser pour chaque sous-ensemble de quatre points le sextuplet constitué par les six représentants cités plus haut, et de mesurer la différence entre les birapports obtenus à l'aide de quatre points par une norme. Dans cette optique, il suffit d'utiliser les trois représentants λ, $\frac{1}{\lambda}$ et $\frac{1}{1-\lambda}$, puisque seules les différences interviennent. Nous avons en premier lieu implémenté cette approche, en utilisant la norme L^2 qui permet de conserver la dérivabilité, néanmoins elle complique un peu les calculs lorsque l'on cherche

à tenir compte de l'incertitude. Les résultats ont été décevants, comparables à ceux obtenus avec la fonction f présentée plus bas.

Examinons à présent les valeurs prises par les birapports. On constate que l'on peut définir une partition de $\mathbb{R}$ en prenant les intervalles bornés par les valeurs successives $-\infty$, -1, 0, $\frac{1}{2}$, 1, 2, $+\infty$, qui possède la propriété suivante: étant donné quatre points, et les six valeurs de birapport obtenues avec les permutations de ces points, chacun des six intervalles précédemment défini contient une et une seule de ces six valeurs. On peut noter que parmi ces valeurs se trouvent celles correspondant à une configuration dans laquelle deux des quatre points sont confondus: $-\infty$, 0, 1, $+\infty$. Le tableau ci-dessous présente l'action des permutations sur les valeurs du birapport.

	1234	1324	2134	2314	3124	3214
	λ	$1-\lambda$	$\frac{1}{\lambda}$	$1-\frac{1}{\lambda}$	$\frac{1}{1-\lambda}$	$1-\frac{1}{1-\lambda}$
λ	$-\infty$ -1	2 $+\infty$	-1 0	1 2	0 $\frac{1}{2}$	$\frac{1}{2}$ 1
$1-\lambda$	-1 0	1 2	$-\infty$ -1	2 $+\infty$	$\frac{1}{2}$ 1	0 $\frac{1}{2}$
$\frac{1}{\lambda}$	0 $\frac{1}{2}$	$\frac{1}{2}$1	2 $+\infty$	$-\infty$ -1	1 2	-1 0
$1-\frac{1}{\lambda}$	$\frac{1}{2}$1	0 $\frac{1}{2}$	1 2	-1 0	2 $+\infty$	$-\infty$ -1
$\frac{1}{1-\lambda}$	1 2	-1 0	$\frac{1}{2}$1	0 $\frac{1}{2}$	$-\infty$ -1	2 $+\infty$
$1-\frac{1}{1-\lambda}$	2 $+\infty$	$-\infty$ -1	0 $\frac{1}{2}$	$\frac{1}{2}$1	-1 0	1 2

De ces remarques, on peut conclure qu'étant donné quatre points, et un des six intervalles, il est toujours possible de ramener leur birapport à l'intérieur de cet intervalle en appliquant une simple permutation sur l'ordre de ces points. Le choix de cette permutation peut être trouvé très simplement en considérant uniquement l'intervalle dans lequel se situe le birapport d'origine et celui voulu finalement. Nous avons incorporé cette seconde approche à notre algorithme de recherche. Son grand avantage est qu'elle ne nécessite qu'une simple comparaison suivie d'une permutation d'indices, et nous permet de réexploiter telles quelles les formules d'incertitude dérivées au paragraphe précédent. Le fait de tenir compte de l'ordre et de ramener tous les birapports dans un intervalle unique nous permet ainsi d'éviter d'obtenir des birapports d'ordre de grandeur différents, dont certains se trouveraient artificiellement favorisés au cours du processus de minimisation. Nous obtenons ainsi des différences comparables. Cependant il est possible d'avoir des très grandes valeurs (si deux points sont presque alignés avec l'épipole). Il faut donc éliminer ces configurations de points (les valeurs interdites sont 0,1,-1,∞). Le choix du représentant entre 0 et 0.5 s'est révélé le meilleur pour réduire l'influence de ces configurations. En effet, si nous choisissons un représentant pouvant atteindre des valeurs arbitrairement grandes (ce qui ce produit également si l'ordre n'est pas contrôlé), la situation "énergétique " est telle que les points choisis induisent un "quadrillage" du plan par leurs droites d'intersection mutuelle, chaque zone du plan étant séparée de ses voisines par une valeur asymptotique du critère.

Une autre approche permettant de résoudre la même difficulté est d'effectuer la minimisation sur les fonctions $f(\lambda) - f(\lambda')$, où λ et λ' sont les birapports et f la fonction algébrique

méthode	bruit 0.2 pixel		bruit 1.0 pixel	
	initialisation		initialisation	
	exacte	linéaire	exact	linéaire
birapports	0.2108	0.2969	0.3446	0.5987
birapports + C	0.1404	0.2200	0.2744	0.5496
birapports + f	0.2315	0.2880	0.3098	0.5847
birapports + P	0.1429	0.2305	0.2651	0.5718
birapports + f + C	0.2104	0.2813	0.3678	0.6297
birapports + P + C	0.1169	0.2158	0.2344	0.5395
matrice fondamentale	0.1006	0.1222	0.2581	0.3437

Tableau 3.10: Moyenne des distances relatives obtenus par la méthode des birapports

naturellement associée au birapport. En effet, la propriété caractéristique est de prendre les mêmes valeurs pour chaque orbite:

$$f(\lambda) = \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(1-\lambda)^2} \tag{3.17}$$

Toutefois, il est à craindre que l'emploi de cette fonction rationnelle de degré élevé en λ introduise des problèmes d'instabilité.

3.4.4 Comparaison: résultats expérimentaux

Nous avons utilisé la procédure expérimentale qui est présentée dans le chapitre suivant pour la comparaison des méthodes itératives. 100 mouvements ont servi à obtenir une caractérisation statistique de la précision du minimum et de certaines propriétés de convergence. La table 3.10 présente les moyennes des distances relatives obtenues avec les différentes versions de la méthode des birapports (C: normalisation de la covariance, f: utilisation de la fonction f définie en 3.17), P : prise en compte des permutations). Cette méthode reposant sur une minimisation non-linéaire, le point d'initialisation est bien sûr de la plus grande importance. Nous avons choisi de partir d'une part avec les épipoles exacts, d'autre part avec ceux obtenus par une méthode linéaire présentée dans le chapitre suivant, qui est la plus rapide (mais non la plus fiable) qui soit. La première initialisation nous permet de tester la stabilité du minimum, la seconde de voir dans quelle mesure les méthodes convergent en partant d'un point d'initialisation qui est assez imprécis.

Ces résultats montrent que la prise en compte de la covariance et des permutations conduit effectivement à une amélioration significative. Cependant l'emploi de la fonction f définie en (3.17) dégrade les résultats, ce qu s'explique par la forme compliquée de cette fonction. Lorsque l'initialisation se fait au point exact, nous observons un minimum qui est d'une précision comparable à celle obtenue par les méthodes de calcul direct de la matrice

bruit	birapports		birapports + C		birapports + P		birapports + C + P	
0	214.062	255.648	214.062	255.648	214.062	255.648	214.062	255.648
0.01	213.992	255.949	214.018	255.651	213.979	255.723	214.018	255.651
0.1	213.758	258.841	213.097	255.321	212.231	258.55	213.592	255.647
0.2	215.425	262.478	212.57	255.342	209.852	261.277	212.895	255.368
0.5	217.774	265.963	211.362	256.181	203.865	272.656	210.455	255.471
1	220.193	266.939	208.379	262.064	206.664	274.999	208.471	257.147
2	222.84	267.821	195.454	279.067	212.881	254.316	212.328	256.518

Tableau 3.11: Minimum exact du critère des birapports, pour le mouvement 1

bruit	birapports		birapports + C		birapports + P		birapports + C + P	
0	2033.81	1409.83	2033.8	1409.83	2033.8	1409.83	2033.8	1409.83
0.01	1903.05	1335.45	2035.55	1411.09	1947.83	1357.77	2035.55	1411.09
0.1	1246.02	960.101	2056.56	1425.43	1423.12	1043.82	2054.92	1424.38
0.2	1767.23	1108.08	2096.17	1450.8	1125.79	860.566	2085.77	1444.16
0.5	2028.21	1493.26	2134.76	1481.26	1001.31	792.913	2114.72	1465.22
1	2093.84	1381.41	2069.09	1446.9	1750.9	1100.31	2053.83	1427.84
2	1388.14	1108.41	2122.09	1484.4	1126.3	789.248	1605.51	1139.08

Tableau 3.12: Minimum exact du critère des birapports pour le mouvement 2

fondamentale, qui sont présentées dans le chapitre suivant. Ces faits sont confirmés par l'examen des minima obtenus pour les deux mouvements de référence utilisés précédemment, présentés tables 3.11 et 3.12. La précision est en particulier nettement supérieure à ce qui pouvait être obtenu par les méthodes précédentes.

En revanche, la méthode des birapports se révèle très sensible à l'initialisation, en tous cas beaucoup plus que la méthode de la matrice fondamentale, ce qui est en pratique un facteur limitant important. En particulier, si nous la lançons à partir des épipoles obtenus par la méthode non-linéaire de la matrice fondamentale (la plus précise), nous provoquons souvent une divergence qui dégrade les résultats.

Chapitre 4

Détermination de la matrice fondamentale

Les techniques linéaires d'estimation de la matrice essentielle admettent une extension naturelle qui permet d'effectuer le calcul direct de la matrice fondamentale à partir d'appariements de points, au moyen d'un critère qui est linéaire. Nous montrons que cette méthode souffre de deux défauts, liés à l'absence de contrainte sur le rang de la matrice recherchée, et à l'absence de normalisation du critère, qui entraînent des erreurs importantes dans l'estimation de la matrice fondamentale et des épipoles. Cette analyse est validée par plusieurs exemples réels. Afin de surmonter ces difficultés, plusieurs nouveaux critères non-linéaires, dont nous donnons des interprétations en termes de distances, sont ensuite proposés, puis plusieurs paramétrisations sont introduites pour rendre compte des contraintes auxquelles doit satisfaire la matrice fondamentale. Un travail expérimental exhaustif est réalisé à l'aide de nombreuses données synthétiques et d'images réelles. En particulier, une méthode statistique fondée sur la variation des déplacements de la caméra est utilisée pour évaluer la stabilité et les propriétés de convergence des différentes méthodes.

4.1 Le critère linéaire

L'algorithme des huit points de Longuet-Higgins admet une extension naturelle qui permet d'effectuer le calcul direct de la matrice fondamentale au moyen d'un critère qui est linéaire. Dans les cas pratiques, la résolution de ce critère s'effectue par des techniques de moindres carrés, dont nous présentons plusieurs implémentations simples et efficaces. Nous effectuons ensuite une analyse de ce critère qui nous permet de mettre en évidence ses deux défauts, dont nous montrons qu'ils entraînent des erreurs importantes dans l'estimation de la matrice fondamentale et des épipoles. L'identification de ces défauts nous donnera le moyen d'y remédier au moyen des critères non-linéaires examinés ultérieurement.

4.1.1 L'algorithme des huit points

Rappelons l'équation de base définissant la matrice fondamentale, et sur laquelle sont fondées toutes les méthodes de détermination étudiées dans ce chapitre:

$$\mathbf{q}'^T\mathbf{F}\mathbf{q} = 0 \tag{4.1}$$

Cette équation peut se réécrire:

$$\mathbf{U}^T\mathbf{f} = 0 \tag{4.2}$$

dans laquelle nous avons posé:

$$\begin{aligned}\mathbf{U} &= [uu', vu', u', uv', vv', v', u, v, 1]^T \\ \mathbf{f} &= [F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}]^T\end{aligned}$$

où $(u, v, 1)^T$ et $(u', v', 1)^T$ sont les coordonnées des points $\mathbf{q}$ et $\mathbf{q}'$ respectivement, qui sont à distance finie puisqu'ils sont mesurés dans les rétines. L'équation est linéaire et homogène en les neuf coefficients de la matrice $\mathbf{F}$. Nous en concluons qu'en général, si nous disposons de huit appariements, nous pouvons déterminer une solution unique pour $\mathbf{F}$, définie à un facteur d'échelle près, en résolvant un système linéaire composé de huit équations (4.2). Cette approche, qui constitue l'algorithme des huit points introduit par Longuet-Higgins [90], et retrouvé indépendamment par Tsai et Huang [163] a été très étudiée dans la littérature pour la détermination de la matrice *essentielle* et des paramètres du mouvement: un grand nombre d'algorithmes en sont partiellement [37, 36, 174] ou entièrement [152, 44] [145] [172, 170] [73] dérivés, et des études de sensibilité ont été menées [89] [172, 170] [86] [72] [128]. Il est en particulier connu que l'algorithme des huit points est très sensible au bruit, comme la majorité des algorithmes calculant le mouvement tridimensionnel à partir de deux vues[1].

Trivedi [155] avait déjà montré que la géométrie épipolaire peut être calculée dans le cas de caméras non calibrées, par une généralisation de l'algorithme des huit points. Quelques travaux récents ont précisé et clarifié [126] [40] [61] cette approche. Nous nous proposons, dans ce cadre de détermination de la matrice *fondamentale*, de mettre en évidence ses limites, et d'en proposer des améliorations.

[1]Pour une vue d'ensemble sur le problème du mouvement, situant les approches mentionnées dans ce cadre plus général, voir [120], [4], et [67].

4.1.2 Implémentations du critère linéaire

Les travaux originaux n'utilisaient que huit points. En pratique, nous disposons en général d'un nombre n bien plus grand d'appariements $(\mathbf{q}_i, \mathbf{q}'_i)$, $1 \leq i \leq n$, et nous résolvons les équations (4.2) aux moindres carrés, ce qui revient à minimiser par rapport à $\mathbf{F}$ le critère:

$$\min_{\mathbf{F}} \sum_i (\mathbf{q}_i'^T \mathbf{F} \mathbf{q}_i)^2 \tag{4.3}$$

qui s'écrit aussi, compte tenu des notations introduites en (4.2):

$$\min_{\mathbf{f}} \|\tilde{\mathbf{U}}\mathbf{f}\|^2 \quad \text{où} \quad \tilde{\mathbf{U}} = \begin{bmatrix} \mathbf{U}_1^T \\ \vdots \\ \mathbf{U}_n^T \end{bmatrix} \tag{4.4}$$

Ce critère est un moindres carrés linéaire et peut donc être minimisé par une solution explicite. La matrice fondamentale $\mathbf{F}$ étant définie à un facteur d'échelle près, deux méthodes de normalisation peuvent être adoptées.

La première méthode consiste à fixer à 1 l'une des neuf composantes du vecteur $\mathbf{f}$ de coefficients de la matrice fondamentale, puis à résoudre par rapport aux huit autres coefficients le critère (4.3). Celui-ci étant quadratique, il existe un minimum unique qui est calculé en résolvant un système linéaire (méthode que nous désignerons par le sigle **M-C**). L'inconvénient est que, comme nous ne disposons pas d'information a priori sur la matrice fondamentale, nous ne sommes pas à l'abri d'un mauvais choix du coefficient mis à 1. Par exemple, dans le cas d'une translation pure, la matrice fondamentale est antisymétrique, en conséquence il faut éviter de normaliser le dernier coefficient à 1, puisque sa valeur devrait être 0.

La seconde méthode est numériquement un peu moins simple, mais ne souffre pas de ce défaut. Elle consiste à fixer à 1 la norme du vecteur $\mathbf{f}$, et non plus une de ses composantes. Le critère (4.3) devient alors une minimisation sous contrainte:

$$\min_{\mathbf{f}} \|\tilde{\mathbf{U}}\mathbf{f}\| \quad \text{avec} \quad \|\mathbf{f}\| = 1 \tag{4.5}$$

La solution de ce problème classique est alors obtenue comme le vecteur propre associé à la plus petite valeur propre de $\mathbf{\Gamma} = \tilde{\mathbf{U}}^T\tilde{\mathbf{U}}$. Deux méthodes de calcul sont envisageables:

- Diagonalisation directe de $\mathbf{\Gamma}$ (méthode désignée par le sigle **DIAG**).
- Décomposition en valeurs singulières de $\tilde{\mathbf{U}}$ en $\mathbf{\Delta}\mathbf{\Lambda}\mathbf{\Theta}^T$ où les matrices $\mathbf{\Delta}$ et $\mathbf{\Theta}$ sont orthogonales et la matrice $\mathbf{\Lambda}$ est diagonale à coefficients positifs. La solution du problème (4.5) est obtenue en prenant le vecteur colonne de $\mathbf{\Theta}$ associé à la plus petite valeur singulière (méthode désignée par le sigle **SVD**).

A ces différents choix s'ajoute celui de normaliser les points projectifs $\mathbf{q}$ et $\mathbf{q}'$ intervenant dans l'équation (4.1) (méthode désignée par le sigle **DIAG-N**), c'est à dire de les prendre tels que $\|\mathbf{q}\| = \|\mathbf{q}'\| = 1$, comme le fait Kanatani [74, 75] au lieu de les conserver sous

la forme $(u, v, 1)^T$, $(u', v', 1)^T$. Nous pouvons donc voir que plusieurs méthodes numériques sont disponibles pour sa résolution, bien que le critère linéaire possède une solution explicite unique. Les résultats obtenus peuvent varier légèrement selon la méthode adoptée, ce qu'illustre la table 4.1.2 ci-dessous. Une expérimentation systématique sera présentée plus bas.

bruit	méthode	coordonnées des épipoles			
	EXACT	169.10	-399.48	286.07	-698.84
0.2 pixel	**SVD**	171.64	-418.82	284.58	-728.35
	DIAG	171.92	-419.18	285.13	-729.39
	M-C	171.96	-419.71	285.10	-730.05
	DIAG-N	175.92	-420.04	283.45	-726.57
0.6 pixel	**SVD**	222.40	-783.76	248.97	-1199.83
	DIAG	221.08	-765.95	255.75	-1184.63
	M-C	221.84	-770.20	255.19	-1190.01
	DIAG-N	251.19	-717.87	270.21	-1118.29
1 pixel	**SVD**	590.24	1836.32	-281.05	2176.95
	DIAG	756.85	2184.98	-500.83	2540.74
	M-C	743.97	2165.48	-485.89	2517.44
	DIAG-N	4509.21	1412.42	-4633.16	480.22

Tableau 4.1: Un exemple illustrant le comportement différent d'implémentations du critère linéaire

4.1.3 Difficultés liées au critère linéaire

Lorsque l'équation (4.1) est exactement satisfaite, la méthode linéaire est parfaitement justifiée. Les problèmes apparaissent lorsque ce n'est pas le cas, ce qui se produit avec des appariements bruités. La méthode linéaire présente l'avantage de conduire à une solution explicite par un calcul simple, cependant elle souffre de deux défauts, qui sont dus respectivement:

- Au fait que la contrainte $\det(\mathbf{F}) = 0$ ne soit pas imposée, ce qui entraîne une incohérence de la géométrie épipolaire au voisinage des épipoles. Ce problème est lié au *paramétrage* par huit coefficients. Celui-ci permet d'obtenir une solution linéaire, mais n'est pas capable d'exprimer la contrainte de rang. Spetsakis et Aloimonos remarquent qu'une contrainte similaire (qu'ils ne formulent pas) est importante dans le cadre du calcul du mouvement. Ils présentent une méthode pour en tenir compte [145]. Jerian et Jain [72] sont plus précis, affirmant que certains cas d'échec de la méthode linéaire sont dus au fait qu'une contrainte polynomiale équivalente à (2.6) n'est pas satisfaite.

- A l'absence de normalisation, qui entraîne un biais dans la localisation des épipoles. Ce problème est lié au *critère*, c'est-à-dire à la fonction d'erreur minimisée, et a déjà été mis en évidence dans le cadre du calcul du mouvement par Daniilidis et Nagel [25], qui ont obtenu des preuves analytiques en se plaçant dans un cas particulier.

L'absence de contrainte de rang Voyons tout de suite quelles sont les conséquences du premier point que nous venons de signaler. Considérons une droite épipolaire l' de la seconde image, calculée à partir d'une matrice fondamentale $\mathbf{F}$ obtenue par le critère linéaire, et d'un point $\mathbf{q} = (u, v, 1)^T$ de la première image. Nous pouvons réécrire $\mathbf{q}$ en faisant intervenir l'épipole dans la première image $\mathbf{e} = (e_1, e_2, 1)^T$, et les écarts x et y. Une représentation projective de l' est alors obtenue par:

$$\mathbf{l}' = \mathbf{F}\mathbf{q} = \mathbf{F}\begin{pmatrix} e_1 - x \\ e_2 - y \\ 1 \end{pmatrix} = \mathbf{F}\mathbf{e} - \underbrace{\mathbf{F}\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}}_{\mathbf{l}_1} \tag{4.6}$$

Si $\det(\mathbf{F}) = 0$, l'épipole $\mathbf{e}$ vérifie exactement $\mathbf{F}\mathbf{e} = 0$, et le dernier membre se réduit alors à son second terme, $\mathbf{l}_1$. Il est aisé de voir que celui-ci définit une droite épipolaire correcte, au sens où elle passe bien par l'épipole $\mathbf{e}'$ de la seconde image. Celui-ci est en effet alors défini par la relation $\mathbf{F}^T\mathbf{e}' = 0$, qui admet également une solution exacte, d'où $\mathbf{e}'^T\mathbf{l}_1 = 0$. En revanche, si la condition sur le déterminant n'est pas exactement satisfaite, nous pouvons observer que $\mathbf{l}'$ s'écrit comme somme d'un vecteur constant $\mathbf{r} = \mathbf{F}\mathbf{e}$, qui devrait être nul, mais ne l'est pas, et du vecteur $\mathbf{l}_1$., dont la norme est bornée par $\sqrt{x^2 + y^2}\|\mathbf{F}\|$. Nous pouvons conclure de cette observation d'une part que lorsque $(x, y) \to (0, 0)$, c'est-à-dire lorsque $\mathbf{q} \to \mathbf{e}$, la droite épipolaire de $\mathbf{q}$ dans la seconde image tend vers une droite fixe, de représentant $\mathbf{r}$, ce qui est absolument incohérent avec la notion de géométrie épipolaire, et d'autre part que plus $\sqrt{x^2 + y^2}$ est faible, c'est-à-dire plus le point $\mathbf{q}$ est proche de l'épipole, plus l'erreur sur sa droite épipolaire sera grande.

Une manière plus précise de voir ce fait est d'introduire une distance euclidienne. Il est bien connu que si $\mathbf{p}$ est un point de coordonnées (x_0, y_0), et l une droite d'équation $l_1x + l_2y + l_3 = 0$, la distance du point $\mathbf{p}$ à la droite l est:

$$d(\mathbf{p}, l) = \frac{|l_1x_0 + l_2y_0 + l_3|}{\sqrt{l_1^2 + l_2^2}}$$

Nous considérons la distance de la droite épipolaire $l'(\mathbf{q})$, définie comme précédemment, et de l'épipole $\mathbf{e}' = (e'_1, e'_2, 1)^T$. Elle s'écrit donc, après regroupement:

$$d(\mathbf{e}', l') = \frac{|r_1e'_1 + r_2e'_2 + r_3 - (F_{11}e'_1 + F_{21}e'_2 + 1)x - (F_{12}e'_1 + F_{22}e'_2 + 1)y|}{\sqrt{(r_1 - F_{11}x - F_{12}y)^2 + (r_2 - F_{21}x - F_{22}y)^2}}$$

Nous pouvons voir que lorsque $(x, y) \to (0, 0)$, $d(\mathbf{e}', l') \to \frac{r_1e'_1 + r_2e'_2 + r_3}{\sqrt{r_1^2 + r_2^2}}$, qui en pratique est une valeur assez grande.

Nous illustrons ces considérations par un exemple réel, obtenu à partir de la première paire d'images présentée plus bas. La matrice fondamentale calculée par le critère linéaire à partir des appariements présentés sur ces images vaut:

$$\mathbf{F} = \begin{pmatrix} .56761\ 10^{-6} & -.000018243 & .0093598 \\ .000018023 & .11202\ 10^{-5} & -.0079245 \\ -.011082 & .0067203 & 1.0000 \end{pmatrix}$$

A partir de cette matrice fondamentale, les épipoles obtenus en minimisant $\|\mathbf{Fe}\|$ et $\|\mathbf{F}^T\mathbf{e}'\|$ sont:

$$\mathbf{e} = (403.90, 603.24, 1)^T \quad \mathbf{e}' = (408.38, 524.63, 1)^T$$

La valeur des vecteurs-résidu $\mathbf{r} = \mathbf{Fe}$ et $\mathbf{r}' = \mathbf{F}^T\mathbf{e}'$ est:

$$\mathbf{r} = (.0000197, .0000274, .630\ 10^{-7})^T \quad \mathbf{r}' = (.0000203, .0000236, .152\ 10^{-6})^T$$

Ces valeurs sont faibles, puisque $\|r\| = 3.3\ 10^{-5}$ toutefois, à titre de comparaison, les résidus obtenus à partir d'un des critères non-linéaires présentés ultérieurement sont: $\|r\| = 2\ 10^{-8}$. La figure 4.1 montre un graphe de la fonction d'erreur constituée par la distance de l'épipolaire $l'_{\mathbf{q}}$ à l'épipole $\mathbf{e}'$, en fonction des écarts x et y. Toutes les unités sont des pixels. Nous pouvons observer qu'il y a un pic marqué au voisinage du point $(x, y) = 0$, qui correspond au point $\mathbf{e}$, et que l'erreur décroît effectivement lorsque le point $\mathbf{q}$ s'éloigne de $\mathbf{e}$, pour tendre vers une valeur assez faible. Sur la figure 4.2, qui permet une lecture plus précise, est représentée la courbe donnant l'erreur suivant l'axe $x = y$. Nous pouvons conclure qu'en particulier *si l'épipole se trouve dans l'image, la géométrie épipolaire décrite par la matrice* $\mathbf{F}$ *obtenue à partir du critère linéaire sera imprécise.*

L'absence de normalisation Donnons une interprétation géométrique du critère (4.3). Considérons la distance euclidienne d'un point $\mathbf{q}'$ de la seconde image à la droite épipolaire $\mathbf{l}' = (l'_1, l'_2, l'_3)^T = \mathbf{Fq}$ du point $\mathbf{q}$ de la première image. L'expression de cette distance est:

$$d(\mathbf{q}', \mathbf{l}') = \frac{|\mathbf{q}'^T\mathbf{l}'|}{\sqrt{(l'_1)^2 + (l'_2)^2}} \tag{4.7}$$

Notons que le terme normalisateur $w = \sqrt{(l'_1)^2 + (l'_2)^2}$ ne s'annule que dans les cas dégénérés où la droite épipolaire est à l'infini. Le critère (4.3) peut alors se réécrire:

$$\min_{\mathbf{F}} \sum_i w_i^2 d^2(\mathbf{q}'_i, \mathbf{l}'_i) \tag{4.8}$$

Cette interprétation fait apparaître le fait qu'une quantité géométriquement significative intervenant dans le critère linéaire est la distance d'un point à la droite épipolaire de son correspondant. Cette quantité se trouve pondérée par les coefficients w définis ci-dessus.

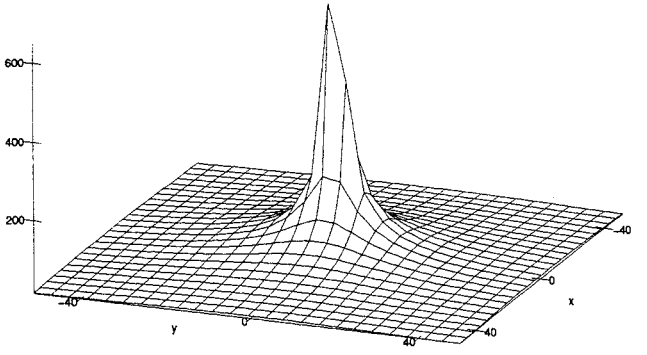


Fig. 4.1: Distances des droites épipolaires à l'épipole, dans le cas du critère linéaire

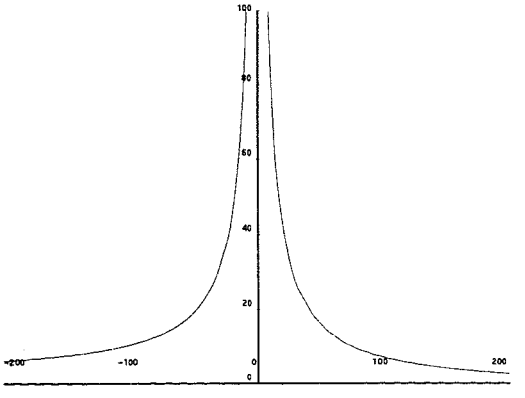


Fig. 4.2: Distances des droites épipolaires à l'épipole, dans le cas du critère linéaire (coupe)

Pour comprendre en quoi ceci peut fausser les résultats, considérons d'abord le cas particulier où le mouvement est une translation pure. Dans ce cas, la matrice fondamentale est antisymétrique et a la forme:

$$\lambda \begin{bmatrix} 0 & 1 & -y \\ -1 & 0 & x \\ y & -x & 0 \end{bmatrix}$$

où $(x, y, 1)^T$ sont les coordonnées affines des épipoles, qui se trouvent être identiques entre les deux images. Soit un point $\mathbf{q}_i$ de la première image, de coordonnées $(u_i, v_i, 1)^T$, alors le coefficient normalisateur vaut, $w_i^2 = \lambda^2((y - v_i)^2 + (x - u_i)^2)$, λ étant une constante. Minimiser le critère (4.8) revient à minimiser, pour chacun des points $\mathbf{q}_i$, à la fois w_i et $d^2(\mathbf{q}'_i, \mathbf{l}'_{\mathbf{q}_i})$. Or minimiser w_i revient à favoriser sans raison les matrices fondamentales qui donnent des épipoles se trouvant dans l'image. Les résultats expérimentaux confirment de manière systématique cette observation. Un exemple est illustré par la table 4.1.3.

bruit (pixel)	e_x	e_y	e'_x	e'_y
0	246.09	1199.34	246.09	1199.34
0.1	246.09	1193.49	245.06	1193.68
0.5	239.63	1027.16	253.20	1027.23
1.0	237.22	758.55	263.33	767.01
2.0	242.54	544.70	263.01	568.16

Tableau 4.2: Un exemple illustrant le comportement du critère linéaire lorsque le mouvement est une translation

Dans le cas général, le coefficient normalisateur vaut:

$$w_i^2 = (a(y - v_i) + b(x - u_i))^2 + (c(y - v_i) + d(x - u_i))^2$$

Pour voir son effet sur la minimisation, supposons les coefficients de l'homographie fixés. En calculant les dérivées partielles $\frac{\partial w_i}{\partial x}$ et $\frac{\partial w_i}{\partial y}$, on constate aisément que le minimum de k_i est atteint pour $x = u_i$ et $y = v_i$. La remarque précédente s'applique également. Les expérimentations que nous avons menées confirment que *le critère linéaire tend à rapprocher les épipoles de l'image.*

Nous pouvons remarquer que la conjonction des deux défauts du critère linéaire conduit à une situation est particulièrement défavorable: la géométrie épipolaire obtenue est d'autant moins correcte que les épipoles sont rapprochés, or le critère linéaire rapproche effectivement les épipoles de l'image.

4.2 Critères non-linéaires

4.2.1 La distance aux droites épipolaires

Pour essayer d'introduire une première approche non linéaire, reprenons l'interprétation géométrique du critère (4.3) que nous venons d'introduire. Cette interprétation a fait apparaître comme quantité géométriquement significative la distance d'un point à la droite épipolaire de son correspondant. Cette quantité est en effet *mesurable dans l'image.*

Nous sommes ainsi amenés à envisager en premier lieu un critère de type suivant:

$$\min_{\mathbf{F}} \sum_i d^2(\mathbf{q}'_i, \mathbf{F}\mathbf{q}_i)$$

Il ressemble à celui proposé par Toscani et Faugeras [153] puis Weng, Ahuja et Huang [170] pour le calcul du mouvement. Cependant, contrairement au critère linéaire, un tel critère n'est pas symétrique par rapport aux deux images, puisqu'il ne détermine que les droites épipolaires dans la seconde image et ne doit donc être utilisé que pour obtenir l'épipole dans la seconde image. Pour obtenir l'épipole dans la première image, il conviendrait de permuter les rôles des $\mathbf{q}_i$ et des $\mathbf{q}'_i$. Une difficulté avec cette approche, déjà rencontrée avec les méthodes projectives du chapitre précédent est l'incohérence de la géométrie épipolaire obtenue entre les deux images. Plus précisément, si l'on calcule $\mathbf{F}$ en minimisant $\sum_i d^2(\mathbf{q}'_i, \mathbf{F}\mathbf{q}_i)$ et $\mathbf{F}'$ en minimisant $\sum_i d^2(\mathbf{q}_i, \mathbf{F}'\mathbf{q}'_i)$, il n'y a aucune garantie pour que les points de la droite épipolaire $\mathbf{F}\mathbf{q}$ *autres que* $\mathbf{q}'$ correspondent aux points de la droite épipolaire $\mathbf{F}'\mathbf{q}'$. Cette remarque est illustrée par la figure 4.11, présentée plus bas dans la section expérimentale. Pour obtenir une géométrie épipolaire cohérente, il est nécessaire et suffisant qu'en inversant le rôle des deux images la matrice fondamentale obtenue soit changée en sa transposée. Ceci conduit au critère suivant, qui opère simultanément sur les deux images:

$$\min_{\mathbf{F}} \sum_i (d^2(\mathbf{q}'_i, \mathbf{F}\mathbf{q}_i) + d^2(\mathbf{q}_i, \mathbf{F}^T\mathbf{q}'_i))$$

que nous pouvons finalement réécrire, compte tenu de (4.7) et du fait que $\mathbf{q}_i'^T\mathbf{F}\mathbf{q}_i = \mathbf{q}_i^T\mathbf{F}^T\mathbf{q}'_i$:

$$\min_{\mathbf{F}} \sum_i \left(\frac{1}{(\mathbf{F}\mathbf{q}_i)_1^2 + (\mathbf{F}\mathbf{q}_i)_2^2} + \frac{1}{(\mathbf{F}^T\mathbf{q}'_i)_1^2 + (\mathbf{F}^T\mathbf{q}'_i)_2^2} \right) (\mathbf{q}_i'^T\mathbf{F}\mathbf{q}_i)^2 \tag{4.9}$$

Il est important de noter que ce critère est clairement *normalisé*, au sens où il ne dépend pas du facteur d'échelle choisi pour $\mathbf{F}$.

4.2.2 La pondération par le gradient

Prise en compte de l'incertitude Une autre manière d'améliorer le critère linéaire est de prendre en compte explicitement l'incertitude, par une méthode analogue à celle que nous avions introduite en 3.4 pour améliorer la méthode des birapports. L'idée est que nous avons à minimiser une somme de termes $C_i = \mathbf{q}_i'^T\mathbf{F}\mathbf{q}_i$ qui ont chacun une variance différente. Il est

naturel de les pondérer de sorte que la contribution de chacun de ces termes au critère total soit inversement proportionnelle à leur variance. La variance de C_i est donnée en fonction de celle des points $\mathbf{q}_i$ et $\mathbf{q}'_i$ par:

$$\sigma^2_{C_i} = \begin{bmatrix} \frac{\partial C_i^T}{\partial \mathbf{q}_i} & \frac{\partial C_i^T}{\partial \mathbf{q}'_i} \end{bmatrix} \begin{bmatrix} \Lambda_{\mathbf{q}_i} & \mathbf{0} \\ \mathbf{0} & \Lambda_{\mathbf{q}'_i} \end{bmatrix} \begin{bmatrix} \frac{\partial C_i}{\partial \mathbf{q}_i} \\ \frac{\partial C_i}{\partial \mathbf{q}'_i} \end{bmatrix} \tag{4.10}$$

où $\Lambda_{\mathbf{q}_i}$ et $\Lambda_{\mathbf{q}'_i}$ sont les matrices de covariance des points $\mathbf{q}$ et $\mathbf{q}'$, respectivement. Ces points ne sont pas corrélés puisqu'ils sont mesurés indépendamment dans des images différentes. Nous faisons de plus l'hypothèse classique que la covariance des points est isotrope et uniforme, c'est-à-dire:

$$\Lambda_{\mathbf{q}_i} = \Lambda_{\mathbf{q}'_i} = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix}$$

L'équation (4.10) se simplifie alors en:

$$\sigma^2_{C_i} = \sigma^2 \|\nabla C_i\|^2$$

où ∇C_i désigne le gradient de C_i par rapport au vecteur de dimension quatre (u_i, v_i, u'_i, v'_i) constitué par les coordonnées affines des points $\mathbf{q}_i$ et $\mathbf{q}'_i$. Il vient:

$$\nabla C_i = ((\mathbf{F}^T\mathbf{q}'_i)_1, (\mathbf{F}^T\mathbf{q}'_i)_2, (\mathbf{F}\mathbf{q}_i)_1, (\mathbf{F}\mathbf{q}_i)_2)^T$$

Ce qui conduit donc au critère suivant, qui est également normalisé:

$$\min_{\mathbf{F}} \sum_i \frac{(\mathbf{q}_i'^T\mathbf{F}\mathbf{q}_i)^2}{(\mathbf{F}\mathbf{q}_i)_1^2 + (\mathbf{F}\mathbf{q}_i)_2^2 + (\mathbf{F}^T\mathbf{q}'_i)_1^2 + (\mathbf{F}^T\mathbf{q}'_i)_2^2} \tag{4.11}$$

Nous pouvons noter la grande similarité de ce critère, dont chacun des termes a la forme $\frac{1}{w^2 + w'^2}C$, avec le critère (4.9) obtenu à partir des distances aux droites épipolaires, puisque ce dernier est une somme de termes de la forme $(\frac{1}{w^2} + \frac{1}{w'^2})C$.

Une interprétation en termes de distances Une seconde interprétation de ce critère est possible, en termes de distances. Nous pouvons considérer le problème du calcul de la matrice fondamentale à partir de la définition (4.1) dans un cadre général d'ajustement de surface[2]. La surface $\mathcal{S}$ est modélisée par une équation implicite $g(\mathbf{x}, \mathbf{f}) = 0$, où $\mathbf{f}$ est le vecteur de paramètres recherché décrivant la surface qui s'ajuste le mieux aux données $\mathbf{x}_i$. Le but est donc de minimiser une quantité de la forme $\sum_i d(\mathbf{x}_i, \mathcal{S})^2$. Dans notre cas, les données sont les quadruplets $\mathbf{x}_i = (u_i, v_i, u'_i, v'_i)^T$ obtenus à partir des coordonnées de points appariés, $\mathbf{f}$ est l'un des paramétrages de dimension 7 présentés dans la section précédente, et g est donnée par (4.1). L'utilisation du critère linéaire est en fait une généralisation de la distance

[2]Shape fitting. Un article de synthèse est [146], dont nous reprenons en partie l'exposition.

introduite par Bookstein [14] pour l'ajustement de coniques planes, l'idée évidente étant de d'approximer la distance réelle du point $\mathbf{x}$ à la surface par la valeur $g(\mathbf{x}, \mathbf{f})$ pour obtenir un calcul explicite. Une meilleure approximation a été ensuite introduite par Sampson [135], fondée sur le développement au premier ordre:

$$g(\mathbf{x}) \simeq g(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0) \cdot \nabla g(\mathbf{x}) = g(\mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\| \, \|\nabla g(\mathbf{x})\| \cos(\mathbf{x} - \mathbf{x}_0, \nabla g(\mathbf{x}))$$

Si $\mathbf{x}_0$ est le point de $\mathcal{S}$ le plus proche de $\mathbf{x}$, nous avons les deux propriétés $g(\mathbf{x}_0) = 0$ et $\cos(\mathbf{x} - \mathbf{x}_0, \nabla g(\mathbf{x}_0)) = 1$, qui jointes à l'approximation consistant à considérer que le gradient a la même direction en $\mathbf{x}$ et en $\mathbf{x}_0$: $\cos(\mathbf{x} - \mathbf{x}_0, \nabla g(\mathbf{x}_0)) \simeq \cos(\mathbf{x} - \mathbf{x}_0, \nabla g(\mathbf{x}))$, permettent d'obtenir l'approximation:

$$d(\mathbf{x}, \mathcal{S}) = \|\mathbf{x} - \mathbf{x}_0\| \simeq \frac{g(\mathbf{x})}{\|\nabla g(\mathbf{x})\|}$$

Il est alors clair que le critère (4.11) s'écrit sous la forme: $\sum_i d(\mathbf{x}_i, \mathcal{S})^2$. Il serait possible d'utiliser une approximation de la distance au second ordre, comme l'ont fait Nalwa et Pauchon [122], cependant nos expérimentations avec le critère (4.11) ne nous encouragent pas tellement dans ce sens, et nous avons préféré, quitte à disposer d'un critère théoriquement intéressant, mais d'une utilité réduite, étudier la distance exacte.

4.2.3 La distance globale aux données: le critère euclidien

L'expérience avec l'ajustement de coniques a montré que si les points ne sont pas bien distribués le long de la conique, les résultats obtenus par l'approximation au premier ordre présentée précédemment peuvent différer notablement de ceux obtenus en utilisant une distance euclidienne exacte. C'est le cas dans lequel nous nous trouvons: les points sont les vecteurs $\mathbf{x}_i = (u_i, v_i, u'_i, v'_i)^T$, construits à partir de coordonnées rétiniennes de points mesurés. Par conséquent, les rétines étant finies, les points ne couvrent qu'une partie de la surface $\mathcal{S}$, puisque celle-ci n'est pas bornée (un point à distance finie dans la première rétine $\mathcal{R}_1$ peut avoir un correspondant à l'infini dans la seconde rétine $\mathcal{R}_2$). Une autre manière de voir les choses est de se souvenir que calculer une matrice fondamentale revient à calculer les épipoles, qui sont des points qui peuvent être très éloignés des points mesurés dans les images. Aussi il a paru intéressant d'étudier un critère fondé sur la distance euclidienne *exacte* du couple $(\mathbf{q}_i, \mathbf{q}'_i)$ à la surface algébrique $\mathcal{S}$.

L'hyper-surface $\mathcal{S}$ définie par l'équation (4.1) dans l'espace $\mathcal{R}_1 \times \mathcal{R}_2$ (la rétine cyclopéenne) est quadratique et contient toutes les droites épipolaires. Le calcul de la distance euclidienne 4D d'un point à $\mathcal{S}$ est fondé sur le fait que les droites 4-D définies par $u = u_i, v = v_i, (u', v') \in l'_{u_i, v_i}$ et $u' = u'_i, v' = v'_i, (u, v) \in l_{u'_i, v'_i}$ sont des sous-ensembles de $\mathcal{S}$. Par conséquent, $\mathcal{S}$ est une surface réglée qui peut être paramétrée par chacune de ces deux familles de droites[3].

Choisissons par exemple de paramétrer $\mathcal{S}$ avec la première famille de droites. Chaque point de la surface peut être représenté par $\mathbf{q}_0 = (u_0, v_0)$ et un point de la droite épipolaire

[3]L'hyperboloïde à une nappe est une surface de l'espace 3D possédant la même propriété

$l'_{\mathbf{q}_0}$, aussi la distance d'un point $(\mathbf{q}, \mathbf{q}') = (u, v, u', v')$ à la surface est donnée par le minimum de

$$d^2(\mathbf{q}_0, \mathbf{q}) + d^2(\mathbf{q}', l'_{\mathbf{q}_0})$$

lorsque $\mathbf{q}_0$ décrit R^2. Le calcul de $\mathbf{F}$ s'effectue donc par la minimisation itérée, dont nous remarquons qu'elle conduit également à un critère normalisé:

$$\min_{\mathbf{F}} \sum_i \min_{\mathbf{q}_0} \{d^2(\mathbf{q}_0, \mathbf{q}_i) + d^2(\mathbf{q}'_i, L'_{\mathbf{q}_0})\}$$

Il ne nous a pas été possible d'expérimenter avec ce critère de manière aussi exhaustive qu'avec les autres, car, à cause de la minimisation itérée, les temps de calcul se chiffrent en heures. Nous avons néanmoins constaté sur quelques exemples que, les résultats se rapprochent le plus de ceux obtenus par le critère gradient, en étant légèrement meilleurs.

4.2.4 Méthodes robustes

Dans toutes les méthodes considérées précédemment la minimisation portait seulement sur la matrice fondamentale, ce qui, en un certain sens, revient à considérer que les données ne contiennent pas d'erreur. Les méthodes présentées dans cette section partent au contraire de l'idée que puisque les données sont bruitées, nous pouvons également les ajuster au cours du processus de minimisation. Ces méthodes *robustes* ont la capacité d'éliminer certaines mesures au cours du processus d'optimisation, et de répondre ainsi au problème posé par les mauvais appariements éventuels.

Moindre carrés total Cette première approche est une généralisation du critère linéaire. Le moindres carrés $\min \|\mathbf{Lx} - \mathbf{b}\|_2$ auquel il conduit peut être formulé:

$$\min_{\mathbf{Lx}=\mathbf{b}+\mathbf{r}} \|\mathbf{r}\|_2$$

ceci fait apparaître qu'une hypothèse implicite est que l'erreur est confinée à l'observation $\mathbf{b}$. Lorsque nous mettons également en doute les données $\mathbf{L}$, une généralisation naturelle est le moindres carrés total de Golub et Van Loan [53]:

$$\min_{(\mathbf{L}+\mathbf{E})\mathbf{x}=\mathbf{b}+\mathbf{r}} \|[\mathbf{E}, \mathbf{r}]\|_F$$

où $[\mathbf{E}, \mathbf{r}]$ désigne la matrice obtenue par concaténation de la matrice $\mathbf{E}$ et du vecteur $\mathbf{r}$, la norme considérée étant celle de Frobenius: $\|\mathbf{L}\|_F^2 = \sum \sum L_{ij}^2$

Nous avons trouvé que les épipoles obtenus avec cette méthode ne sont pas du tout fiables. Ils sont nettement moins précis que ceux obtenus par le critère linéaire. Cette méthode, que nous avons mentionnée pour mémoire, ne présente dans notre cas aucun intérêt. Nous avons alors essayé des méthodes itératives, dont un avantage est de pouvoir disposer des critères non-linéaires.

Ajustement des points La première approche a des points communs avec celle de Trivedi [156]: nous cherchons à effectuer une correction optimale sur les points. A chaque itération, dans une première étape, une matrice fondamentale $\mathbf{F}$ est obtenue en minimisant l'un des critères $\sum_i C(\mathbf{F}, \mathbf{q}_i, \mathbf{q}'_i)^2$ présentés précédemment. Dans une seconde étape, les points sont autorisés à varier pour minimiser le critère:

$$\sum_i C(\mathbf{F}, \mathbf{q}_i^{(f)}, \mathbf{q}_i^{'(f)})^2 + \lambda \sum_i \{d^2(\mathbf{q}_i^{(f)}, \mathbf{q}_i^{(i)}) + d^2(\mathbf{q}_i^{'(f)}, \mathbf{q}_i^{'(i)})\}$$

où λ est un coefficient croissant avec la confiance accordée aux mesures. Pour cela, chacun des appariements initiaux $\mathbf{q}_i^{(i)}, \mathbf{q}_i^{'(i)}$ est ajusté en les points finaux $\mathbf{q}_i^{(f)}, \mathbf{q}_i^{'(f)}$ en minimisant le critère:

$$C(\mathbf{F}, \mathbf{q}_i^{(f)}, \mathbf{q}_i^{'(f)})^2 + \lambda\{d^2(\mathbf{q}_i^{(f)}, \mathbf{q}_i^{(i)}) + d^2(\mathbf{q}_i^{(f)}, \mathbf{q}_i^{(i)})\}$$

sous les contraintes $d(\mathbf{q}_i^{(f)}, \mathbf{q}_i^{(i)}) \leq t$ et $d(\mathbf{q}_i^{'(f)}, \mathbf{q}_i^{'(i)}) \leq t$, t étant un seuil de tolérance définissant l'erreur maximale acceptable. La procédure est stoppée lorsque les points de données ne varient plus de manière significative.

Notre procédure avec ajustement des points échoue même dans des cas apparemment très favorables: données non bruités, un seul point très faux[4]. Il n'est pas totalement inintéressant de voir ce qui ce produit: le point très faux est un peu déplacé (quelques pixels), et *tous les autres points* le sont également (fractions de pixel). Ce très faible déplacement suffit à réduire le critère de un à plusieurs ordres de grandeur, sans que $\mathbf{F}$ et la position des épipoles ne soit modifiée de manière significative. Deux itérations sont la plupart du temps suffisantes. Nous pouvons néanmoins retenir de ces expériences que l'introduction d'un seul point très faux peut modifier *entièrement* les résultats, et qu'un déplacement *minime* des autres points permet de trouver une géométrie épipolaire présentant une compatibilité élevée.

Régression robuste Nous avons également implémenté des méthodes de régression robustes classiques. La méthode de "moindre moyenne des carrés" étant plutôt conçue pour éliminer les faux points de données (elle n'est pas très robuste au bruit), nous avons finalement implémenté une méthode de la famille des M-estimateurs[5]. A chaque étape:

- un moindres carrés pondéré $\min \sum_i p_\sigma(r_i) C(\mathbf{F}, \mathbf{q}_i, \mathbf{q}'_i)^2$ est utilisé pour estimer $\mathbf{F}$. Initialement, les coefficients $p_\sigma(r_i)$ sont égaux à 1.

- Une estimation de l'erreur $r_i = |C(\mathbf{F}, \mathbf{q}_i, \mathbf{q}'_i)|$ à l'appariement i est calculée, ainsi que σ, l'écart type de ces erreurs.

[4] outlier

[5] Une référence classique est le livre de Huber [68], voir également la revue de Meer, Mintz et Rosenfeld [110] pour des références récentes.

- Chaque équation est pondérée par un nouveau coefficient $p_\sigma(r)$ décroissant avec l'éloignement de l'appariement i de la zone trouvée par le moindres carrés:

$$p_\sigma(r) = \begin{cases} 1 & r \leq \sigma \\ \frac{\sigma}{r} & \sigma \leq r \leq c\sigma \\ 0 & c\sigma < r \end{cases}$$

où c désigne un seuil (fixé à 3 dans nos expériences).

- La procédure est stoppée lorsque la différence des erreurs normalisées

$$\sqrt{(\sum_i p_\sigma(r_i) r_i^2) / \sum_i p_\sigma(r_i)}$$

devient inférieure à un seuil.

Nous avons trouvé que cette approche a une certaine efficacité, ce que montrent les résultats détaillés dans la section expérimentale.

4.3 Paramétrages de la matrice fondamentale

4.3.1 Une matrice définie à un facteur d'échelle près

L'idée la plus naturelle pour paramétrer $\mathbf{F}$ est d'utiliser ses neuf coefficients. $\mathbf{F}$ n'étant définie qu'à un facteur d'échelle près, il convient de fixer un de ces coefficients à 1, puisque seul le critère linéaire permet d'exprimer simplement une normalisation d'un autre type. Nous obtenons donc un paramétrage de $\mathbf{F}$ par huit valeurs, qui sont le quotient des huit autres coefficients par le coefficient normalisateur.

En pratique, le choix du coefficient à normaliser a une certaine importance vis-à-vis de la stabilité numérique des résultats dans le cas des critères non-linéaires. Ces critères, présentés section 4.2 ont en effet la forme:

$$\min_{\mathbf{F}} \sum_i \frac{Q_i(F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})}{R_i(F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23})}$$

où Q_i et R_i sont des formes quadratiques prenant des valeurs nulles à l'origine. Il est connu qu'alors, au voisinage de l'origine la fonction Q_i/R_i ne possède pas de propriétés de régularité, ce qui aurait pour conséquence d'induire des instabilités numériques lors du processus de minimisation, lequel utilise les dérivées. Cette première remarque conduit donc à imposer des normalisations à 1 d'un des six premiers coefficients de $\mathbf{F}$, qui sont ceux apparaissant dans l'expression de R_i. Nous avons expérimentalement constaté que ces choix ne sont pas équivalents, si les ordres de grandeur des éléments de la matrice fondamentale $\mathbf{F}$ sont très différents. Dans ce cas, les meilleurs résultats sont obtenus lorsque le coefficient normalisé à 1 est un de ceux ayant finalement la plus grande valeur absolue, la normalisation à 1 des coefficients de plus petite valeur absolue conduisant généralement à de mauvais résultats.

Avec un paramétrage de ce type, il n'y a pas de garantie que la matrice $\mathbf{F}$ obtenue soit singulière, si bien que la détermination de la transformation épipolaire ne peut pas s'effectuer à l'aide des formules de décomposition (2.14), mais doit utiliser la méthode approchée (2.15). Le fait que la matrice $\mathbf{F}$ ne soit pas de rang deux, mais trois signifie que les différentes droites épipolaires trouvées ne forment pas un faisceau, c'est à dire qu'elles n'ont pas de manière exacte de point commun. Comme nous l'avons remarqué précédemment, ceci est la source d'incohérences dans la géométrie épipolaire. Un exemple réel est donné plus bas, 4.4.2.

4.3.2 Une matrice singulière

L'inconvénient de l'approche précédente est de ne pas tenir compte du fait que $\mathbf{F}$, en tant que matrice fondamentale, est de rang deux et ne dépend que de sept paramètres. Une première approche, est d'effectuer une minimisation sous la contrainte $\det(\mathbf{F}) = 0$, qui est polynomiale du troisième degré en les coefficients de $\mathbf{F}$. Nous avons d'abord implémenté une telle méthode, cependant les programmes de minimisation numérique sous contrainte que nous avons employés se sont révélés peu efficaces.

Une remarque nous permet d'obtenir le même effet en conservant une minimisation sans contrainte: l'idée est d'écrire la matrice $\mathbf{F}$ sous la forme:

$$\mathbf{F} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7a_1 + a_8a_4 & a_7a_2 + a_8a_5 & a_7a_3 + a_8a_6 \end{pmatrix} \tag{4.12}$$

Le fait que la troisième ligne soit une combinaison linéaire des deux premières assure le fait que $\mathbf{F}$ soit singulière. Cette représentation d'une matrice de rang deux est presque générale, le seul cas où elle n'est pas valide étant celui où les deux premiers vecteurs-ligne sont liés. Son avantage est de paramétrer la matrice fondamentale par le nombre exact de variables indépendantes, qui est de sept, une fois que le facteur d'échelle a été pris en compte comme précédemment. En revanche, la symétrie des paramètres n'est pas maintenue, et il est nécessaire de faire appel à des techniques de minimisation non-linéaire. Ce dernier point n'est pas très important, car de toute manière les critères les plus fiables sont déjà non linéaires, ce que nous pressentons depuis la section précédente, et que la section suivante confirmera.

Voyons sur les deux cas représentatifs ce que sont les sept paramètres de cette représentation, compte tenu du fait que nous devons également tenir compte d'une normalisation.

- normalisation d'un des six premiers éléments, par exemple F_{11}:

$$\mathbf{F} = \begin{pmatrix} 1 & x_1 & x_2 \\ x_3 & x_4 & x_5 \\ x_6 + x_7x_3 & x_6x_1 + x_7x_4 & x_6x_2 + x_7x_5 \end{pmatrix}$$

- normalisation d'un des trois derniers éléments, par exemple F_{31}:

$$\mathbf{F} = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7x_1 & 1 & x_7x_3 + (1 - x_7x_2)/x_5 \end{pmatrix}$$

Il faut noter que les matrices servant à initialiser la recherche ne vérifiant généralement pas la propriété sur le rang ne peuvent pas être représentées de cette manière. Nous devons donc commencer par déterminer une matrice de rang 2 qui en soit proche, puis calculer les sept paramètres à partir de cette matrice.

4.3.3 Une matrice fondamentale aux épipoles finis

La représentation précédente ne tient compte que du fait que la matrice recherchée soit singulière. Il est également possible de tenir compte de sa structure de matrice fondamentale pour obtenir un paramétrage par les éléments de la transformation épipolaire, qui sont les quantités qui nous intéressent finalement. Nous nous plaçons dans le cas pratique où les épipoles sont à distance finie, si bien que les formules (2.13) peuvent s'appliquer, d'où:

$$\mathbf{F} = \begin{pmatrix} b & a & -ay - bx \\ -d & -c & cy + dx \\ dy' - bx' & cy' - ax' & -cyy' - dy'x + ayx' + bxx' \end{pmatrix} \tag{4.13}$$

Les paramètres utilisés sont dans ce cas les coordonnées (x, y) et (x', y') des deux épipoles, et trois des quatre coefficients de l'homographie. Ceux-ci sont ceux de la sous-matrice 2×2 obtenue en supprimant la troisième ligne et la troisième colonne, et ne sont définis qu'à une constante multiplicative près. Les trois paramètres sont obtenus en choisissant le coefficient de plus grande valeur absolue, puis en divisant les trois coefficients restants par celui-ci, ce qui revient à le fixer à 1. L'approximation initiale est effectuée en recherchant les épipoles aux moindres carrés selon (2.15).

4.4 Une comparaison expérimentale des méthodes

4.4.1 Approche statistique: données synthétiques

Nous avons présenté une approche du calcul de la matrice fondamentale par minimisation. Nous avons vu que plusieurs paramétrages pouvaient être utilisés, et que plusieurs critères étaient envisageables. L'objet de cette section est d'effectuer une comparaison statistique des résultats obtenus avec les différentes combinaisons possibles.

Méthode Nous partons d'un nuage de points répartis aléatoirement dans l'espace dans un cube de dimensions réalistes, et d'une matrice de projection perspective réaliste $\mathbf{P}$. Tous les exemples de cette section ont été obtenus avec:

$$\mathbf{P} = \begin{bmatrix} 0.3510195804 & 0.008150611256 & -0.1542164157 & 108.8789035 \\ 0.08887767453 & 0.5295757929 & 0.1027246520 & 205.5086623 \\ 0.0003935363917 & 0.000007856647581 & 0.0003971828308 & 1 \end{bmatrix}$$

Chaque essai consiste à:

- choisir aléatoirement un déplacement rigide $\mathbf{D}$, et calculer la matrice fondamentale $\mathbf{F}$ associée à $\mathbf{D}$ et $\mathbf{P}$,
- calculer la matrice de projection $\mathbf{P}'$ obtenue par application de $\mathbf{D}$ à $\mathbf{P}$,
- prendre l'image des points 3D par chacune des caméras en projetant à l'aide de $\mathbf{P}$ et $\mathbf{P}'$ puis en éliminant les points sortant des rétines,
- bruiter les points 2D obtenus,
- effectuer le calcul de la matrice fondamentale,
- comparer les coordonnées des épipoles trouvés avec les valeurs exactes.

La mesure de distance relative que nous utilisons consiste à considérer, pour chacune des deux coordonnées de chacun des deux épipoles, l'erreur relative:

$$\min\{\frac{|x - x_0|}{\min(|x|, |x_0|)}, 1\}$$

Comme le montre par exemple la table 4.4.1, les erreurs relatives sur les quatre coordonnées citées varient de manière cohérente, ce qui permet d'exprimer la distance relative sous la forme d'un seul nombre, obtenu en prenant la moyenne des quatre erreurs relatives. Nous avons également calculé systématiquement la covariance, mais nous ne la mentionnerons pas si elle n'apporte pas d'information particulière. Il est très important de choisir une procédure de test consistant à varier les déplacements et non pas seulement le bruit (en travaillant à mouvement fixé, et en évaluant alors moyenne et covariance) car le comportement des différentes méthodes d'estimation est très dépendant du mouvement choisi. Nous étudierons cette dépendance ultérieurement. Dans ce chapitre, nous voulons nous centrer uniquement sur la comparaison de méthodes de calcul, d'où le choix de notre procédure expérimentale.

Les implémentations du critère linéaire Nous avons d'abord comparé les différentes implémentations du critère linéaire qui ont été présentées plus haut. La table 4.4.1 présente les moyennes des différences relatives obtenues à trois niveaux de bruit différents. Chacun des 300 essais est effectué avec 50 points. Les conclusions que l'on peut tirer de ces essais sont que:

- La normalisation des coordonnées projectives change beaucoup les résultats, en les dégradant.
- Les deux méthodes de calcul **DIAG** et **M-C** donnent des résultats très proches, qui semblent les plus précis.
- La différence des trois premiers critères est faible, en regard des erreurs commises.

La dernière remarque est conforme au fait qu'en théorie le minimum soit unique. En conséquence de ces essais, nous utilisons comme critère linéaire **DIAG**.

bruit	critère	distances relatives				variances			
		e_x	e_y	e'_x	e'_y	$\sigma(e_x)$	$\sigma(e_y)$	$\sigma(e'_x)$	$\sigma(e'_y)$
0.2 pixel	**LIN**	0.1913	0.1824	0.1955	0.1671	0.0994	0.0915	0.1055	0.0871
	DIST-L	0.1115	0.0875	0.0689	0.0638	0.0649	0.0408	0.0287	0.0206
		0.1133	0.0893	0.0788	0.0738	0.0649	0.0409	0.0372	0.0292
		0.0018	0.0018	0.0100	0.0100	0.0003	0.0003	0.0099	0.0099
	DIST-T	0.0670	0.0423	0.0463	0.0417	0.0323	0.0054	0.0137	0.0071
		0.1299	0.1056	0.0864	0.0827	0.0804	0.0566	0.0410	0.0347
		0.0634	0.0640	0.0406	0.0419	0.0563	0.0563	0.0303	0.0305
	GRAD-L	0.1114	0.0870	0.0687	0.0634	0.0649	0.0408	0.0286	0.0205
		0.1131	0.0891	0.0788	0.0736	0.0648	0.0409	0.0372	0.0292
		0.0017	0.0020	0.0102	0.0103	0.0002	0.0003	0.0099	0.0099
	GRAD-T	0.0672	0.0425	0.0465	0.0416	0.0323	0.0054	0.0137	0.0070
		0.1201	0.0895	0.0847	0.0852	0.0723	0.0406	0.0417	0.0383
		0.0532	0.0469	0.0393	0.0443	0.0461	0.0379	0.0314	0.0345
1.0 pixel	**LIN**	0.4816	0.4827	0.4444	0.4585	0.1765	0.1764	0.1657	0.1755
	DIST-L	0.2662	0.2554	0.2135	0.2086	0.1213	0.1170	0.0918	0.0856
		0.3117	0.2928	0.2753	0.2467	0.1493	0.1414	0.1335	0.1089
		0.1040	0.1034	0.1042	0.0962	0.0923	0.0924	0.0923	0.0843
	DIST-T	0.1470	0.1359	0.1322	0.1188	0.0560	0.0492	0.0508	0.0387
		0.3387	0.3155	0.3162	0.2894	0.1638	0.1525	0.1537	0.1347
		0.2366	0.2320	0.2342	0.2245	0.1672	0.1632	0.1626	0.1549
	GRAD-L	0.2517	0.2453	0.1953	0.1978	0.1136	0.1108	0.0795	0.0795
		0.3114	0.3040	0.2689	0.2561	0.1421	0.1461	0.1236	0.1139
		0.1028	0.1101	0.1154	0.1170	0.0824	0.0905	0.0981	0.1005
	GRAD-T	0.1464	0.1337	0.1317	0.1161	0.0556	0.0490	0.0505	0.0378
		0.3533	0.3217	0.3076	0.2869	0.1700	0.1583	0.1512	0.1341
		0.2536	0.2427	0.2197	0.2186	0.1797	0.1743	0.1564	0.1521

Tableau 4.3: Comparaison des quatre coordonnées des épipoles

Les critères non-linéaires de base Nous avons présenté deux paramétrages section 4.3. Des trois critères non-linéaires présentés section 4.2, nous n'avons retenu que les deux premiers, l'étude systématique du critère de distance euclidienne exacte n'étant pas possible en raison du temps de calcul. Cela conduit à quatre combinaisons, dont les sigles sont présentés dans la table 4.4.1.

Un problème commun aux critères non-linéaires est celui de l'initialisation. Comme le critère linéaire donne de manière explicite une solution, il est commode de l'utiliser à cet effet. La matrice obtenue n'est, en général, pas singulière, et ne peut donc être représentée avec sept paramètres seulement. Lors du changement de paramétrage, nous calculons donc l'une des représentations à sept paramètres qui permet d'obtenir une matrice fondamentale la plus proche possible de la matrice initiale, sachant que celles-ci vont nécessairement différer.

La comparaison que nous effectuons porte sur trois points:

1. *La stabilité du minimum correspondant à la solution correcte.* En présence de bruit, la surface représentant la valeur du critère en fonction des paramètres se déforme, si bien

bruit		distances relatives			
		SVD	**DIAG**	**M-C**	**DIAG-N**
0.2 pixel	**EXACT**	0.1598	0.1543	0.1550	0.1828
	SVD		0.0544	0.0543	0.1511
	DIAG			0.0034	0.1375
	M-C				0.1380
1 pixel	**EXACT**	0.4623	0.4590	0.4590	0.4929
	SVD		0.1077	0.1089	0.3535
	DIAG			0.0040	0.3486
	M-C				0.3488
1.8 pixel	**EXACT**	0.6375	0.6334	0.6336	0.6587
	SVD		0.1442	0.1444	0.4575
	DIAG			0.0055	0.4465
	M-C				0.4465

Tableau 4.4: Comparaison des critères linéaires

sigle	critère	paramétrage
LIN	linéaire	normalisation par $\|\mathbf{F}\|$
DIST-L	distance aux épipolaires (4.9)	linéaire singulière (4.12)
DIST-T	distance aux épipolaires	transformation épipolaire (4.13)
GRAD-L	pondération par gradient (4.11)	linéaire singulière
GRAD-T	pondération par gradient	transformation épipolaire

Tableau 4.5: Critères utilisées pour le calcul de la matrice fondamentale

que les coordonnées du minimum changent également. Une mesure de cette variation est constituée par les distances relatives entre les épipoles corrects, et ceux obtenus par minimisation en utilisant comme point de départ les épipoles corrects.

2. *Les propriétés de convergence.* La question est ici de voir dans quelle mesure il est possible d'obtenir un résultat correct en partant d'une initialisation plausible, qui est la matrice obtenue par le critère linéaire. L'écart est mesuré par la distance relative des épipoles ainsi obtenus aux épipoles corrects.

3. *La stabilité du critère.* Lors de la déformation de la surface représentant la valeur du critère en fonction des paramètres, des minima locaux ne correspondant pas à la solution exacte peuvent prendre des valeurs plus faibles que le minimum correspondant à celle-ci, ce qui empêcherait alors toute méthode de la retrouver. Une telle situation se produit lorsque la valeur du critère obtenue par le test 1 est supérieure à celle obtenue par le test 2.

Aussi, lors de chaque essai consiste à:

1. effectuer le calcul de la matrice fondamentale en initialisant avec **F**
2. effectuer le calcul de la matrice fondamentale en initialisant avec le critère linéaire

Une petite difficulté a été le choix d'un algorithme de minimisation. Nous avons fini par utiliser la procédure E04FCF de la librairie Numerical Algorithm Group [1], qui effectue une minimisation aux moindres carrés non-linéaire par une variante de la méthode de Newton. Les procédures de cette même librairie exigeant l'expression des dérivées se sont avérées délicates à employer. Nous avons également utilisé la procédure de minimisation "Powell" présentée dans Numerical Recipes [129] , cependant celle-ci s'est révélée un peu moins efficace que la première procédure mentionnée, comme le montre la table 4.6.

bruit	1: initialisé exactement		2: initialisé par linéaire		différence 1 - 2	
	NAG	NR	NAG	NR	NAG	NR
0.2	0.0495	0.0499	0.0949	0.1060	0.0459	0.0586
0.4	0.0795	0.0841	0.1642	0.1783	0.0921	0.1170
0.6	0.1054	0.1173	0.2140	0.2553	0.1163	0.2069
0.8	0.1251	0.1511	0.2552	0.3001	0.1646	0.2256
1.0	0.1320	0.1793	0.3174	0.3345	0.2337	0.2591
1.2	0.1635	0.2139	0.3727	0.3936	0.2747	0.3097
1.4	0.1852	0.2188	0.4104	0.3964	0.3029	0.3122
1.6	0.1945	0.2596	0.4286	0.4434	0.3125	0.3488
1.8	0.2140	0.2495	0.4166	0.4515	0.2971	0.3558
2.0	0.2215	0.2725	0.4648	0.4994	0.3412	0.3994
2.2	0.2342	0.2916	0.4653	0.4973	0.3439	0.3796
2.4	0.2545	0.3353	0.5144	0.5332	0.4016	0.4383
2.6	0.2628	0.3333	0.5396	0.5643	0.4316	0.4549

Tableau 4.6: Comparaison des résultats de deux algorithmes de minimisation

Toutes les statistiques présentées ont été établies en effectuant 100 essais, et en sélectionnant 50 appariements au maximum. La figure 4.3 présente les distances relatives obtenues en initialisant les minimisations avec les valeurs exactes. La figure 4.4 présente les distances relatives obtenues en les initialisant avec le critère linéaire. La distance relative des épipoles obtenus par les deux calculs précédents est représentée sur la figure 4.5. La figure 4.6 comptabilise le nombre de cas où la valeur du minimum obtenue avec la première initialisation se révèle moins bonne que celle obtenue avec la seconde. Les conclusions que nous tirons de cette comparaison sont:

- Les critères non-linéaires sont toujours plus performants que le critère linéaire. En partant d'une solution donnée par le critère linéaire et en appliquant un critère non-linéaire, nous améliorons toujours les résultats.

- La différence de comportement due au choix du critère (**DIST** ou **GRAD**) est peu significative, par rapport à celle due au choix du paramétrage (**L** ou **T**).
- Le paramétrage **T** conduit à des minima plus stables, mais ses propriétés globales (convergence et stabilité du critère) sont moins bonnes.
- En conséquence, en partant du critère linéaire, chacune des quatre combinaisons donne des résultats à peu près comparables, la méthode **DIST-L** étant légèrement supérieure.
- Le calcul reste sensible: un bruit gaussien de variance 1 conduit à une erreur relative qui est de 30%.

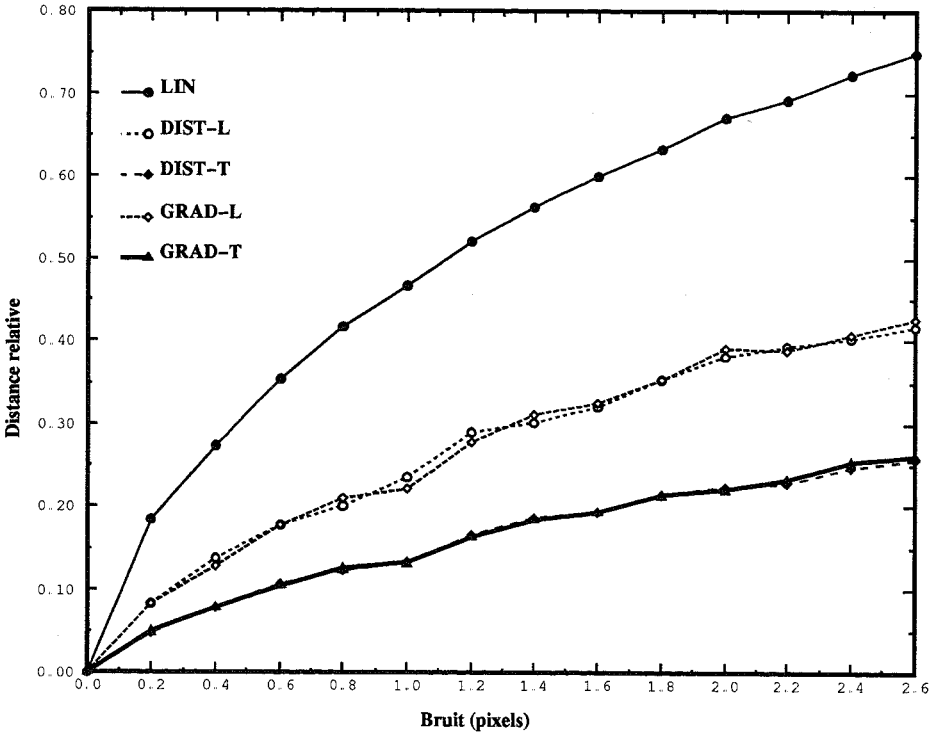


Fig. 4.3: Distances relatives obtenues en initialisant avec les valeurs exactes

La méthode robuste Nous avons ensuite évalué l'apport de la méthode itérative du M-estimateur, en lançant un calcul fondé sur le critère **DIST-L** avec initialisation par le critère linéaire. Les essais statistiques ont été menés en utilisant 200 mouvements, les *outliers* étant à chaque fois introduites différemment, avec une variance de 50 pixels. Les résultats sont présentés table 4.7. Chaque case de cette table contient, outre les moyennes des distances relatives (première ligne), le pourcentage d'essais pour lequel les épipoles ont été trouvés avec moins de 5% d'erreur (premier rapport), ainsi que les pourcentages d'essais pour lesquels l'une ou l'autre des méthodes s'est avérée significativement plus précise (second rapport).

Les principales conclusions sont:

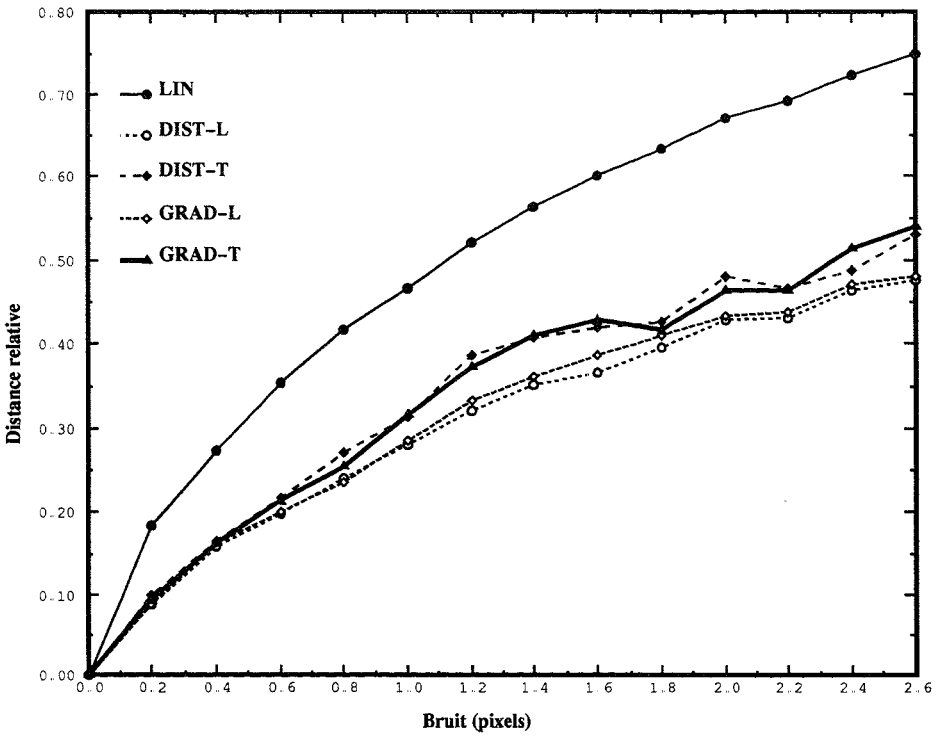


Fig. 4.4: Distances relatives obtenues en initialisant avec le critère linéaire

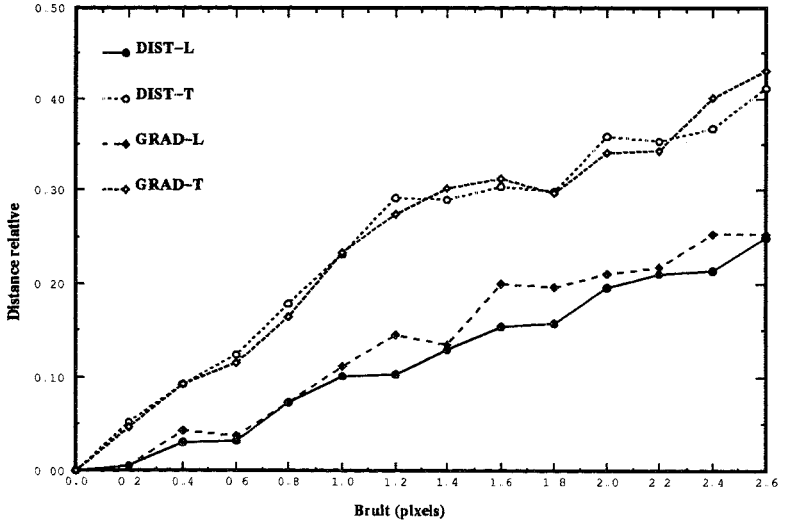


Fig. 4.5: Distances relatives entre les résultats obtenus avec les deux initialisations

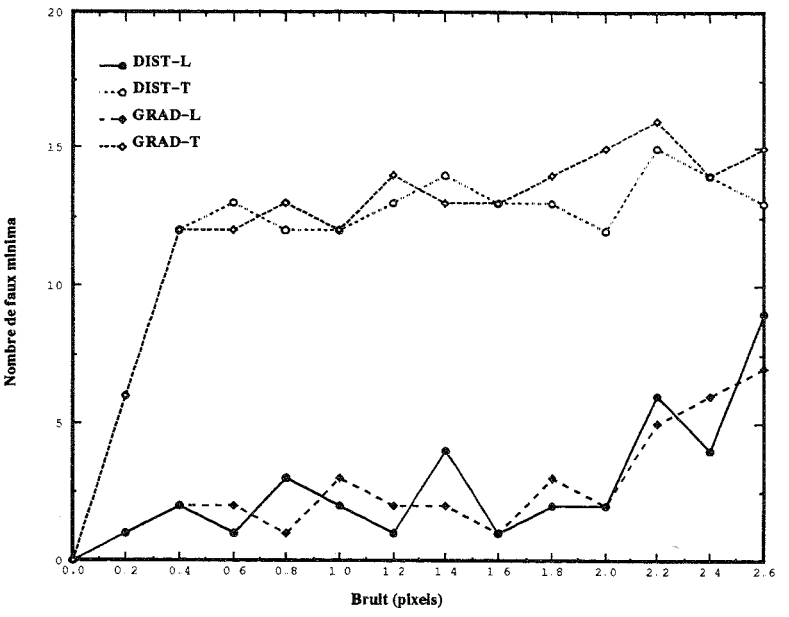


Fig. 4.6: Nombre de faux minima

- En l'absence d'outliers, la méthode itérative dégrade les résultats, à tous les niveaux de bruit.
- En présence d'un pourcentage d'outliers assez modeste, la méthode itérative améliore notablement les résultats *en moyenne.*
- Lorsque ce pourcentage devient assez important ($> 10\%$), la méthode itérative ne suffit plus, et les résultats sont de toute manière très mauvais.

Dans les cas réels, les faux appariements ne sont pas si loin, et la méthode se révèle alors relativement efficace. Un exemple est donné dans le dernier chapitre, figure 9.13.

bruit	pourcentage d'outliers							
(pixels)	0		2		5		10	
0	0.0000	0.0390	0.5172	0.3623	0.7107	0.5417	0.7893	0.7116
	100/96	4/0	18/72	10/46	8/37	8/39	5/17	12/27
0.2	0.0633	0.1565	0.5707	0.4245	0.6956	0.5560	0.7831	0.6810
	83/73	13/2	17/49	9/42	9/33	9/37	5/15	11/27
0.6	0.1532	0.1986	0.5875	0.4814	0.7130	0.5759	0.7899	0.6869
	64/59	11/5	15/40	12/36	8/28	7/36	5/15	9/27
1.0	0.2034	0.2849	0.5964	0.5104	0.7092	0.5765	0.7844	0.6848
	53/45	19/6	14/36	12/34	8/24	9/36	5/13	10/29

Tableau 4.7: L'apport de la procédure itérative du M-estimateur.

4.4.2 Quelques exemples réels

Nous présentons dans cette section quelques exemples destinés à valider nos analyses et à illustrer les difficultés rencontrées en pratique.

La figure 4.7 permet de voir que les faisceaux d'épipolaires obtenus avec le critère linéaire et ceux obtenus avec un critère non-linéaire présentent de grandes différences. Nous pouvons observer que si nous considérons un des points qui a servi a effectuer le calcul, sa droite épipolaire passe bien à proximité de son correspondant dans les deux cas. Cependant un agrandissement comme celui présenté figure 4.8 montre qu'avec le critère non-linéaire, nous obtenons des distances de points à l'épipolaire du correspondant qui sont meilleures qu'avec le critère linéaire. Ceci s'explique aisément par le fait que le critère non-linéaire minimise précisément la somme de ces distances. Le défaut le plus grave du critère linéaire est illustré par la figure 4.9, qui montre que l'intersection des droites épipolaires obtenues à partir de ce critère ne se réduit pas à un point, comme nous l'avions signalé plus haut. En conséquence, nous remarquons que si nous considérons un point autre que ceux ayant servi pour le calcul, sa droite épipolaire peut passer loin du point correspondant, l'écart étant d'autant plus marqué que nous nous rapprochons de l'épipole. Ce phénomène ne se produit bien sûr pas avec le critère non-linéaire, puisque la matrice fondamentale obtenue est exactement de rang deux, ce qui équivaut au fait que l'intersection des droites épipolaires se réduit à un point. La figure 4.10 montre que, si dans le haut de l'image la cohérence est à peu près respectée, dans le bas de l'image il est possible en revanche d'avoir avec le critère linéaire des écarts très importants. Par contre, la géométrie épipolaire obtenue avec le critère non-linéaire ne présente pas ce défaut.

La figure 4.11 montre des épipolaires obtenues à partir du critère non linéaire de distance aux droites épipolaires en effectuant un calcul indépendant dans chaque image. Nous remarquons que les problèmes observés précédemment ne se produisent plus: quel que soient les points de la première image, les droites épipolaires associées de la seconde image passent bien par leur correspondants, et le fait que la matrice fondamentale utilisée soit bien de rang deux garantit que les épipoles forment un faisceau. Il est cependant évident sur ces figures que les seuls points de correspondance des droites épipolaires sont ceux qui ont servi à les définir, alors que tous les points devraient se correspondre mutuellement. Dans d'autres exemples, l'incohérence est toutefois moins visible, tout en restant détectable par la simple application de la contrainte d'ordre épipolaire.

Une deuxième illustration est fournie par la figure 4.12. L'image est celle prise par la caméra centrale d'un système de stéreovision trinoculaire dont les caméras sont disposés le long de deux axes à peu près perpendiculaires pour obtenir des images à peu près rectifées. Ce cas n'est pas du tout favorable pour l'estimation de la matrice fondamentale, néanmoins nous pouvons voir sur la partie droite que des faisceaux à peu près corrects sont obtenus par le critère non-linéaire, ce qui n'est pas du tout le cas du critère linéaire, qui raproche une fois de plus les épipoles. Pour ces images d'extérieur, les points d'intérêt ont été extraits par une méthode de corrélation.

La figure 4.13 permet enfin de voir qu'il est possible d'obtenir une géométrie épipolaire qui

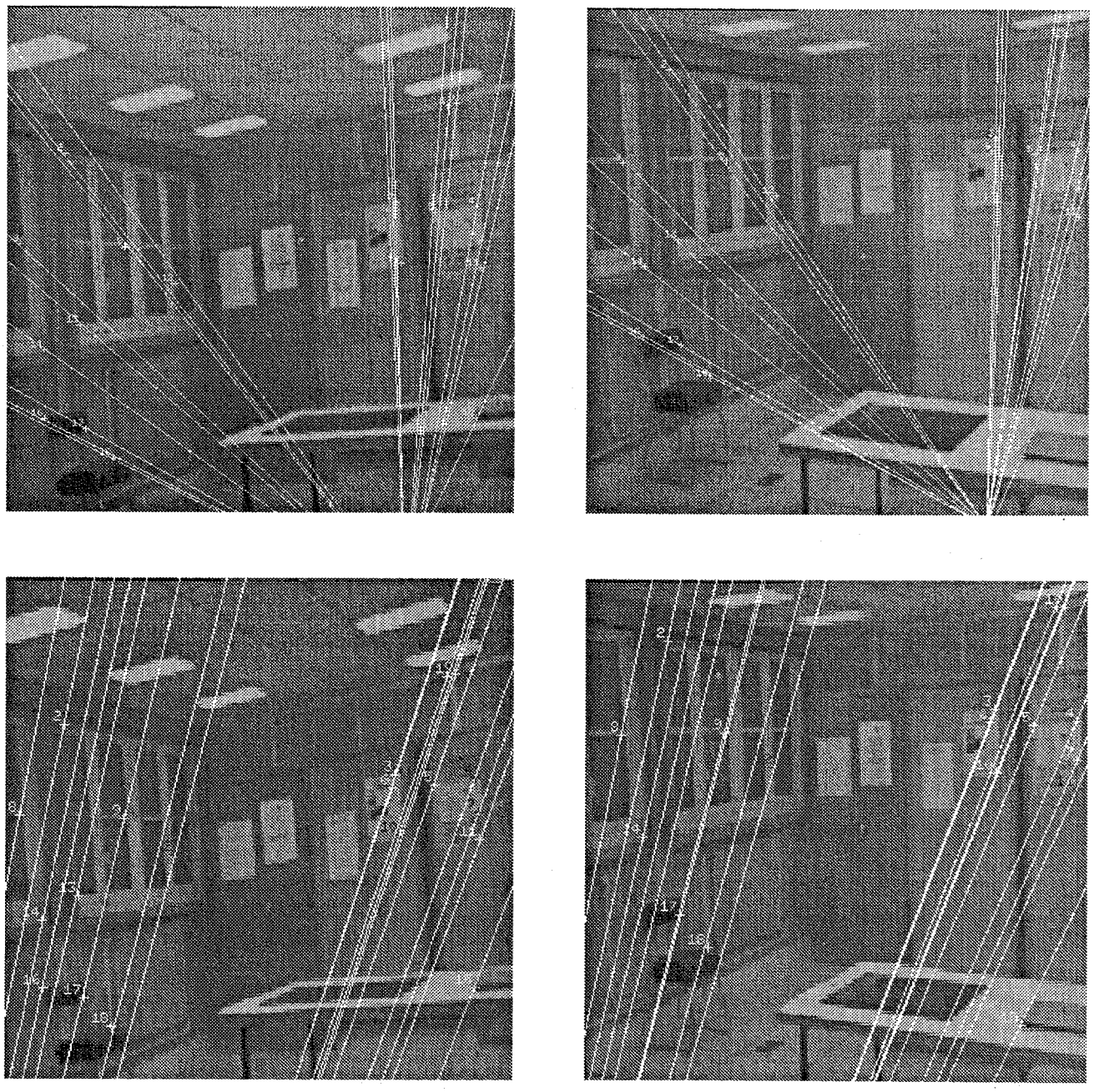


Fig. 4.7: Epipolaires obtenues avec le critère linéaire (haut), et avec le critère non-linéaire (bas)

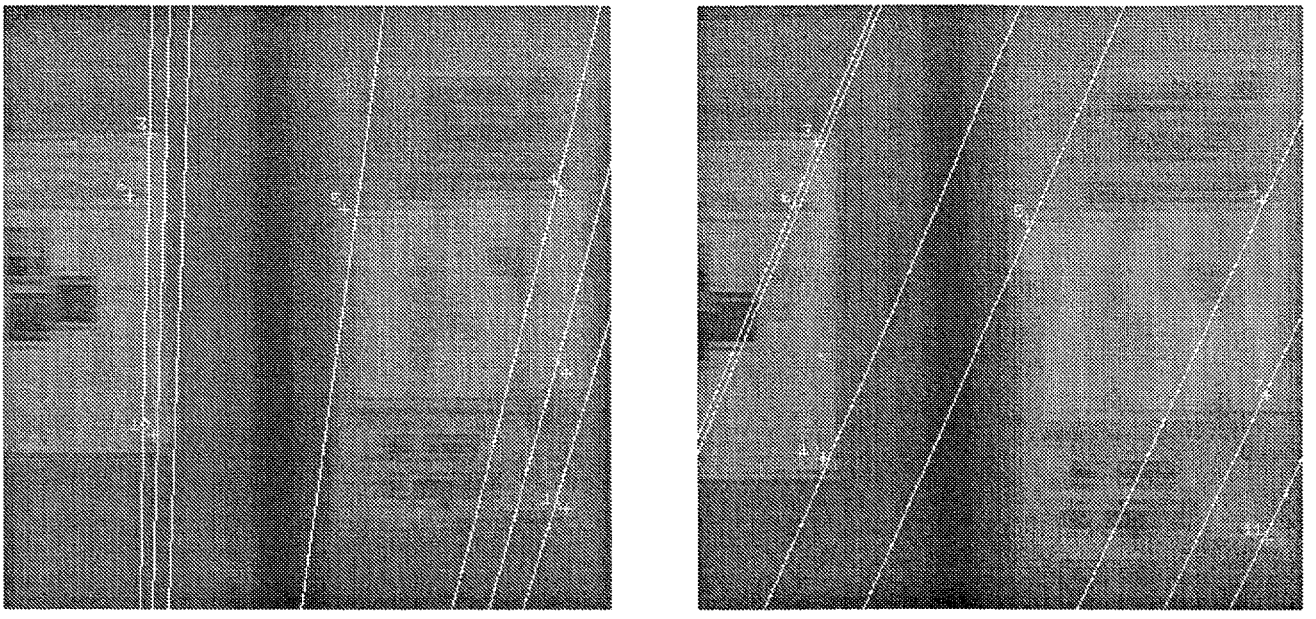


Fig. 4.8: Détail montrant l'ajustement avec le critère linéaire (gauche) et le critère non-linéaire (droite)

semble parfaitement correcte (en particulier les écarts aux points de référence sont largement inférieurs au pixel, la moyenne étant 0.192 pixel) alors que la position des épipoles n'est pas du tout correcte: la position estimée est

$$\mathbf{e} = (421.27, 854.52)^T \quad \mathbf{e}' = (600.39, 798.02)^T$$

alors que la position obtenue par calcul à partir des matrices de projection résultant de la calibration sur les mires présentes dans les images est:

$$\mathbf{e} = (1529.551201.08)^T \quad \mathbf{e}' = (1964.51, 1475.68)^T$$

Nous pouvons remarquer que les écarts aux épipolaires obtenus avec la géométrie épipolaire correcte sont légèrement supérieurs (0.206 pixel). Dans la partie supérieure de ces figures, les épipolaires sont calculés à l'aide de la matrice fondamentale trouvée par le critère de distance aux droites épipolaires opérant dans les deux images. La partie inférieure de ces figures représente les droites épipolaires correctes. Ce phénomène, l'existence de plusieurs géométries épipolaires plausibles, a été rencontré très fréquemment lors de nos expérimentations, et est une preuve de la difficulté du problème du calcul de la matrice fondamentale. Un second exemple est visible dans le chapitre suivant, figure 5.3. *Si obtenir une géométrie épipolaire qui semble satisfaisante à l'intérieur des images est assez simple, trouver la vraie position des épipoles l'est beaucoup moins.*

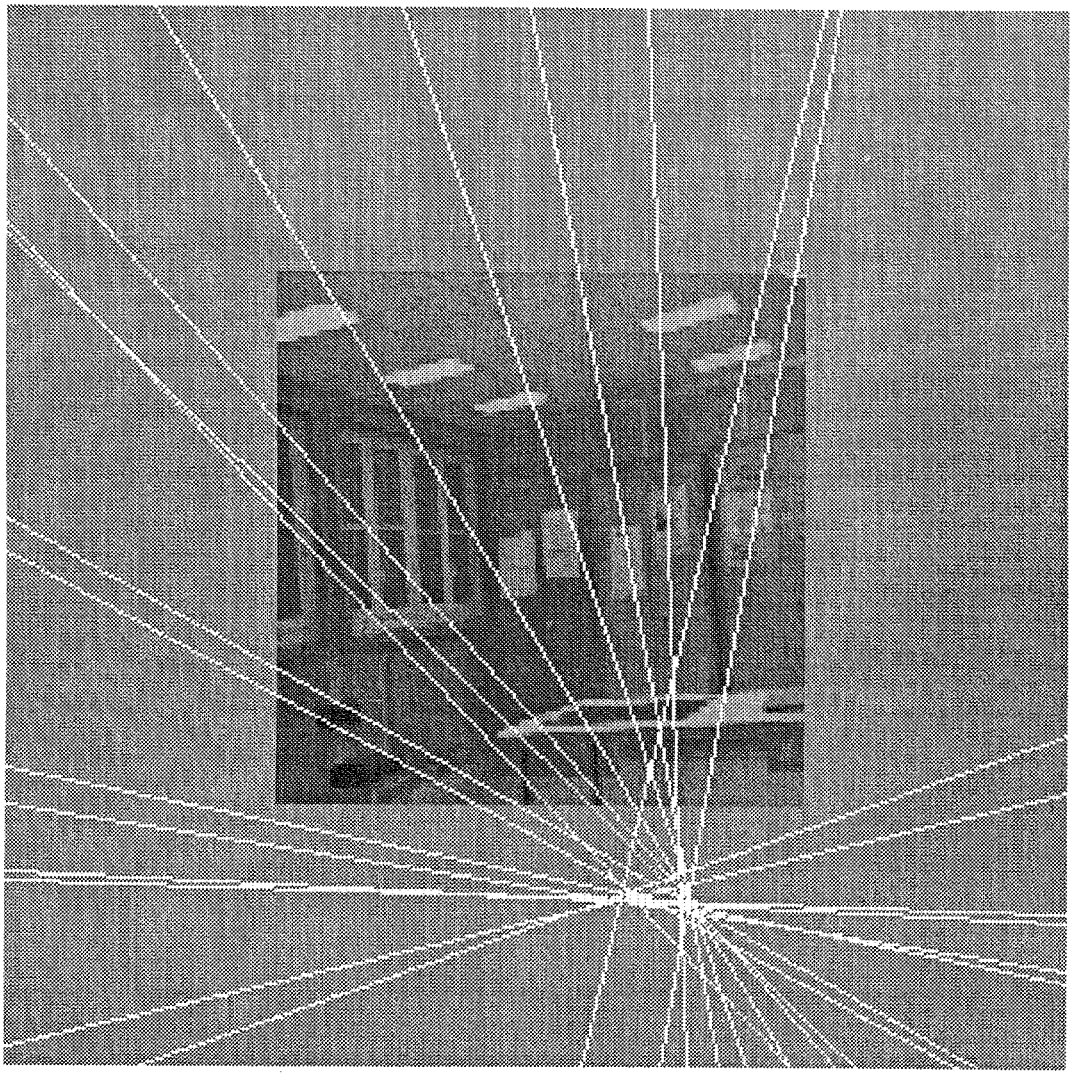


Fig. 4.9: Intersection des droites épipolaires obtenues avec le critère linéaire

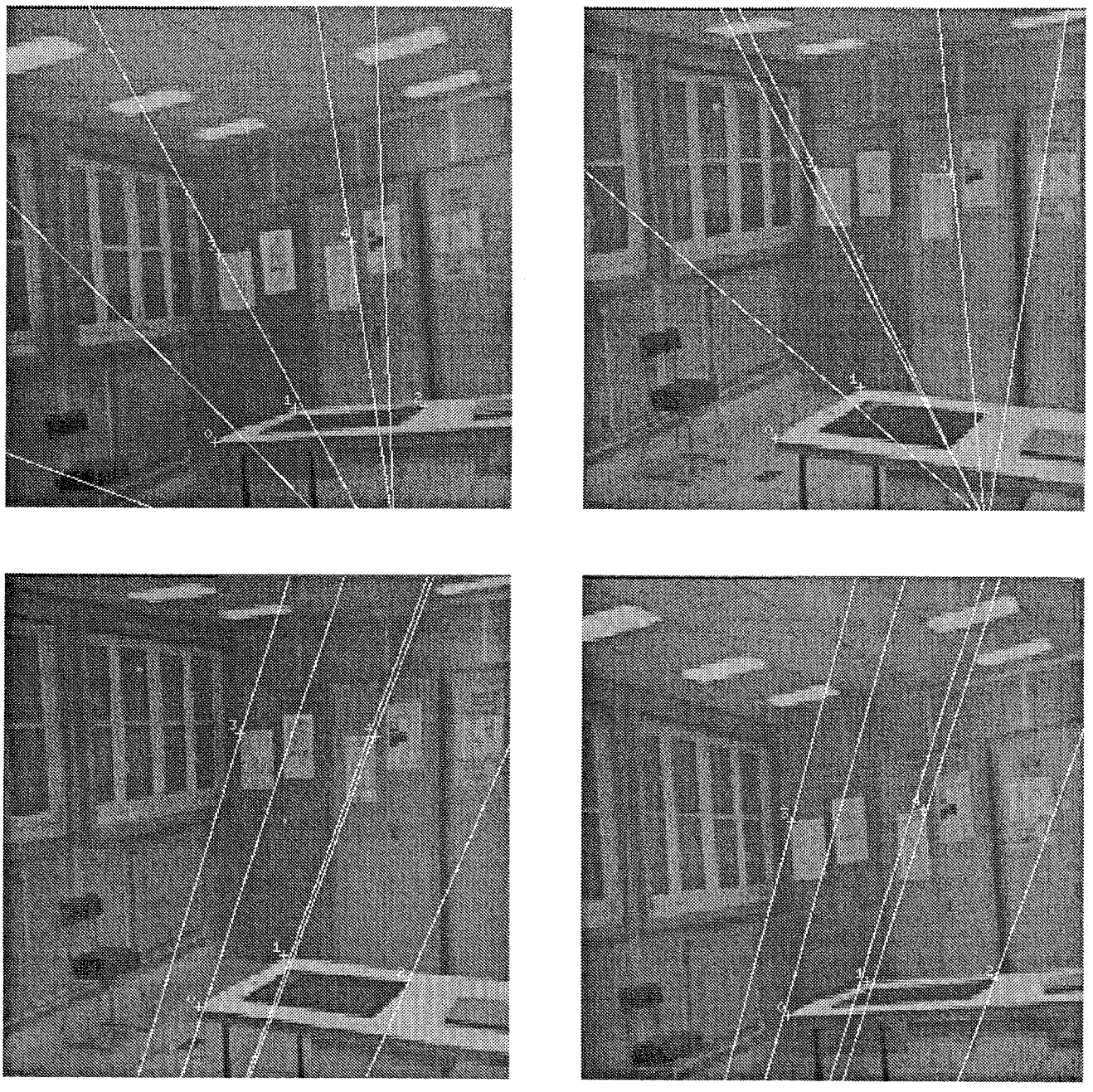


Fig. 4.10: Epipolaires supplémentaires obtenues avec le critère linéaire (haut), et avec le critère non-linéaire (bas)

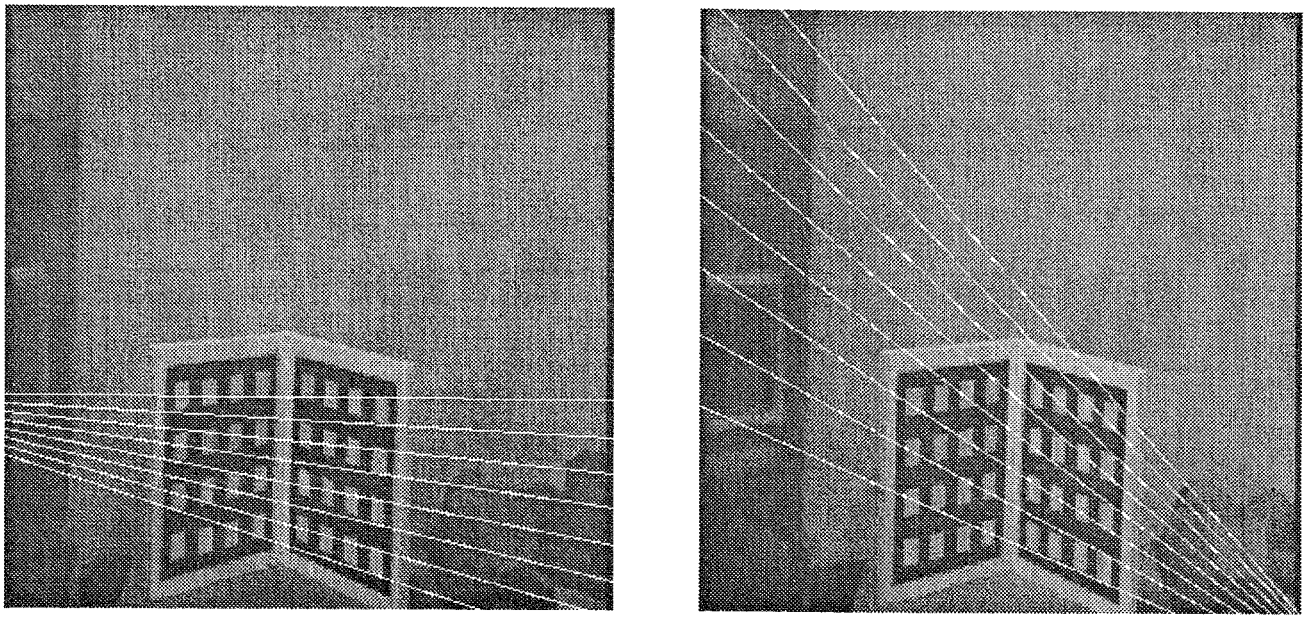


Fig. 4.11: Un exemple de géométrie épipolaire incohérente obtenue par calcul indépendant sur chaque image

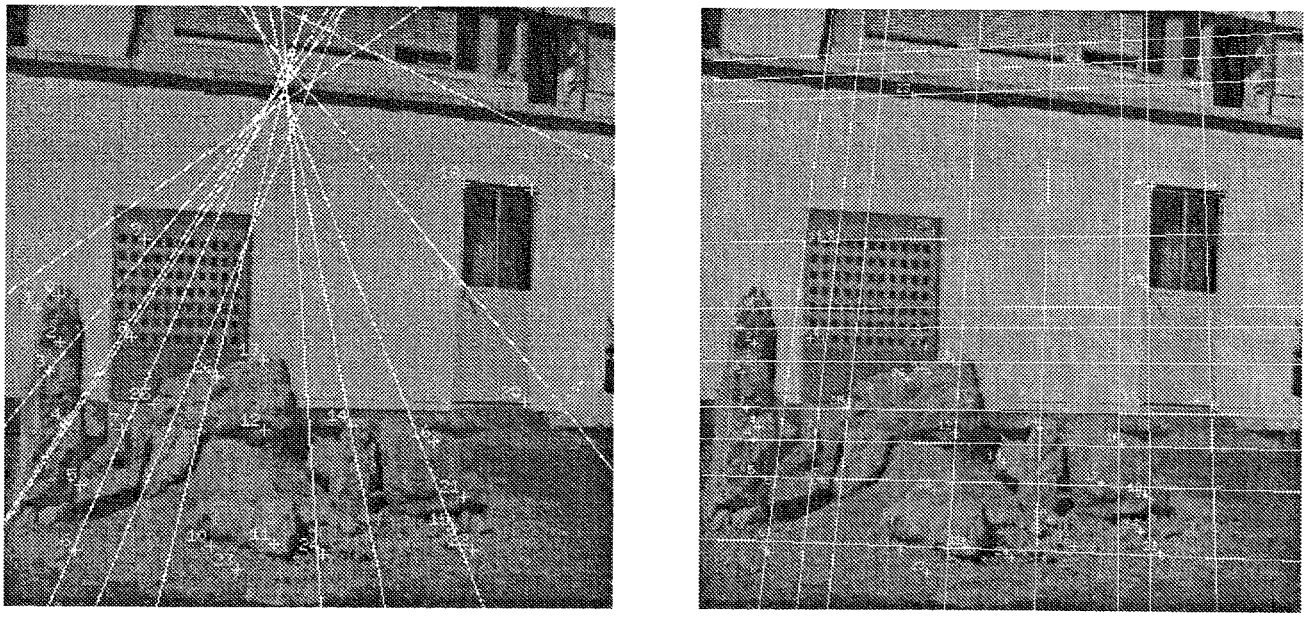


Fig. 4.12: Un exemple de calcul sur des scènes d'extérieur. Linéaire à gauche, non-linéaire (2 faisceaux) à droite.

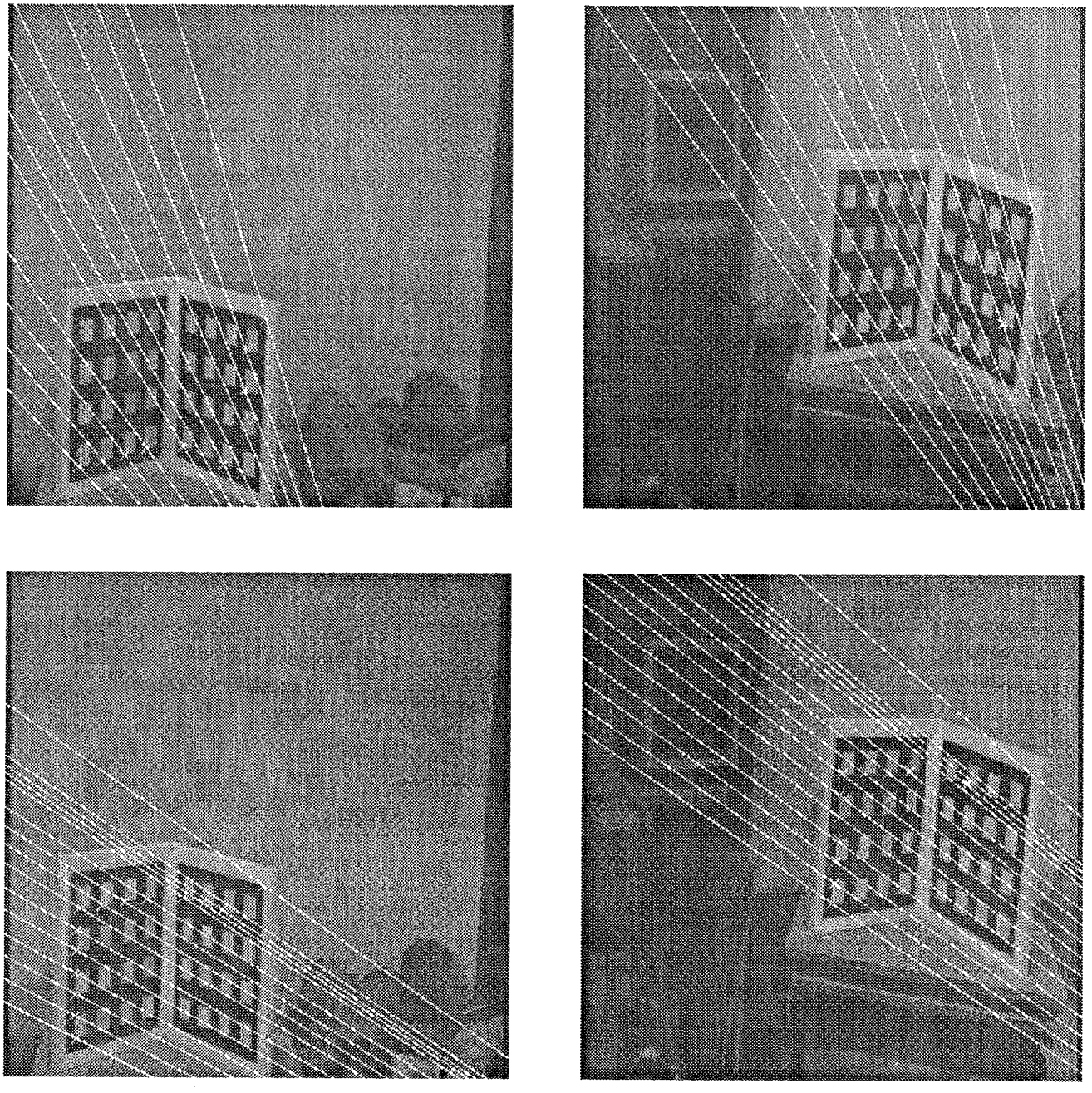


Fig. 4.13: Plusieurs géométries plausibles: Epipolaires estimées (haut) et épipolaires correctes (bas)

Chapitre 5

Une configuration spatiale particulière: les plans

Il est connu que les techniques d'estimation du mouvement par l'intermédiaire de la matrice essentielle échouent si tous les points sont coplanaires [89], et qu'alors d'autres techniques, fondées sur l'existence d'une transformation homographique entre coordonnées de points, sont applicables [162] [41]. Nous allons voir que la situation est assez analogue pour le calcul de la matrice fondamentale.

Nous commençons par exposer, dans un cadre projectif, la correspondance entre les images d'un plan, qui est une simple homographie, et par étudier numériquement son calcul. Nous établissons ensuite une relation nouvelle et très simple entre les homographies définies par des plans et la matrice fondamentale. La première conséquence de l'existence de cette relation est une perte de précision lors du calcul de la matrice fondamentale par l'une des méthodes présentées au chapitre précédent, lorsque les points sont confinés dans un nombre réduit de plans, fait que nous illustrons par des simulations. Nous mettons également en évidence à l'aide d'exemples réels des configurations ambigües obtenues à proximité de plans. Une seconde conséquence de la relation entre homographies et matrice fondamentale est que nous pouvons proposer plusieurs méthodes de celle-ci, qui sont fondées sur la détection des structures planes. Le chapitre s'achève par des expérimentations de ces méthodes, et une comparaison avec les méthodes générales.

5.1 Plans et matrice fondamentale

Nous montrons d'abord que si tous les points observés se trouvent dans un plan, la correspondance entre les images est une homographie, et nous examinons comment nous pouvons calculer cette homographie. Puis nous mettons en évidence une relation très simple entre les homographies définies par des plans et la matrice fondamentale.

5.1.1 La correspondance entre les images d'un plan

Considérons des points $\mathbf{M}_i$ de l'espace se trouvant en configuration générale dans un même plan, et leur image $\mathbf{q}_i$ par une relation projective linéaire de $\mathcal{P}^3$ vers $\mathcal{P}^2$. Celle-ci dans ce cas précis se restreint à une relation entre deux plans, qui, à titre de relation projective de $\mathcal{P}^2$, est une homographie h. Si nous disposons de deux images $\mathbf{q}_i$ et $\mathbf{q}'_i$ des mêmes points $\mathbf{M}_i$ d'un plan, nous pouvons considérer la relation $h' \circ h^{-1}$ entre ces deux images. C'est également une relation homographique, c'est-à-dire qu'il existe une matrice 3×3 inversible $\mathbf{H}$, telle que nous ayons l'égalité projective:

$$\mathbf{q}'_i = \mathbf{H}\mathbf{q}_i \tag{5.1}$$

Le fait que $\mathbf{H}$ soit inversible découle aussi du fait qu'une relation du même type doit exister si nous échangeons les rôles des deux images. L'homographie inverse a pour matrice $\mathbf{H}^{-1}$.

Le fait que nous disposions d'une telle relation entre les coordonnées d'appariements signifie que nous pouvons identifier des plans uniquement à partir de mesures effectuées dans l'image. Une technique de prédiction et vérification a déjà été mise en oeuvre par Lustman [99], qui permet d'identifier un certain nombre de plans à partir de deux images. Récemment des algorithmes assez similaires ont été proposés par Sinclair et collègues [142], ainsi que par Robert [132]. L'idée est de choisir quatre points[1], de calculer l'homographie, dont la connaissance permet de prédire les coordonnées dans la seconde image de tout nouveau point du plan. La position prédite et la position mesurée sont comparées à l'aide d'un seuil sur la distance pour décider si le nouveau point est sur le plan défini par les quatre premiers points.

5.1.2 Calcul de l'homographie

Examinons à présent comment nous pouvons calculer les paramètres de cette homographie à partir d'appariements de points appartenant tous à un même plan. Il est connu qu'une homographie du plan est déterminée par la donnée de deux bases projectives du plan, c'est-à-dire par deux ensembles de quatre points tels que trois quelconques d'entre eux ne soient pas alignés. Il nous faut donc exactement quatre appariements $(\mathbf{q}_i,\mathbf{q}'_i)$, $1 \leq i \leq 4$ afin de pouvoir calculer l'homographie, ce qui est conforme au fait que $\mathbf{H}$, qui est définie à un facteur d'échelle près dépende exactement de huit paramètres indépendants, et que chaque appariement nous fournisse deux équations. Celles-ci sont obtenues à partir de la relation (5.1) en écrivant que

[1] Seuls trois sont nécessaires si la matrice fondamentale est déjà connue.

deux des conditions de proportionnalité de $\mathbf{q}'$ et de $\mathbf{Hq}$ sont satisfaites. On obtient ainsi les deux équations:

$$\begin{cases} \mathbf{V}_1\mathbf{h} = 0 \\ \mathbf{V}_2\mathbf{h} = 0 \end{cases} \tag{5.2}$$

dans lesquelles nous avons posé

$$\begin{aligned} \mathbf{V}_1 &= [u, v, 1, 0, 0, 0, -uu', -vu', -u'] \\ \mathbf{V}_2 &= [0, 0, 0, u, v, 1, -uv', -vv', -v'] \\ \mathbf{h} &= [H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33}]^T \end{aligned}$$

où $(u, v, 1)^T$ et $(u', v', 1)^T$ sont les coordonnées des points $\mathbf{q}$ et $\mathbf{q}'$ respectivement, qui sont à distance finie puisqu'ils sont mesurés dans les rétines. Ces équations sont par conséquent linéaires et homogènes en les coefficients de $\mathbf{H}$. En pratique, nous disposons d'un nombre plus élevé d'appariements, et nous sommes conduits à résoudre les équations (5.2) au sens d'un moindres carrés linéaire, c'est-à dire à minimiser:

$$\min_{\mathbf{h}} \sum_i \{(\mathbf{V}_1\mathbf{h})^2 + (\mathbf{V}_2\mathbf{h})^2\}$$

Cette approche est la plus simple, toutefois en pratique, tout comme dans le cas du calcul de la matrice fondamentale, elle souffre du fait que les quantités minimisées ne soient pas directement significatives en termes de quantités mesurables dans l'image: le critère n'étant pas normalisé est biaisé. Nous pouvons, comme au chapitre précédent, introduire plusieurs critères non-linéaires en minimisant:

1. **DIST:** les distances euclidiennes entre points mesurés dans la seconde image et points calculés à partir de l'homographie, critère que nous symétrisons entre les deux images:

$$\min_{\mathbf{H}} \sum_i \{d^2(\mathbf{q}'_i, \mathbf{Hq}_i) + d^2(\mathbf{q}_i, \mathbf{H}^{-1}\mathbf{q}'_i)\}$$

2. **GRAD:** les équations linéaires pondérées par l'incertitude qui leur est associée

$$\min_{\mathbf{h}} \sum_i \{\frac{(\mathbf{V}_{1i}\mathbf{h})^2}{\|\nabla \mathbf{V}_{1i}\mathbf{h}\|^2} + \frac{(\mathbf{V}_{2i}\mathbf{h})^2}{\|\nabla \mathbf{V}_{2i}\mathbf{h}\|^2}\}$$

Nous avons comparé ces différents critères au moyen d'une méthode statistique. Lors de chacun des 100 essais, nous avons généré des points 3D dans un plan, puis un déplacement rigide. Les 30 appariements sont obtenus en projetant les points du plan à partir des positions de la caméra avant et après déplacement. Nous avons initialisé à chaque fois les critères non linéaires en utilisant le critère linéaire (**LIN**) et les valeurs exactes (**EX**), afin d'effectuer des tests du même type que dans le chapitre précédent. La figure 5.1 présente la moyenne des moyennes des erreurs relatives obtenues sur les neuf coefficients. La table 5.1.2 donne les moyennes des différences réciproques entre les homographies calculées par différentes méthodes. Nous pouvons conclure de ces comparaisons:

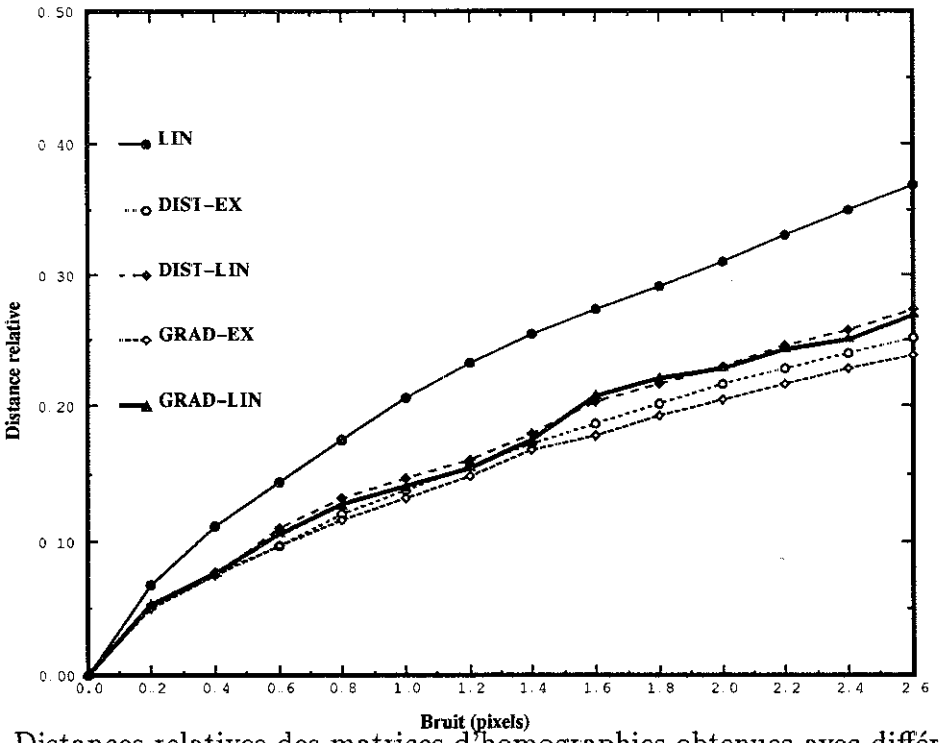


Fig. 5.1: Distances relatives des matrices d'homographies obtenues avec différentes méthodes de calcul

- Les critères non-linéaires donnent des résultats sensiblement meilleurs que le critère linéaire.
- Les résultats obtenus avec les critères non-linéaires sont assez voisins.
- Le calcul est relativement stable[2], il y a moins de problèmes de minima locaux que lors du calcul des *épipoles*. En effet, comme il n'y a pas de contrainte de rang à prendre en compte, nous pouvons utiliser un paramétrage très simple en nous contentant de fixer un coefficient à 1.

5.1.3 Relation entre les homographies de plans et la corrélation fondamentale

Nous savons que la connaissance de la matrice fondamentale permet d'établir une relation projective entre deux rétines, qui est une corrélation: deux points correspondants $\mathbf{q}$ et $\mathbf{q}'$ vérifient $\mathbf{q}'^T\mathbf{F}\mathbf{q} = 0$. Nous venons de voir que l'observation d'un plan Π permet également d'établir une relation projective entre les rétines, qui est une homographie: deux points du plan Π ont des images $\mathbf{q}$ et $\mathbf{q}'$ qui vérifient $\mathbf{q}' = \mathbf{H}\mathbf{q}$. Ces deux relations sont étroitement liées. En effet, étant donné un point quelconque $\mathbf{q}$ de la première rétine nous pouvons toujours

[2]Tout comme pour le calcul de l'homographie épipolaire présenté en (4.4).

bruit		distances relatives				
		LIN	**DIST-EX**	**DIST-LIN**	**GRAD-EX**	**GRAD-LIN**
0.6 pixel	**EXACT**	0.1455	0.0980	0.1111	0.0977	0.1065
	LIN		0.1094	0.1019	0.1063	0.1015
	DIST-EX			0.0415	0.0185	0.0480
	DIST-LIN				0.0502	0.0150
	GRAD-LIN					0.0368
1.6 pixel	**EXACT**	0.2735	0.1878	0.2030	0.1792	0.2085
	LIN		0.2410	0.2321	0.2381	0.2345
	DIST-EX			0.0690	0.0574	0.0939
	DIST-LIN				0.0912	0.0591
	GRAD-LIN					0.0537

Tableau 5.1: Différences réciproques entre les matrices d'homographies calculées par différentes méthodes

le considérer comme l'image d'un point virtuel $\mathbf{M}_\Pi$ du plan Π. L'homographie nous permet alors de calculer l'image $\mathbf{q}'$ de $\mathbf{M}_\Pi$ sur la seconde rétine. Nous obtenons ainsi une paire de points correspondants, d'où $(\mathbf{Hq})^T\mathbf{Fq} = 0$. Cette relation peut se réécrire $\mathbf{m}^T\mathbf{H}^T\mathbf{Fq} = 0$ et doit être satisfaite quel que soit le point $\mathbf{q}$. Nous pouvons en conclure que la matrice $\mathbf{H}^T\mathbf{F}$ est antisymétrique, d'où la relation:

$$\mathbf{H}^T\mathbf{F} + \mathbf{F}^T\mathbf{H} = 0 \tag{5.3}$$

Une seconde manière de retrouver cette relation est d'utiliser les valeurs explicites de la matrice fondamentale et de l'homographie, en fonction des paramètres du mouvement et des paramètres intrinsèques: par définition, $\mathbf{F} = \mathbf{A}'^{-1T}\mathbf{TRA}^{-1}$. Il est connu [99, 39] que, $\mathbf{n}$ étant la normale au plan et d la distance de celui-ci à l'origine:

$$\mathbf{H} = \mathbf{A}'^{-1T}(\mathbf{R} + \frac{1}{d}\mathbf{tn}^T)\mathbf{A}^{-1} \tag{5.4}$$

Nous avons donc à vérifier que:

$$\mathbf{A}^{-1T}\mathbf{R}^T\mathbf{T}^T(\mathbf{R} + \frac{1}{d}\mathbf{tn}^T)\mathbf{A}^{-1}$$

Comme $\mathbf{Tt} = 0$, cette expression se réduit à $(\mathbf{A}^{-1T}\mathbf{R}^T)\mathbf{T}^T\mathbf{RA}^{-1}$, qui est une matrice antisymétrique, puisque $\mathbf{T}$ est antisymétrique.

Notons que si nous échangeons le rôle des deux images, nous obtenons des relations qui sont équivalentes à (5.3). En effet, la matrice d'homographie $\mathbf{H}$ est alors changée en son inverse $\mathbf{H}' = \mathbf{H}^{-1}$ et la matrice fondamentale $\mathbf{F}$ en sa transposée $\mathbf{F}' = \mathbf{F}^T$. La relation (5.3) peut aussi s'écrire:

$$\mathbf{H}^T\mathbf{F} = -\mathbf{F}^T\mathbf{H}$$

En multipliant cette relation à gauche par $\mathbf{H}^{-1T}$ et à droite par $\mathbf{H}^{-1}$, il vient:

$$\mathbf{F}\mathbf{H}^{-1} = -\mathbf{H}^{-1T}\mathbf{F}^T \tag{5.5}$$

ce qui montre que la matrice $\mathbf{H}'^T\mathbf{F}'$ est également antisymétrique.

5.2 Stabilité du calcul de la matrice fondamentale en présence de structures planes

Nous donnons un argument simple, utilisant la relation mise en évidence précédemment, pour expliquer que le calcul de la matrice fondamentale est moins stable si les points sont situés dans des plans. Nous présentons des résultats expérimentaux systématiques illustrant ce fait, et remarquons pour finir que de telles situations ont une importance pratique non négligeable.

5.2.1 L'observation de base

Nous allons à présent voir que si les points utilisés pour le calcul de la matrice fondamentale se situent dans *plusieurs* plans, celui-ci s'effectuera de manière beaucoup moins stable si nous utilisons l'une des méthodes présentées dans le chapitre précédent.

Plaçons nous tout d'abord dans le cas idéal où nous disposons d'appariements qui sont mesurés de manière exacte. Soit donc quatre paires de points appariés qui sont situés sur un même plan. Nous avons donc les relations: $\mathbf{q}'_i = \mathbf{H}\mathbf{q}_i$, qui déterminent entièrement $\mathbf{H}$, puisque chacune de ces relations permet d'obtenir deux équations, et que $\mathbf{H}$ possède 8 paramètres indépendants. Nous avons vu que la relation $\mathbf{q}^T\mathbf{H}^T\mathbf{F}\mathbf{q} = 0$ est alors vérifiée pour tout $\mathbf{q}$. La conséquence est que si nous disposons d'un cinquième appariement $(\mathbf{q}_5, \mathbf{q}'_5)$ obtenu à partir d'un point du même plan nous aurons nécessairement $\mathbf{q}'_5\mathbf{F}\mathbf{q}_5 = 0$. Nous pouvons interpréter ce fait en disant que quel que soit le nombre de points dont nous pouvons disposer, si ceux-ci sont tous situés sur un plan, ils ne génèrent pas davantage de contraintes que quatre points. Les contraintes sont de la forme (5.3), et sont au nombre de six, ce qui est en accord avec le fait que le rang de la matrice $\tilde{\mathbf{U}}$ introduite dans le chapitre précédent en (4.4) pour le calcul par le critère linéaire soit alors de six, comme l'a montré Lustman [99]. Nous retrouvons le fait qu'un ensemble de points situés dans un plan constitue une configuration dégénérée pour le critère de Longuet-Higgins [89].

Dans le cas de données bruitées, il est connu que la covariance d'une estimée aux moindres carrés est, en général proportionnelle aux nombre d'équations de mesure utilisées. L'idée est que, si le problème aux moindres carrés peut s'écrire $\min \|\mathbf{L}\mathbf{x} - \mathbf{b}\|^2$, la solution est donnée par: $\mathbf{x} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{b}$. Avec certaines approximations classiques, $\Lambda_\mathbf{x}$ la matrice de covariance

de x est alors proportionnelle à $(\mathbf{L}^T\mathbf{L})^{-1}$. Si nous écrivons L sous la forme:

$$\mathbf{L} = \begin{pmatrix} \mathbf{l}_1^T \\ \vdots \\ \mathbf{l}_n^T \end{pmatrix}$$

où n est le nombre de mesures, nous voyons que $\mathbf{L}^T\mathbf{L} = \sum_{i=1}^{n} \mathbf{l}_i\mathbf{l}_i^T$ Chaque matrice $\mathbf{l}_i\mathbf{l}_i^T$ est de rang 1. Si chaque vecteur $\mathbf{l}_i$ est indépendant, alors $\Lambda_{\mathbf{x}}$ est approximativement proportionnel à n. Maintenant, suite à la remarque précédente, nous voyons qu'en ajoutant des points à un plan qui est déjà défini, nous ajoutons moins d'information qu'en ajoutant un point indépendant: si il y a déjà quatre points qui définissent effectivement un plan, tout point additionnel de ce plan conduit à des équations $\mathbf{a}_i$ qui sont linéairement dépendantes des équations obtenues avec les quatre points initiaux. L'addition de nouvelles matrices de la forme $\mathbf{a}_i\mathbf{a}_i^T$, ne fait pas progresser le rang de $\mathbf{A}^T\mathbf{A}$. La conclusion est que même si nous avons beaucoup de correspondances situées dans un même plan, elles ne contribuent pas beaucoup plus qu'un faible nombre de correspondances situées dans ce plan.

5.2.2 Etude expérimentale: simulations statistiques

Nous avons vérifié cette observation expérimentalement, à l'aide d'un très grand nombre de simulations. Nous avons généré des ensembles de points 3D placés aléatoirement dans des nombres variables de plans aléatoires, et utilisé un nombre variable de ces points pour l'estimation des épipoles par les meilleures méthodes présentées dans le chapitre précédent. Seuls les points tombant dans une rétine de dimension 512 × 512 ont été utilisés. Nous avons ensuite, en utilisant une méthodologie analogue à celle qui a été employée dans ce même chapitre, obtenu un écart relatif moyen en utilisant pour chaque configuration 100 mouvements. Les résultats sont présentés table 5.2.2, pour deux niveaux de bruit différents, sous forme de tableau à double entrée. Au vu de ces expérimentations, nous pouvons en effet constater que:

- La qualité des estimations croît, à nombre de plans constants, avec le nombre de points, ce qui est classique.
- La qualité des estimations croît, a nombre de points constants, avec le nombre de plans sur lesquels ceux-ci sont répartis.
- La second croissance est plus rapide que la première.
- Des meilleures estimations sont obtenues avec un faible nombre de points répartis sur un nombre élevé de plans plutôt qu'avec un grand nombre de points répartis sur un faible nombre de plans.

	nombre total de points (bruit 0.2 pixel)						
plans	10	20	30	40	50	60	70
2	0.7308	0.6437	0.4836	0.4780	0.4720	0.4690	0.4669
3	0.8368	0.5220	0.4827	0.4529	0.4485	0.4793	0.4640
4	0.5407	0.3138	0.2436	0.2656	0.2473	0.2037	0.2143
5	0.6022	0.3091	0.2634	0.2700	0.2434	0.2088	0.2228
6	0.4960	0.3107	0.2312	0.2235	0.1756	0.1866	0.1910
7	0.3741	0.2114	0.1625	0.1589	0.1423	0.1420	0.1186
8	0.3743	0.1829	0.1349	0.1166	0.1249	0.1011	0.1009
50	0.3541	0.0995	0.0935	0.0778	0.0875	0.0974	0.0848
	nombre total de points (bruit 0.7 pixel)						
plans	10	20	30	40	50	60	70
2	0.8442	0.7465	0.7096	0.6751	0.6441	0.5922	0.6123
3	0.8727	0.7213	0.6825	0.6494	0.6319	0.6289	0.6295
4	0.7161	0.5592	0.4838	0.3638	0.3853	0.3153	0.3537
5	0.7635	0.5240	0.5056	0.4336	0.3889	0.3743	0.3489
6	0.7569	0.5522	0.3843	0.3874	0.3343	0.3080	0.3036
7	0.6047	0.4224	0.3849	0.3261	0.3488	0.3236	0.2889
8	0.6250	0.3979	0.3200	0.2973	0.2984	0.2586	0.2385
50	0.5596	0.2975	0.2632	0.2176	0.2386	0.2069	0.2056

Tableau 5.2: Résultats d'estimation de la matrice fondamentale lorsque les points sont situés dans des plans

5.2.3 Et en pratique ?

Ceci explique les résultats instables que nous avons subis de manière répétée en utilisant la mire de calibration bi-plan pour des tâches d'auto-calibration, bien que les points d'intérêt y soient détectés avec une très bonne précision.

Nous faisons une remarque plus générale, qui est que même si les points utilisés pour le calcul ne se trouvent pas de manière exacte dans une configuration instable, mais en sont suffisamment proches, l'instabilité aura tout de même lieu. Ainsi, il suffit que les points utilisés soient suffisamment proches d'un plan pour que nous observions des phénomènes d'ambiguïté. Montrons sur un exemple typique d'images réelle ce phénomène. La partie supérieure de la figure 5.2 présente une paire d'images entre lesquelles des points ont été appariés. Bien que ces points ne soient pas coplanaires, nous les avons considérés comme tels, et nous avons calculé une homographie entre les deux ensembles de points. La partie inférieure de la figure 5.2 montre le résultat de l'application de l'homographie obtenue (respectivement de son inverse) à la première (respectivement seconde) image. Le très faible écart que nous pouvons observer entre les points réels et les points prédits par l'homographie montre que bien que les points ne soient pas coplanaires, ils sont bien approximés par un plan. Une conséquence

est l'ambiguïté de la géométrie épipolaire, illustrée par la figure 5.3, où nous pouvons voir plusieurs faisceaux d'épipolaires, que nous avons obtenu en minimisant différents critères, qui sont également cohérents[3], et significativement différents les uns des autres.

Dans les images d'environnements robotiques, des scènes du type de celle qui nous a servi d'exemple, c'est-à-dire où l'intervalle de profondeur dans lequel se trouvent les objets observés est assez faible, sont assez courantes. Il est nécessaire d'éviter d'utiliser de telles images pour des tâches nécessitant la détermination de la matrice fondamentale. Un test simple consiste à calculer une homographie, et à comparer le résidu de l'évaluation à un seuil. A titre d'exemple, avec les points de la paire d'images 5.2, l'erreur résiduelle constituée par la moyenne des distances des points prédits par une homographie aux points effectivement mesurés, est de 0.5 pixel. Dans les images 4.7 du chapitre précédent, qui contiennent des points qui ont une disparité plus marquée, puisque la profondeur de l'image est celle d'un bureau, l'erreur résiduelle est de 4 pixels. Ce n'est pas une valeur très élevée, néanmoins l'ordre de grandeur de différence suffit pour éviter les problèmes d'ambiguïté dans ce cas. En revanche, il est un peu plus compliqué de tester la proximité de plusieurs plans.

5.3 Calcul de la matrice fondamentale à partir de plans

Nous détaillons une méthode de détermination de la matrice fondamentale à partir de matrices d'homographies, dont nous donnons également une interprétation géométrique en termes de méthode de détermination des épipoles. Nous présentons ensuite à titre de résultats expérimentaux une comparaison avec les méthodes générale, et discutons des raisons de la meilleure stabilité de celles-ci.

5.3.1 Une méthode de calcul direct de la matrice fondamentale

La relation matricielle (5.3) nous fournit six équations qui sont linéaires et homogènes en les coefficients de $\mathbf{F}$. Supposons à présent que nous disposons de deux homographies de matrice $\mathbf{H}_1$ et $\mathbf{H}_2$, obtenues en observant deux plans distincts de la scène. En joignant les deux systèmes d'équations (5.3), dont la matrice de chacune a la forme:

$$\begin{pmatrix} H_{11} & 0 & 0 & H_{21} & 0 & 0 & H_{31} & 0 & 0 \\ H_{12} & H_{11} & 0 & H_{22} & H_{21} & 0 & H_{32} & H_{31} & 0 \\ H_{13} & 0 & H_{11} & H_{23} & 0 & H_{21} & H_{33} & 0 & H_{31} \\ 0 & H_{12} & 0 & 0 & H_{22} & 0 & 0 & H_{32} & 0 \\ 0 & H_{13} & H_{12} & 0 & H_{23} & H_{22} & 0 & H_{33} & H_{32} \\ 0 & 0 & H_{13} & 0 & 0 & H_{23} & 0 & 0 & H_{33} \end{pmatrix} \tag{5.6}$$

nous obtenons un système $\mathcal{S}$ linéaire homogène de douze équations linéaires en les coefficients de la matrice fondamentale $\mathbf{F}$. Ces douze équations doivent former un système linéaire de

[3]Les distances des droites épipolaires aux correspondants sont largement inférieures au pixel dans les quatre cas. Nous n'avons tracé qu'une partie des droites épipolaires pour des raisons de clarté.

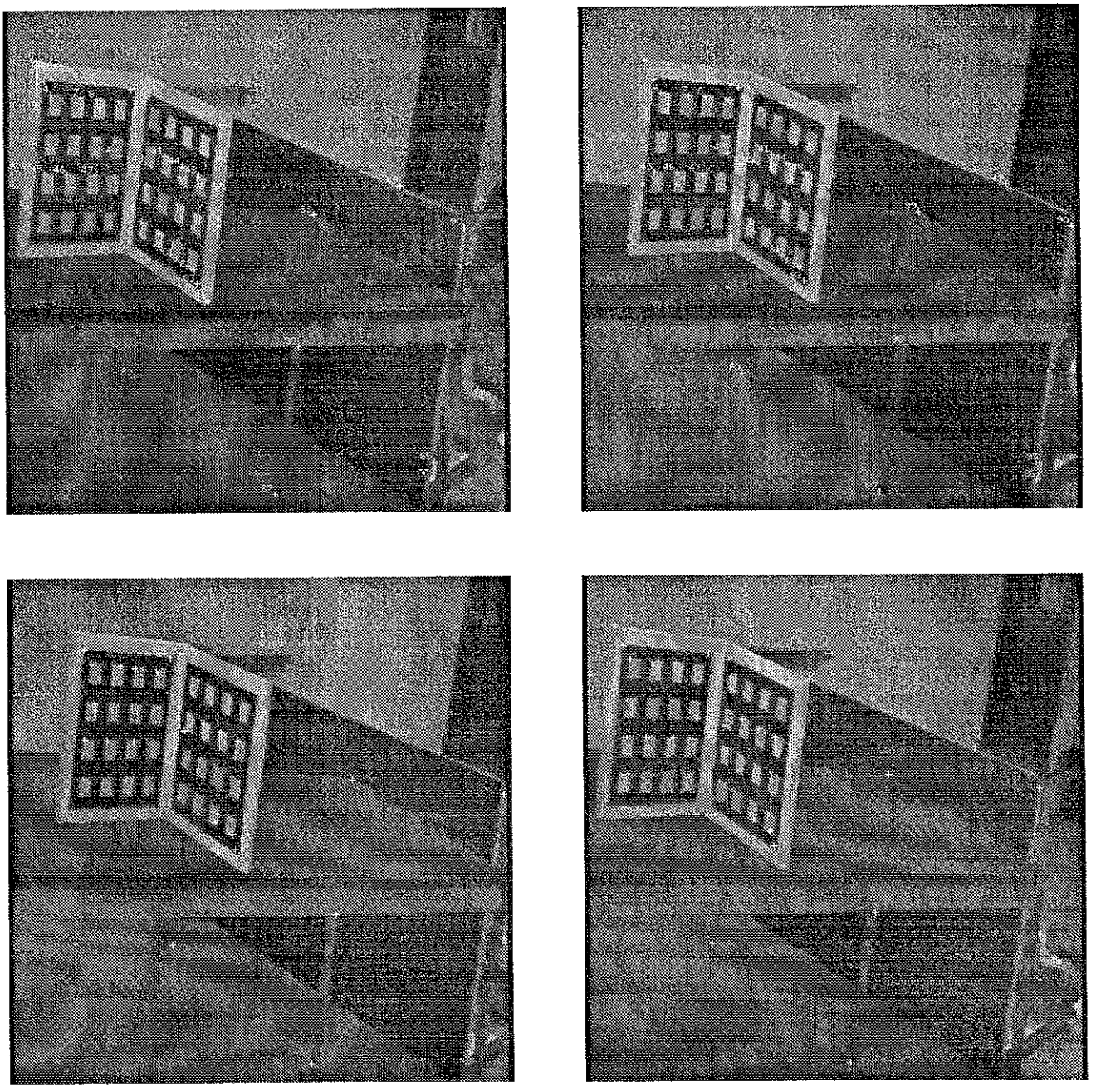


Fig. 5.2: Appariements réels (haut), et prédits en utilisant une homographie (bas)

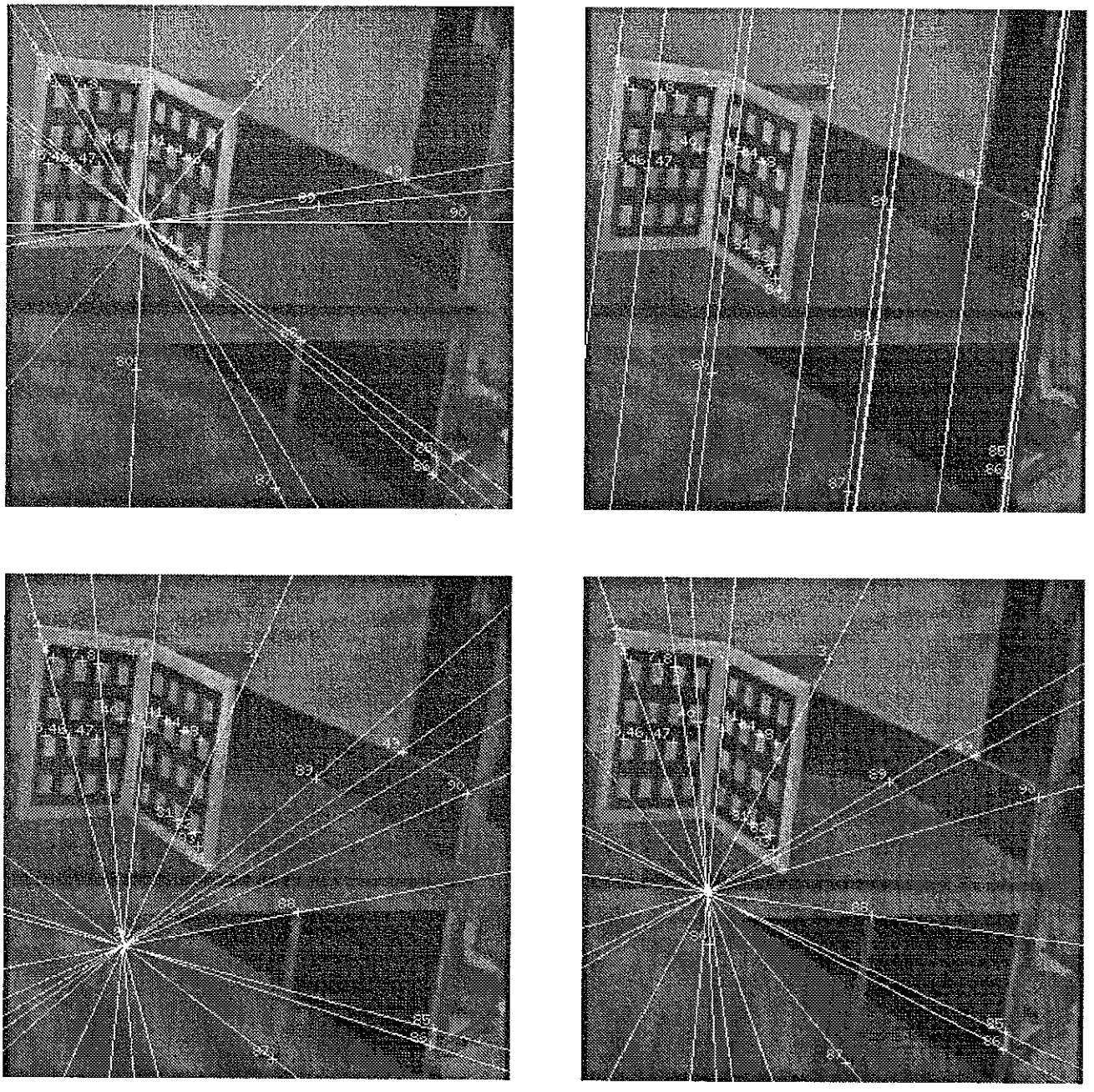


Fig. 5.3: Ambiguïté: quatre exemples de faisceaux d'épipolaires cohérents dans la première image

rang réduit à huit pour qu'existe une solution, ce qui introduit quatre contraintes algébriques portant sur les coefficients des matrices d'homographie $\mathbf{H}_1$ et $\mathbf{H}_2$, qui peuvent être obtenues en écrivant que quatre sous-matrices 8×8 choisies pour être indépendantes dans la matrice de $\mathcal{S}$ ont un déterminant nul. La matrice de $\mathcal{S}$ est de dimension 12×9, aussi nous pouvons choisir par exemple les matrices formées des lignes (1,...,9), (2,...,10), (3,...,11), et (4,...,12). L'existence de ces contraintes peut aussi s'interpréter par le fait que, les deux caméras étant fixées, l'expression (5.4), limite les matrices d'homographies possibles.

En pratique, il n'est pas facile d'exploiter ces contraintes algébriques, car même après simplification elles se révèlent de degré encore élevé. Nous pouvons cependant exploiter les équations (5.3) de manière simple en vue de calculer la matrice fondamentale. Il suffit de disposer d'un nombre N d'homographies supérieur ou égal à deux. Chaque homographie nous fournissant six équations linéaires en les coefficients de la matrice fondamentale $\mathbf{F}$, il est alors possible de résoudre les $6N$ équations au sens des moindres carrés. Une première amélioration de cette approche consiste à utiliser l'un des paramétrages non-linéaire (**NON-LIN**) de la matrice fondamentale présentés dans le chapitre précédent qui permettent de tenir compte des propriétés de celle-ci, ce qui conduit, avec les mêmes équations, à un moindres carrés non-linéaire. Les raisons pour lesquelles il est important de tenir compte de ces propriétés ont été largement discutées dans le chapitre précédent. Une seconde amélioration consiste à symétriser le critère de calcul de la matrice fondamentale (**SYM**). En effet les équations (5.3) sont obtenues en écrivant des contraintes portant seulement sur la seconde image, et comme le montre l'exemple du tableau 5.3.1, les épipoles obtenus dans la première image à l'aide de ces seules équations (**NON-SYM**) sont très nettement moins bien localisés que ceux obtenus dans la seconde image. Notre approche est donc de minimiser simultanément les équations (5.3) et les équations (5.5), bien que ces dernières soient redondantes, comme nous l'avons déjà signalé plus haut. Un exemple permettant de constater l'effet sur les résultats est présenté dans le tableau 5.3.1. Notons que si nous disposons d'un seul plan et d'un minimum de deux points appariés n'appartenant pas à ce plan, nous pouvons également obtenir un nombre d'équations linéaires suffisant pour déterminer la matrice fondamentale: six équations (5.3) à partir de l'homographie, et deux équations (4.1) à partir des points.

5.3.2 Interprétation géométrique: structures planes et détermination des épipoles

Nous partons de l'observation suivante, illustrée par la figure 5.4: supposons que nous ayons deux caméras, et que nous observons dans la seconde caméra les images $\mathbf{q}'_1$ et $\mathbf{q}'_2$ de deux points $\mathbf{M}_1$ et $\mathbf{M}_2$ de l'espace ayant une image identique par la première caméra. Alors $\mathbf{q}'_1\mathbf{q}'_2$ est une droite épipolaire de la seconde caméra. En effet, le fait que $\mathbf{M}_1$ et $\mathbf{M}_2$ aient une image identique par la première caméra signifie que la droite $\mathbf{M}_1\mathbf{M}_2$ est un rayon optique de cette caméra. La droite $\mathbf{q}'_1\mathbf{q}'_2$, image d'un rayon optique de la première caméra, est donc une droite épipolaire de la seconde caméra. L'idée derrière toutes les méthodes de calcul examinées ici est que si nous observons un plan, et que nous connaissons l'homographie reliant l'image des points de ce plan par chacune des caméras, nous pouvons *calculer* de tels points $\mathbf{q}'_i$ à

EXACT	780.08	-427.04	1823.82	-1255.57
	LIN			
0.1 pixel	677.93	-867.29	2104.51	-1031.93
0.5 pixel	452.82	-1263.39	1628.43	59.84
1.0 pixel	330.67	-1567.84	1467.22	378.28
	NON-LIN,NON-SYM			
0.1 pixel	711.67	-645.84	1835.76	-1272.76
0.5 pixel	576.56	-865.67	1859.75	-1328.78
1.0 pixel	227.31	-1237.21	1838.19	-1438.38
	NON-LIN,SYM			
0.1 pixel	782.70	-433.09	1845.02	-1283.58
0.5 pixel	781.70	-455.56	1870.00	-1362.63
1.0 pixel	760.32	-490.37	1887.33	-1436.37

Tableau 5.3: Un exemple illustrant les résultats obtenus avec différentes méthodes de calcul de $\mathbf{F}$ à partir de matrices $\mathbf{H}_i$.

partir de leur images par la première caméra $\mathbf{q}_i$, en considérant des points $\mathbf{M}_i$ virtuels. Nous pouvons exploiter cette idée de deux manières différentes.

Une première possibilité est d'utiliser deux plans Π_1 et Π_2. Soit un point $\mathbf{M}$ de l'espace, ayant pour image $\mathbf{q}$ par la première caméra. Nous pouvons considérer les points $\mathbf{M}_1$ et $\mathbf{M}_2$ de Π_1 et Π_2, respectivement, ayant pour image $\mathbf{q}$ par la première caméra. L'image de $\mathbf{M}_i$ dans la seconde caméra est $h_{\Pi_i}(\mathbf{q})$, avec $i = 1, 2$. Nous obtenons donc une droite épipolaire $h_{\Pi_1}(\mathbf{q})h_{\Pi_2}(\mathbf{q})$ à partir de tout point $\mathbf{m}$ de la première image. Cette approche ne nécessite donc que l'identification de deux fois quatre points coplanaires. Montrons qu'elle permet d'obtenir directement les coordonnées des épipoles en fonction des matrices d'homographies. Supposons que nous ayons pu identifier ainsi deux homographies de plans, h_1 et h_2, de matrices respectives $\mathbf{H}_1$ et $\mathbf{H}_2$. L'épipole $\mathbf{e}'$ est le point vérifiant:

$$\forall \mathbf{q} \qquad \mathbf{e}'^T(\mathbf{H}_1\mathbf{q} \wedge \mathbf{H}_2\mathbf{q}) = 0$$

Cette équation définit une forme bilinéaire en les coordonnées de $\mathbf{q}$ et peut par conséquent se réécrire sous la forme:

$$\forall \mathbf{q} \qquad \mathbf{q}^T\mathbf{B}\mathbf{q} = 0$$

où $\mathbf{B}$ est une matrice homogène 3×3 dépendant quadratiquement de $\mathbf{H}_1$, $\mathbf{H}_2$, et linéairement de $\mathbf{e}'$. De la dernière relation, nous déduisons que la matrice $\mathbf{B}$ doit être antisymétrique, ce qui nous donne un système $\mathcal{S}_0$ de six équations linéaires en les coordonnées de l'épipole $\mathbf{e}'$ que nous pouvons résoudre en pratique au sens des moindres carrés. Si nous notons $\mathbf{h}_1^i$ (resp.

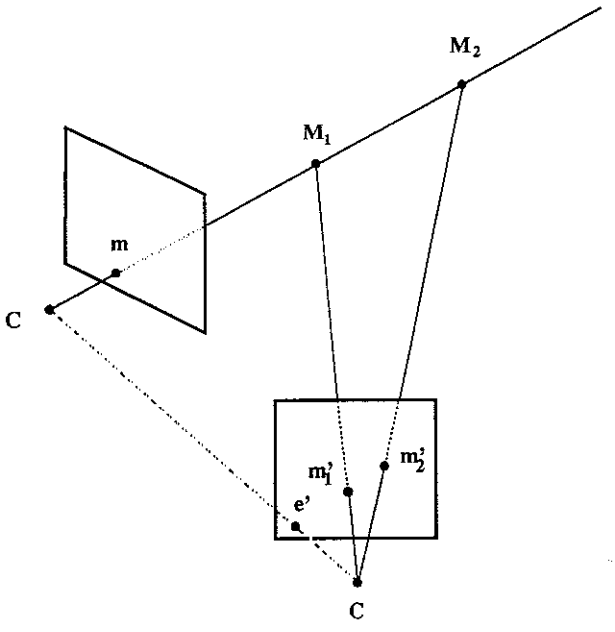


Fig. 5.4: Droite épipolaire obtenue en observant deux points particuliers

$\mathbf{h}_2^i$) les vecteurs colonne de $\mathbf{H}_1$ (resp. $\mathbf{H}_2$), la matrice de ce système s'écrit:

$$\begin{pmatrix} (\mathbf{h}_1^1 \wedge \mathbf{h}_2^1)^T \\ (\mathbf{h}_1^2 \wedge \mathbf{h}_2^2)^T \\ (\mathbf{h}_1^3 \wedge \mathbf{h}_2^3)^T \\ (\mathbf{h}_1^1 \wedge \mathbf{h}_2^2)^T + (\mathbf{h}_1^2 \wedge \mathbf{h}_2^1)^T \\ (\mathbf{h}_1^1 \wedge \mathbf{h}_2^3)^T + (\mathbf{h}_1^3 \wedge \mathbf{h}_2^1)^T \\ (\mathbf{h}_1^2 \wedge \mathbf{h}_2^3)^T + (\mathbf{h}_1^3 \wedge \mathbf{h}_2^2)^T \end{pmatrix} \tag{5.7}$$

Notons toutefois qu'en théorie, ces six équations doivent former un système linéaire de rang réduit à deux pour qu'existe une solution, ce qui introduit quatre contraintes algébriques portant sur les coefficients des matrices d'homographie $\mathbf{H}_1$ et $\mathbf{H}_2$, qui peuvent être obtenues en écrivant que quatre sous-matrices 3×3 choisies de façon à être indépendantes dans la matrice de $\mathcal{S}_0$ ont un déterminant nul. La matrice de $\mathcal{S}_0$ est de dimension 6×3, aussi nous pouvons choisir par exemple les matrices formées des lignes (1,2,3), (2,3,4), (3,4,5), et (4,5,6). Nous avons montré, à l'aide du système MAPLE, que les deux ensembles de quatre contraintes algébriques sont équivalents. Ces relations peuvent être utilisées pour la recherche de plans en contraignant les homographies possibles. Cette remarque fait le lien avec la méthode de calcul direct de la matrice fondamentale que nous avons présentée plus haut.

Une seconde possibilité ne nécessite que l'emploi d'un seul plan Π, Soit un point $\mathbf{M}$ de l'espace, ne se situant pas sur Π, ayant pour image $\mathbf{q}$ par la première caméra et $\mathbf{q}'$ par la seconde caméra. Considérons $\mathbf{M}_1$, point de Π ayant même image par la première caméra que

M. Nous pouvons, à l'aide de l'homographie h_Π associée au plan Π, calculer l'image de $\mathbf{M}_1$ dans la seconde caméra, qui est $h_\Pi(\mathbf{q})$. D'après la remarque précédente, $\mathbf{q}' h_\Pi(\mathbf{q})$ est une droite épipolaire. Dans le cas général, deux droites épipolaires sont suffisantes pour déterminer l'épipole, si bien que le nombre minimum d'appariements requis par cette approche est six, dont quatre points coplanaires et deux points n'appartenant pas au plan, ce qui correspond à la remarque que nous avions faite à propos de la méthode de calcul direct de la matrice fondamentale que nous avons présentée plus haut. En pratique, nous disposons de bien plus de points n'appartenant pas au plan, et nous calculons l'intersection par une minimisation aux moindres carrés. L'épipole est cherché comme le point e qui minimise la somme des distances aux droites épipolaires:

$$\sum_i d^2(\mathbf{e}, l_i)$$

Le critère étant non-linéaire, nous l'initialisons à l'intersection de deux droites épipolaires choisies arbitrairement.

Nous avons implémenté cette dernière approche de la manière suivante:

- Calculer une homographie entre images de points coplanaires.
- Pour chaque point apparié de la première image n'appartenant pas au plan, calculer sa droite épipolaire dans la seconde image.
- calculer l'intersection au sens des moindres carrés de toutes les droites épipolaires ainsi générées.

Les résultats obtenus sont extrêmement sensibles au bruit, comme le montre le tableau 5.3.2, qui présente des statistiques obtenus avec les données synthétiques. Nous avons juste indiqué la moyenne sur les quatre coordonnées des épipoles de la moyenne des erreurs relatives obtenues en effectuant 100 essais en variant le mouvement et les points, 30 points étant à chaque fois situés dans un plan et 20 points situés en dehors. La grande sensibilité au bruit

	0.4 pixel	0.8 pixel	1.2 pixel	1.6 pixel
erreur	0.4864	0.6101	0.6846	0.7287

Tableau 5.4: Résultats de détermination de l'épipole en utilisant un seul plan et des points extérieurs

de cet algorithme s'explique par la difficulté, déjà rencontrée avec l'une des méthodes de Sturm, de localiser le point d'intersection de courbes bruitées passant toutes par ailleurs dans une zone fixe du plan, l'image. Cette remarque est illustrée par la figure 5.5. Le haut de cette figure, obtenu avec les données réelles extraites (avec une bonne précision) de la mire, montre un ensemble de droites épipolaires calculées à partir d'une homographie et d'appariements, pour lequel le point d'intersection (430,368) n'est pas déterminé avec une bonne précision. L'image en bas à gauche permet de comparer ces droites avec le faisceau

d'épipolaires correct, passant par l'épipole (640,213), obtenu par un calcul identique, mais en partant des points exacts obtenus à l'aide du modèle de la mire. Celle en bas à droite permet de constater que les droites épipolaires (bruitées) obtenues à l'aide de l'homographie passent pourtant par les points correspondants avec une très bonne précision.

5.3.3 Une comparaison: résultats expérimentaux

Voyons d'abord un exemple plutôt favorable à la méthode des plans: la mire de calibration bi-plan. Le nombre élevé de points dans chaque plan (64) permet une détermination précise des homographies. En revanche le faible nombre de plans (2) et le fait que les points soient confinés dans une région limitée de l'espace est défavorable à la méthode générale. La table 5.3.3 montre des résultats obtenus avec les trois mouvements réels conduisant aux quatre images présentées figure 5.6. Nous pouvons constater que s'il y a une amélioration très nette lors du premier mouvement, il y a une dégradation (peu significative) lors du troisième.

	$1 \rightarrow 2$			
EXACT	780.08	-427.04	1823.82	-1255.57
GEN	-31.68	709.76	174.26	543.61
PLANS	781.70	-455.56	1870.00	-1362.63
	$2 \rightarrow 3$			
EXACT	1634.30	-1910.35	3859.44	-3830.30
GEN	744.03	-573.65	1112.87	-950.32
PLANS	943.96	-727.65	1386.38	-1028.40
	$3 \rightarrow 4$			
EXACT	517.49	203.88	349.39	208.09
GEN	523.04	205.67	353.92	209.81
PLANS	538.12	206.70	366.34	210.94

Tableau 5.5: Résultats obtenus sur les images de mire avec le critère général et celui fondé sur les plans.

Nous avons ensuite procédé de manière statistique, en générant des ensembles d'appariements de 50 à 100 points situés dans un nombre variable de plans, comme dans la section précédente. Nous avons comparé, à deux niveaux de bruit, les résultats obtenus par la méthode générale et par la méthode des plans. Ces résultats sont résumés sur le tableau 5.3.3. Nous pouvons conclure que dès que le nombre de plans disponibles est supérieur ou égal à trois, la méthode de calcul général donnera de meilleurs résultats que la méthode des plans. L'augmentation du nombre de plans permet ensuite d'améliorer de manière significative les résultats de la méthode générale, alors que ceux de la méthode des plans se dégradent légèrement.

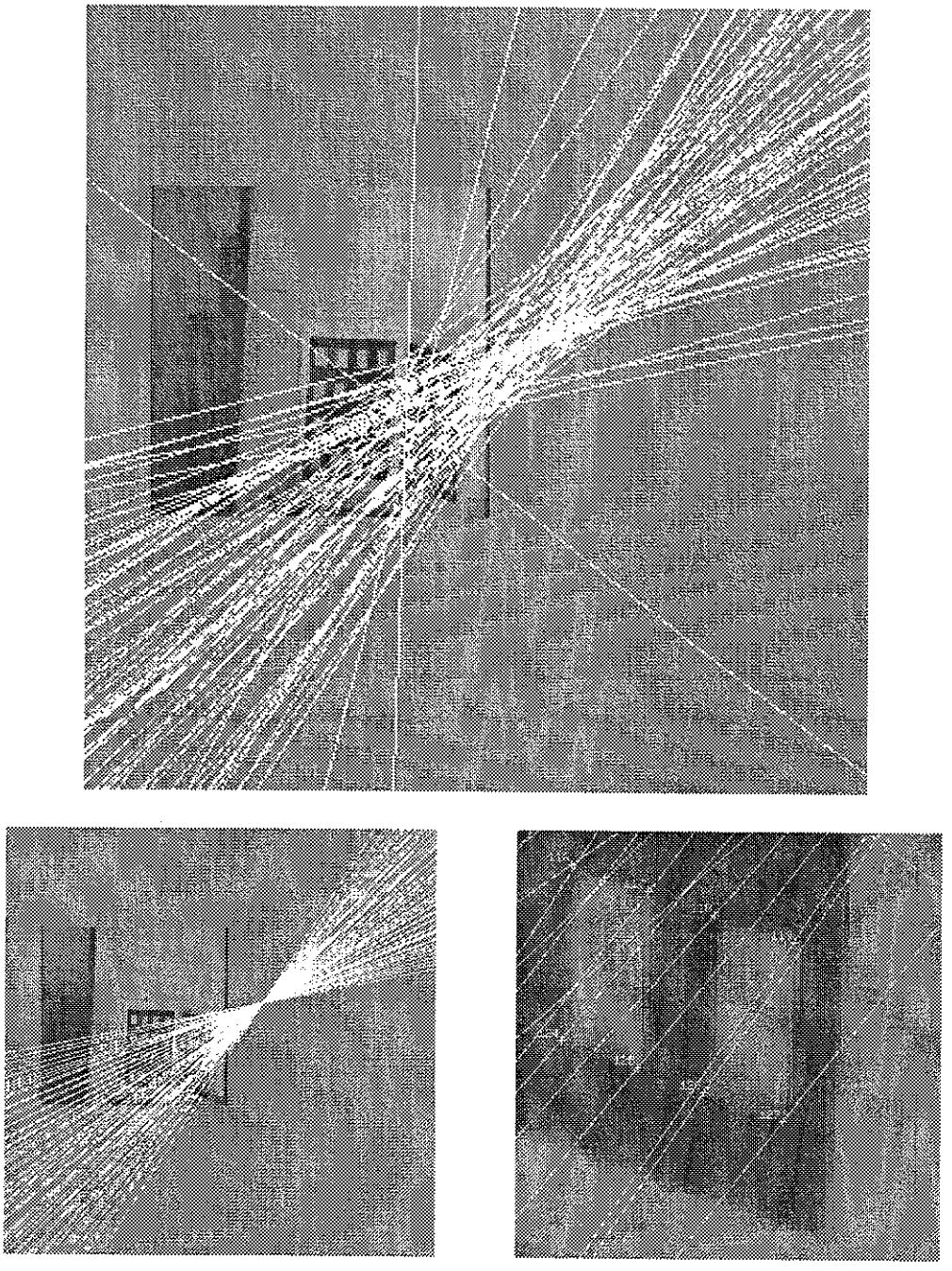


Fig. 5.5: Ensemble d'épipolaires obtenues par homographie

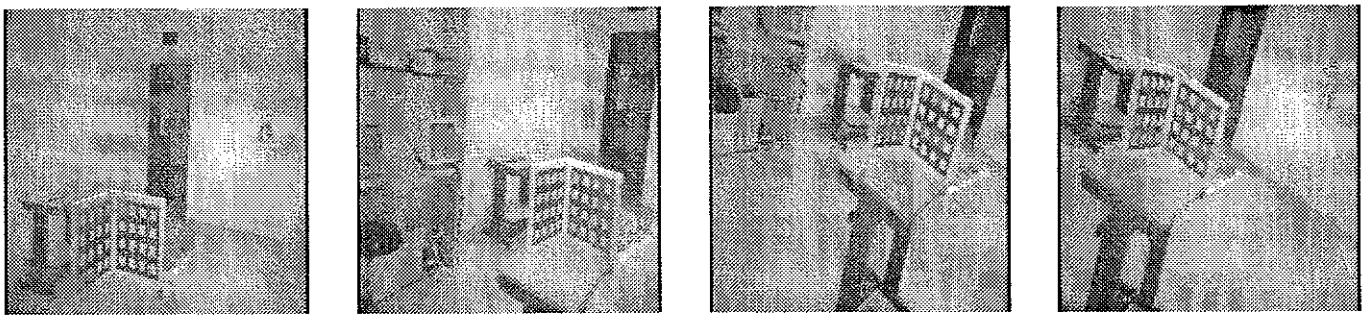


Fig. 5.6: Les quatre images de mire utilisées à titre d'exemple

		nombre de plans						
bruit	méthode	2	3	4	5	6	7	8
0.2 pixel	**PLANS**	0.1961	0.2419	0.3056	0.3145	0.2613	0.2996	0.3146
	GEN	0.2508	0.1713	0.1158	0.0943	0.0697	0.0702	0.0683
1 pixel	**PLANS**	0.4199	0.4462	0.5955	0.5331	0.5668	0.5806	0.5893
	GEN	0.4853	0.4017	0.3390	0.2926	0.3004	0.2749	0.2651

Tableau 5.6: Comparaison statistique de la méthode des plans et de la méthode générale

Nous pouvons donner une explication de la faiblesse de ces résultats par un calcul d'erreur simplifié. Dans la méthode des plans, la matrice fondamentale est finalement obtenue à partir d'équations du type (5.3), qui ont une forme similaire aux équations (4.1). Afin de déterminer $\mathbf{F}$ avec une aussi bonne précision à partir des équations (5.3) qu'à partir des équations (4.1), les coefficients de (5.3), qui sont les éléments des matrices d'homographies, devraient être déterminés avec la même précision que ne le sont les coefficients de (4.1), qui sont des fonctions très simples des coordonnées des points $\mathbf{q}$ mesurés. Plaçons nous, pour fixer les idées, dans le cas où le bruit a pour déviation standard $\sigma = 1$ pixel. Considérons à les coefficients des équations (4.1). Ils sont de la forme $C_1 = ab$ ou $C_2 = a$, a et b étant des coordonnées affines de points appariés. En faisant l'hypothèse d'une covariance isotrope, de valeur σ, sur les points appariés, nous obtenons des covariances de la forme $\sigma_{C_1} = \sigma\sqrt{a^2+b^2}$, $\sigma_{C_2} = \sigma$. L'erreur relative σ_i/C_i est donc une fonction décroissante de la distance à l'origine. Le cas moyen correspond au milieu de l'image, qui est avec les images 512×512 que nous utilisons, le point (255,255). L'erreur relative moyenne pour les coefficients de la forme C_1 (respectivement C_2) est donc $\frac{\sqrt{2}}{255} = 5.5\%$ (respectivement $\frac{1}{255} = 3.9\%$). D'après les courbes de la figure 5.1, nous ne pouvons estimer une matrice d'homographie qu'avec une précision relative de 13% sur les coefficients en moyenne. Ceci correspond à un cas relativement favorable où nous disposons de 30 appariements par plan, ce qui est déjà

beaucoup

Comme dans nos expérimentations le nombre total de points est fixé, lorsque nous augmentons le nombre de plans, nous diminuons le nombre de points par plan, ce qui diminue d'autant la précision du calcul des homographies associées à chaque plan. Nous obtenons donc davantage d'équations, mais celles-ci sont plus bruitées, ce qui explique que la précision ne s'améliore pas.

Nous pouvons résumer cette discussion en concluant que dans le cas de la méthode des plans, nous remplaçons des coefficients qui sont des coordonnées, ou produits de coordonnées, par des coefficients de matrices d'homographies que nous estimons avec une plus grande erreur relative. Pour ce qui est du nombre des équations, nous obtenons pour chaque plan six équations (5.3) indépendantes, pour remplacer un nombre bien plus élevé d'équations non indépendantes. Ces raisons expliquent que cette méthode soit davantage sensible au bruit que la méthode générale.

Chapitre 6

L'influence des mouvements sur la stabilité

Après avoir étudié l'effet de certaines configurations de points, nous étudions à présent l'influence sur la stabilité du calcul de la matrice fondamentale, du type du déplacement effectué par la caméra.

Nous montrons d'abord qu'une bonne mesure de la stabilité de l'estimation est obtenue par des calculs de covariance. A cet effet, nous comparons expérimentalement les covariances obtenues par simulation avec les covariances prédites par deux méthodes différentes, l'une, classique, fondée sur le critère linéaire, l'autre, plus nouvelle, sur les minimisations de critères quelconques. Les calculs de covariance nous permettent d'estimer expérimentalement la stabilité, mais non de comprendre les situations instables.

La seconde idée que nous présentons est celle de surface critique. La proximité des points de l'espace d'une telle surface peut être caractérisée par des calculs effectués à partir des seules images, en déterminant une approximation par une transformation quadratique, calcul pour lequel nous présentons plusieurs méthodes linéaires et non-linéaires.

Nous sommes alors en mesure de tenter une caractérisation des situations instables. Celles-ci se produisent pour trois types de mouvements: ceux dont la composante translationnelle est petite, ceux pour lesquels elle est parallèle au plan image, et, dans une certaine mesure, les translations pures. Le deuxième type d'instabilité, plus subtil, résulte de l'*interaction* du mouvement et de la structure tridimensionnelle. Il se produit à proximité de la surface critique. Cette caractérisation est validée par des simulations statistiques analytiques et synthétiques.

6.1 Une caractérisation probabiliste de la stabilité

Le but de cette section est de montrer que le calcul de covariance est un outil qui nous permet de caractériser de manière probabiliste la stabilité d'une façon assez fiable. Deux méthodes de calcul de la covariance sont d'abord présentées.

6.1.1 Le calcul de l'incertitude: cas linéaire

Nous utilisons le formalisme classique des matrices de covariance pour caractériser l'incertitude associée à l'estimation. Si les mesures sont modelisées par $\mathbf{x}$, vecteur aléatoire de $\mathbb{R}^P$ de moyenne $\mathbf{x}_0$ et de covariance $\mathbf{\Lambda_x} = E((\mathbf{x}-\mathbf{x}_0)^T(\mathbf{x}-\mathbf{x}_0))$, alors le vecteur $\mathbf{y} = f(\mathbf{x})$ est aussi aléatoire, de moyenne $f(\mathbf{x}_0)$ et de covariance:

$$\mathbf{\Lambda_y} = \mathbf{J}_f(\mathbf{x}_0)\mathbf{\Lambda_x}\mathbf{J}_f(\mathbf{x}_0)^T \tag{6.1}$$

Dans cette expression, $\mathbf{J}_f(\mathbf{x}_0)$ est la matrice jacobienne de f, calculée au point $\mathbf{x}_0$.

Dans le cas du critère linéaire, nous possédons une solution explicite, ce qui nous permet d'utiliser le résultat précédent. Des calcul similaires ont été introduits par Weng, Huang et Ahuja [172], ainsi que par Philip [128], cependant notre formulation est plus concise. Si nous utilisons la normalisation consistant à fixer à 1 l'un des coefficients de $\mathbf{F}$, le critère (4.3) peut se mettre sous la forme:

$$\min_{\mathbf{y}} \|\mathbf{L}\mathbf{y} - \mathbf{b}\|^2 \tag{6.2}$$

où $\mathbf{L}$ est une matrice $N \times 8$, $\mathbf{b}$ un vecteur $N \times 1$, et $\mathbf{y}$ le vecteur 8×1 de paramètres recherchés, N étant de nombre de points appariés utilisés pour le calcul. Il est connu que le problème (6.2) admet une solution unique donnée par:

$$\mathbf{y} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{b} \tag{6.3}$$

Cette formule explicite nous permet de calculer les dérivées de $\mathbf{y}$ par rapport aux mesures $w_j, j = 1 \ldots N$, (w_j étant u_j, v_j, u'_j, v'_j) par la formule:

$$\frac{\partial \mathbf{y}}{\partial w_j} = \left\{ -(\mathbf{L}^T\mathbf{L})^{-1}\left(\frac{\partial \mathbf{L}}{\partial w_j}^T\mathbf{L} + \mathbf{L}^T\frac{\partial \mathbf{L}}{\partial w_j}\right)(\mathbf{L}^T\mathbf{L})^{-1} \right\}\mathbf{L}^T\mathbf{b} + (\mathbf{L}^T\mathbf{L})^{-1}\left(\frac{\partial \mathbf{L}}{\partial w_j}\mathbf{b} + \mathbf{L}^T\frac{\partial \mathbf{b}}{\partial w_j}\right) \tag{6.4}$$

Voyons à présent comment effectuer les calculs de manière efficace. Dans notre cas, chaque ligne de $\mathbf{L}$ a la même forme $\mathbf{l}_j^T$, et ne dépend que des mesures (u_j, v_j, u'_j, v'_j). Nous avons donc les simplifications:

$$\frac{\partial \mathbf{L}}{\partial w_j}^T\mathbf{L} = \frac{\partial \mathbf{l}_j}{\partial w_j}^T\mathbf{l}_j \quad , \quad \mathbf{L}^T\frac{\partial \mathbf{L}}{\partial w_j} = \mathbf{l}_j^T\frac{\partial \mathbf{l}_j}{\partial w_j} \quad , \quad \frac{\partial \mathbf{L}}{\partial w_j}\mathbf{b} = \frac{\partial \mathbf{l}_j}{\partial w_j}\mathbf{b} \tag{6.5}$$

En remarquant que dans notre cas $\mathbf{b}$ est un vecteur constant, et en effectuant la substitution (6.3), il vient:

$$\frac{\partial \mathbf{y}}{\partial w_j} = -(\mathbf{L}^T\mathbf{L})^{-1}\underbrace{\left\{\left(\frac{\partial \mathbf{l}_j}{\partial w_j}^T\mathbf{l}_j + \mathbf{l}_j^T\frac{\partial \mathbf{l}_j}{\partial w_j}\right)\mathbf{y} - \frac{\partial \mathbf{l}_j}{\partial w_j}\mathbf{b}\right\}}_{\mathbf{N}_j} \tag{6.6}$$

Nous faisons de plus l'hypothèse que la covariance des points mesurés est isotrope et uniforme. La matrice $4N \times 4N$ $\Lambda_{\mathbf{x}}$ est alors scalaire, si bien que la covariance de $\mathbf{F} = \mathbf{y}$ est obtenue par:

$$\Lambda_{\mathbf{y}} = \sigma^2 \sum_j \frac{\partial \mathbf{y}}{\partial w_j} \frac{\partial \mathbf{y}}{\partial w_j}^T = \sigma^2 (\mathbf{L}^T\mathbf{L})^{-1} \left\{ \sum_j \mathbf{N}_j \mathbf{N}_j^T \right\} (\mathbf{L}^T\mathbf{L})^{-1T} \tag{6.7}$$

Lorsque les épipoles sont calculés par les formules (2.14), la covariance de ceux-ci est obtenue à partir de celle de $\mathbf{F}$ en utilisant de nouveau (6.1). En revanche, lorsque l'approximation (2.15) est utilisée, nous devons également considérer les épipoles comme solution d'un problème aux moindres carrés et utiliser la formule (6.4).

6.1.2 Le calcul de l'incertitude: cas non-linéaire

Nous présentons une méthode de calcul d'incertitude fondée sur le théorème des fonctions implicites [39]. Le but cette approche est de pouvoir calculer l'incertitude associée à des quantités obtenues par minimisation, sans avoir à tenir compte du processus de minimisation utilisé.

Considérons un critère $C(\mathbf{y}, \mathbf{m})$ dépendant du vecteur de données $\mathbf{m}$ de dimension N et du vecteur de paramètres à calculer $\mathbf{y}$ de dimension p. A chaque vecteur de données $\mathbf{m}_0$, on peut associer le vecteur de paramètres $\mathbf{y}_0$ obtenu en minimisant le critère C. Montrons que l'on définit ainsi localement une fonction ϕ si, et seulement si, le hessien $\mathbf{H}$ de C par rapport à $\mathbf{y}$ est inversible en $(\mathbf{y}_0, \mathbf{m}_0)$. La fonction ϕ est définie implicitement par la condition qu'au minimum, le gradient du critère par rapport à $\mathbf{y}$ est nul. Nous pouvons donc appliquer le théorème des fonctions implicites à la fonction $\mathbf{D}(\mathbf{y}, \mathbf{m}) = \frac{\partial C}{\partial \mathbf{y}}(\mathbf{f}, \mathbf{m})^T$ au point $(\mathbf{y}_0, \mathbf{m}_0)$, d'où la condition

$$\mathbf{H}(\mathbf{y}_0, \mathbf{m}_0) = \frac{\partial \mathbf{D}}{\partial \mathbf{y}}(\mathbf{y}_0, \mathbf{m}_0) \neq 0$$

Ceci nous donne de plus, que la différentielle de ϕ est obtenue par:

$$\mathrm{d}\phi = -\underbrace{\left(\frac{\partial \mathbf{D}}{\partial \mathbf{y}}\right)^{-1}}_{p \times p} \underbrace{\frac{\partial \mathbf{D}}{\partial \mathbf{m}}}_{p \times N} \qquad \text{avec} \quad \mathbf{D} = \underbrace{\left(\frac{\partial C}{\partial \mathbf{y}}\right)^T}_{p \times 1}$$

Dans les cas qui nous intéressent, nous cherchons à minimiser une somme de carrés ayant tous la même forme, donc le critère s'écrit:

$$C_T = \sum_{i=1}^{n} C_i^2 \qquad \text{où} \quad C_i = C(\mathbf{y}, \mathbf{m}_i)$$

Nous montrons que les calculs précédents peuvent être menés de manière assez simple, au moyen d'opérations répétitives ne faisant intervenir que des matrices de taille $p \times p$ au plus. Il vient en effet:

$$\mathbf{D} = \sum_{i=1}^{n} \mathbf{D}_i \qquad \text{avec} \quad \mathbf{D}_i = 2 \left(C_i \frac{\partial C_i}{\partial \mathbf{f}} \right)^T$$

$$\mathbf{H} = \sum_{i=1}^{n} \mathbf{H}_i \quad \text{avec} \quad \mathbf{H}_i = \frac{\partial \mathbf{D}_i}{\partial \mathbf{y}}$$

$$\mathbf{G} = \frac{\partial \mathbf{D}}{\partial \mathbf{m}} = \left[\begin{array}{ccc} \frac{\partial \mathbf{D}_1}{\partial \mathbf{m}_1} & \cdots & \frac{\partial \mathbf{D}_n}{\partial \mathbf{m}_n} \end{array} \right]$$

Supposons de plus que la variance des points $\mathbf{m}_i$ et $\mathbf{m}'_i$ soit isotrope et égale à σ. Nous avons alors:

$$\Lambda_{\mathbf{y}} = \sigma^2 \mathbf{H}^{-1} \mathbf{G} \mathbf{G}^T \mathbf{H}^{-1T} = \sigma^2 \mathbf{H}^{-1} \left\{ \sum_{i=1}^{n} \frac{\partial \mathbf{D}_i}{\partial \mathbf{m}_i} \frac{\partial \mathbf{D}_i}{\partial \mathbf{m}_i}^T \right\} \mathbf{H}^{-1T} \tag{6.8}$$

Nous pouvons voir immédiatement que la formule (6.8) a une forme similaire à la formule (6.7). Vérifions qu'elles sont effectivement équivalentes si nous nous plaçons dans le cas d'un critère linéaire. Dans ce cas:

$$C_i = \mathbf{l}_i^T \mathbf{y} - \mathbf{b}_i \quad \mathbf{D}_i = 2\mathbf{l}_i(\mathbf{l}_i^T \mathbf{y} - \mathbf{b}_i) \quad \mathbf{H}_i = 2\mathbf{l}_i \mathbf{l}_i^T$$

nous avons donc:

$$\mathbf{H} = 2\mathbf{L}^T \mathbf{L}$$

de plus, il est aisé de voir, après regroupement, que l'expression:

$$\frac{\partial \mathbf{D}_i}{\partial w_i} = 2 \left\{ \frac{\partial \mathbf{l}_i}{\partial w_i} (\mathbf{l}_i^T \mathbf{y} - \mathbf{b}_i) + \mathbf{l}_i \left(\frac{\partial \mathbf{l}_i}{\partial w_i}^T \mathbf{y} - \frac{\partial \mathbf{b}_i}{\partial w_i} \right) \right\}$$

est proportionnelle à $\mathbf{N}_i$ définie en (6.6), ce qui achève la vérification.

6.1.3 Une validation expérimentale

Nous commençons par présenter les résultats obtenus avec la première méthode de calcul de la covariance.

Quelques exemples Nous commençons par illustrer la validité des calculs de covariance présentés précédemment, à l'aide des deux mouvements définis au chapitre 3, table 3.1. Nous partons d'ensembles d'appariements non bruités, à partir desquels nous calculons la matrice de covariance des épipoles. La diagonalisation de cette matrice définie positive nous permet de tracer une ellipse à l'intérieur de laquelle les épipoles se trouvent, avec une probabilité fixée à un certain seuil (90% dans les représentations qui suivent). La direction des axes est donnée par les vecteurs propres. Le grand axe et le petit axe sont les racines carrées des valeurs propres[1]. D'autre part, nous effectuons 100 calculs d'épipoles en bruitant de manière différente les appariements avec un bruit gaussien de 2 pixels, ce qui nous donne un nuage de points que nous représentons. Les figures 6.1 et 6.2 montrent les ellipses d'incertitude et les épipoles obtenus dans la première image, avec le critère linéaire, pour les mouvements 1 et 2, respectivement, le carré représentant les bords de l'image, et le point

[1]Pour des détails, voir par exemple [176] pp 51-54

triangulaire l'épipole correct. Il est clair que l'estimation de l'épipole est biaisée. Pour le second mouvement, le critère linéaire rapproche les épipoles, conformément à notre analyse effectuée en 4.1.3. Le éléments représentés figures 6.3 et 6.4 sont identiques, à l'exception du fait que le critère non-linéaire a été utilisé. Le biais est mieux corrigé, mais la covariance est plus grande, dû à quelques faux minima locaux. Nous observons que l'ellipse prévoit encore de manière correcte l'ordre de grandeur de l'incertitude, qui est beaucoup plus importante pour le second mouvement, comme nous l'avons déjà remarqué au chapitre 3. La direction suivant laquelle les différents épipoles sont répartis est également très bien prédite. Nous remarquons que pour le second mouvement, la répartition des épipoles n'est pas du tout isotrope: ceux-ci sont a peu près répartis le long d'une droite[2], ce qui est un fait qui avait déjà été noté et expliqué dans le contexte de l'étude du flot optique par Scott [136] et Maybank [102]. L'ellipse d'incertitude associée est effectivement très allongée, puisque le rapport des deux valeurs propres est supérieur à 10^3.

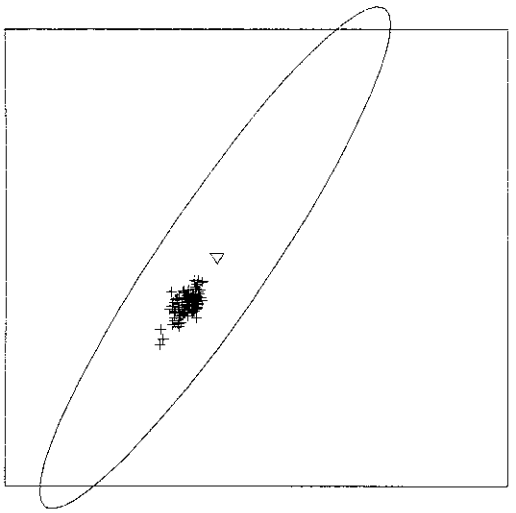


Fig. 6.1: Ellipse d'incertitude et épipoles bruités (critère linéaire), pour le mouvement 1

Des statistiques Nous avons ensuite comparé les covariances obtenues et les covariances prédites au moyen d'un test statistique. Pour chacun des 200 déplacements, nous avons estimé les épipoles en variant le bruit 20 fois, celui-ci étant gaussien, de déviation 1 pixel, ce qui nous a permis de calculer une covariance. Nous avons alors classé les déplacements par ordre de covariance croissante. Les abscisses sur la figure 6.5 représentent le rang obtenu lors de ce classement. Les ordonnées sont une fonction croissante[3] de la covariance

[2] en fait une courbe algébrique

[3] Nous avons pris un logarithme afin de pouvoir représenter toutes les valeurs de manière lisible

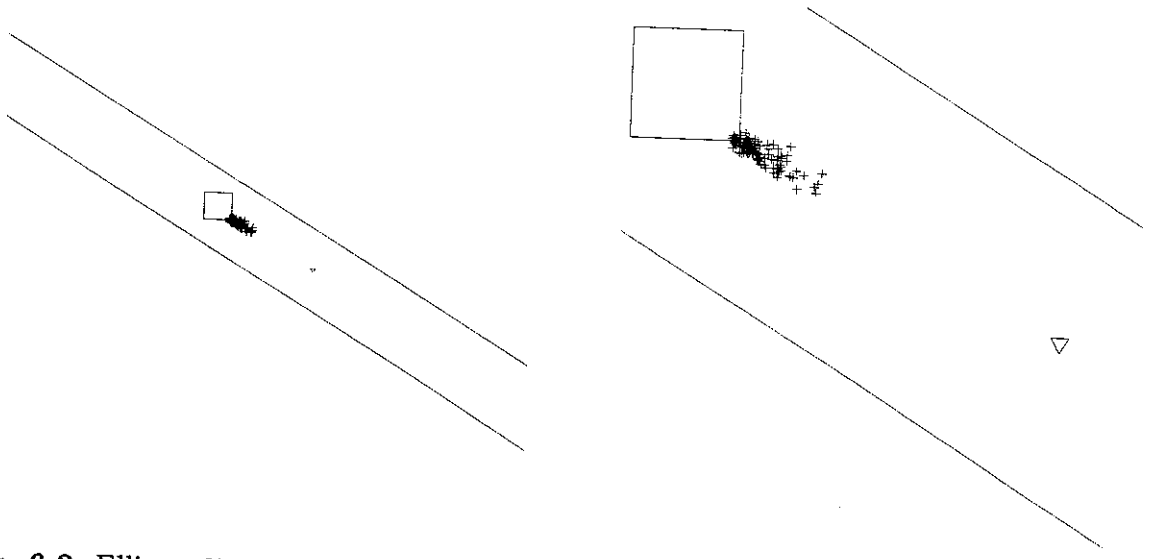


Fig. 6.2: Ellipse d'incertitude et épipoles bruités (critère linéaire), pour le mouvement 2

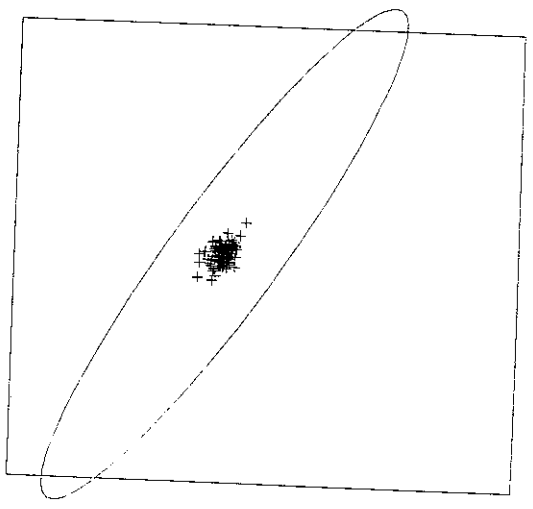


Fig. 6.3: Ellipse d'incertitude et épipoles bruités (critère non-linéaire), pour le mouvement 1

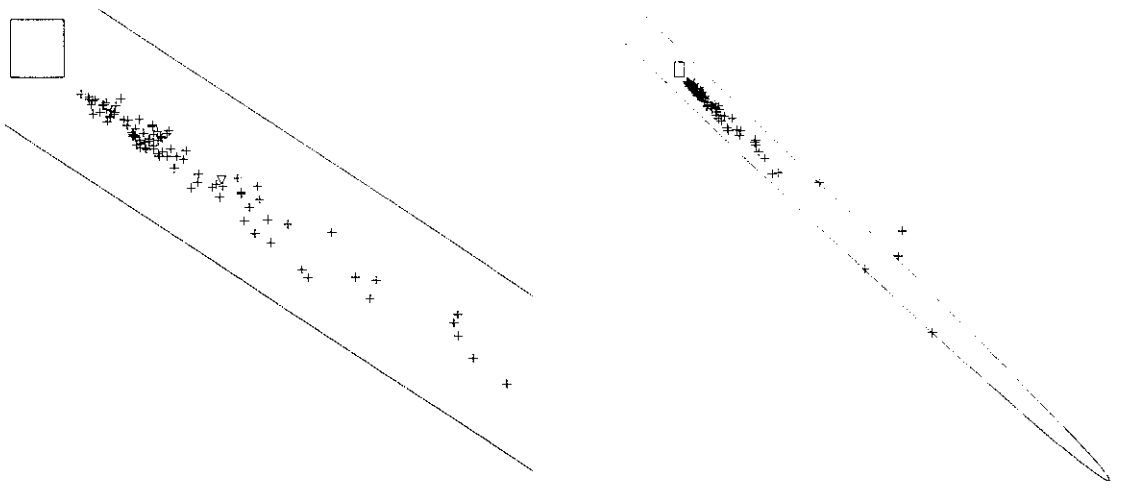


Fig. 6.4: Ellipse d'incertitude et épipoles bruités (critère non-linéaire), pour le mouvement 2

prédite par la méthode décrite en 6.1.1. Nous avons, pour vérifier la robustesse au bruit de ces estimations, effectué celles-ci, d'une part à partir des points non-bruités (points clairs sur la figure), d'autre part à partir des points bruités (points noirs sur la figure), et superposé les deux résultats. Nous pouvons conclure:

- Il y a une forte corrélation entre la covariance prédite et celle effectivement obtenue sur un échantillon.
- Cette corrélation est plus forte lorsque la covariance est calculée à partir des points non-bruités, cependant elle demeure tout à fait significative même lorsque le calcul est effectué à partir des points bruités.

Le calcul de covariance est donc un outil qui nous permet de caractériser de manière probabiliste la stabilité d'une façon assez fiable.

6.2 Surface critique et ambiguïté

Le but de cette section est d'introduire une notion importante, celle de surface critique, et surtout de présenter des algorithmes qui rendent cette notion opérationnelle. A cet effet, nous montrons comment, à partir d'appariements de points, il est possible de retrouver la distance à une surface critique approximante au moyen de plusieurs algorithmes linéaires et non-linéaires.

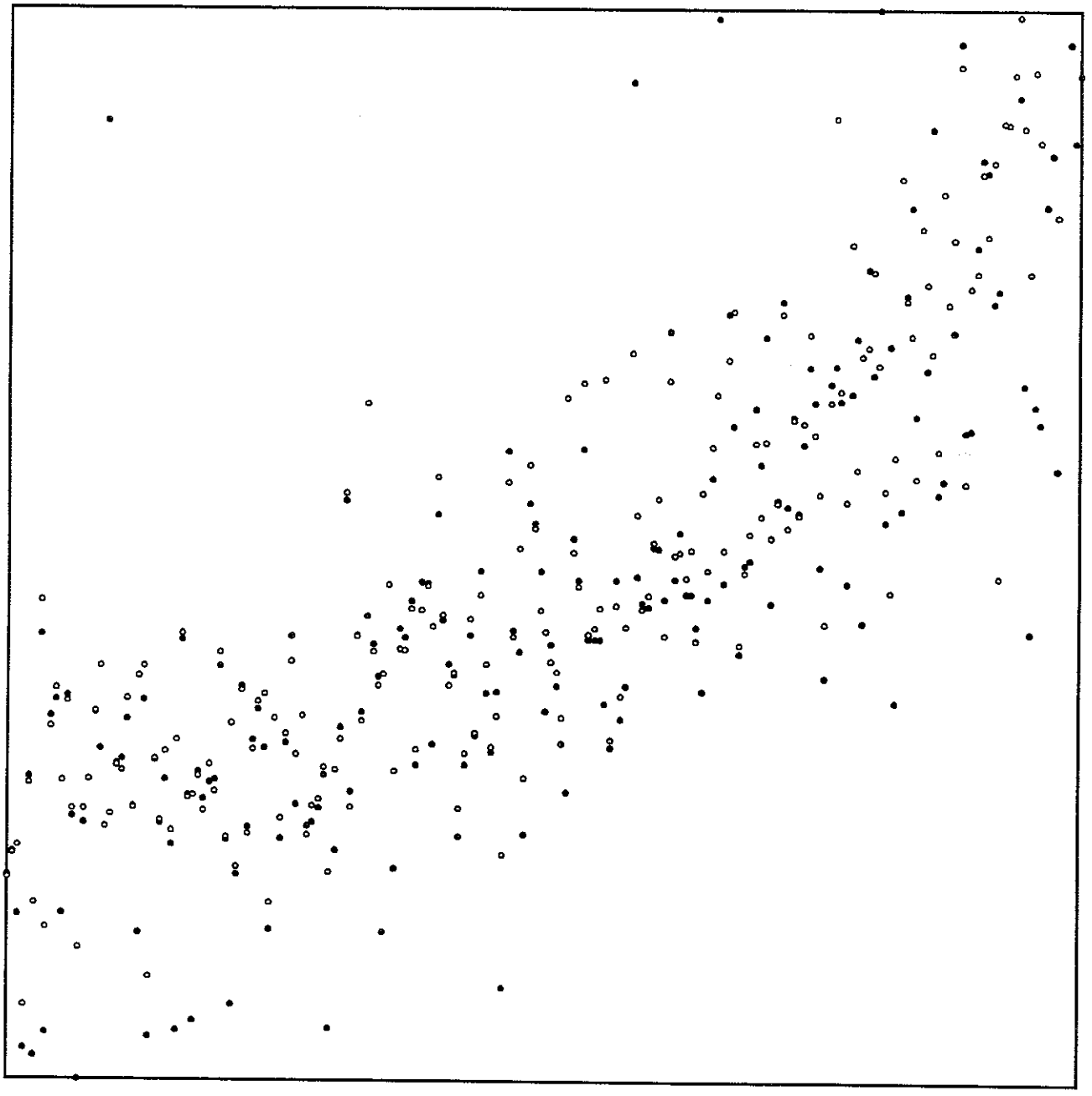


Fig. 6.5: Corrélation entre covariances calculées et covariances prédites.

6.2.1 La surface critique

Le problème du calcul de la matrice fondamentale à partir d'appariements de points, tel que nous l'avons posé au chapitre précédent, n'admet pas nécessairement une solution unique. C'est la difficulté posée par *l'ambiguité*: l'existence de plusieurs matrices fondamentales compatibles avec la contrainte de base: $\mathbf{q}'^T\mathbf{F}\mathbf{q} = 0$. Chacune de ces matrices fondamentales définit alors un mouvement de la caméra tel que les paires de vues obtenues à partir des mouvements respectifs ne sont pas distinguables Il est clair que cette situation ne peut pas se produire si nous disposons d'un nombre arbitraire de points en configuration générale, mais seulement pour certaines configurations de points dans l'espace

Deux ensembles de N points $\mathbf{m}^{(i)}$ et $\mathbf{m}'^{(i)}$ appartenant à deux projections sont *ambigüs* si ils peuvent être obtenus à partir de deux mouvements différents, tout en contenant un nombre arbitraire de points. Cette dernière précision est nécessaire, car il est clair que si N est faible plusieurs solutions peuvent être trouvées[4]. Cette remarque étant faite, nous abandonnons l'exposant $^{(i)}$ pour simplifier l'écriture. Il existe donc deux mouvements $(\mathbf{R}_1, \mathbf{t}_1)$ et $(\mathbf{R}_2, \mathbf{t}_2)$, tels que si $\mathbf{m}$ (resp. $\mathbf{m}_1$, $\mathbf{m}_2$) représente les coordonnées normalisées du point vu par la caméra dans la position initiale (resp. après les mouvements 1 et 2), exprimées dans le repère caméra initial (resp. après les mouvements 1 et 2):

$$\begin{aligned} \lambda'_1\mathbf{m}_1 &= \mathbf{R}_1\lambda_1\mathbf{m} + \mathbf{t}_1 \\ \lambda'_2\mathbf{m}_2 &= \mathbf{R}_2\lambda_2\mathbf{m} + \mathbf{t}_2 \end{aligned}$$

avec, par définition de l'ambiguïté, $\mathbf{m}_1 = \mathbf{m}_2 = \mathbf{m}'$. Nous en déduisons que:

$$\mathbf{R}_2\lambda_2\mathbf{m} + \mathbf{t}_2 = (\lambda'_1\lambda_1/\lambda'_2)\mathbf{R}_1\mathbf{m} + (\lambda'_1/\lambda'_2)\mathbf{t}_1$$

donc les trois vecteurs $\mathbf{R}_2\lambda_2\mathbf{m} + \mathbf{t}_2$, $\mathbf{t}_1$, et $\mathbf{R}_1\mathbf{m}$ sont coplanaires. En exprimant leur produit mixte, et en utilisant la matrice essentielle $\mathbf{E}_1 = \mathbf{T}_1\mathbf{R}_1$, il vient la relation:

$$(\mathbf{R}_2\lambda_2\mathbf{m} + \mathbf{t}_2)^T\mathbf{E}_1\mathbf{m} = 0 \tag{6.9}$$

Cette relation est vérifiée si, et seulement si:

$$\lambda_2 = -\frac{\mathbf{t}_2^T\mathbf{E}_1\mathbf{m}}{(\mathbf{R}_2\mathbf{m})^T\mathbf{E}_1\mathbf{m}}$$

En posant $\mathbf{M}_2 = \lambda_2\mathbf{m}$, nous trouvons un point de l'espace ayant pour image $\mathbf{m}$ par la caméra prise dans sa position initiale, et $\mathbf{m}'$ par la caméra après le mouvement 2. Si nous échangeons les indices 1 et 2, nous voyons qu'en posant $\mathbf{M}_1 = \lambda_1\mathbf{m}$, avec:

$$\lambda_1 = -\frac{\mathbf{t}_1^T\mathbf{E}_2\mathbf{m}}{(\mathbf{R}_1\mathbf{m})^T\mathbf{E}_2\mathbf{m}}$$

nous trouvons un second point de l'espace ayant pour image $\mathbf{m}$ par la caméra prise dans sa position initiale, et $\mathbf{m}'$ par la caméra après le mouvement 1.

[4] Il a été montré [42] qu'il y a 10 solutions si $N = 5$

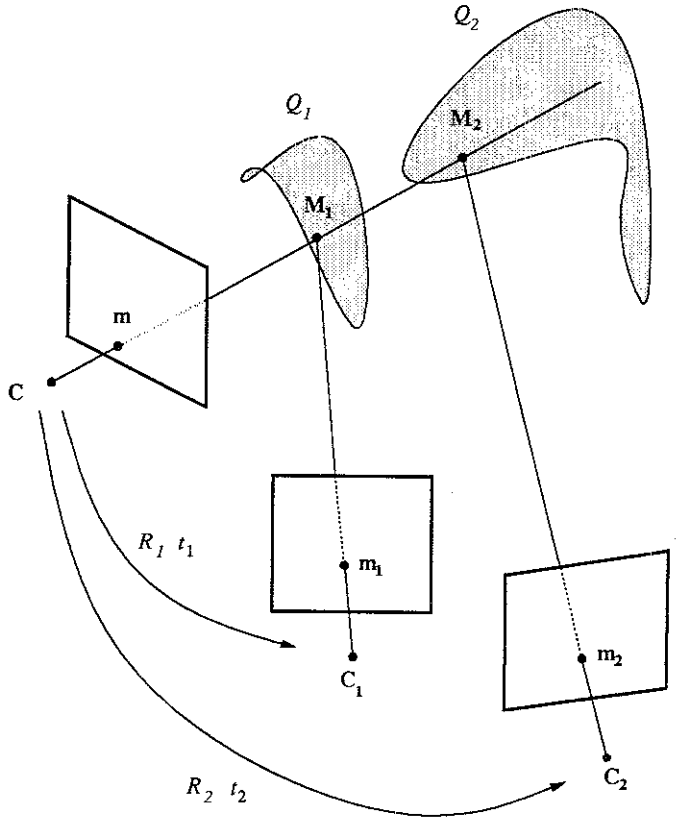


Fig. 6.6: Construction des surfaces critiques

Le lieu des points $\mathbf{M}_2$ est donc la surface $\mathcal{Q}_2$ d'équation:

$$(\mathbf{R}_2\mathbf{M} + \mathbf{t}_2)^T\mathbf{E}_1\mathbf{M} = 0 \tag{6.10}$$

En échangeant les indices 1 et 2, il vient le lieu $\mathcal{Q}_1$ des points $\mathbf{M}_1$:

$$(\mathbf{R}_1\mathbf{M} + \mathbf{t}_1)^T\mathbf{E}_2\mathbf{M} = 0 \tag{6.11}$$

Il s'agit de deux quadriques passant par le centre optique initial et celui après déplacement, dites de Maybank [101], qui sont des *surfaces critiques*, au sens où il n'est pas possible de savoir si l'on regarde des points de $\mathcal{Q}_1$ en ayant effectué le mouvement 1, ou[5] des points de $\mathcal{Q}_2$ en ayant effectué le mouvement 2. La situation est schématisée figure 6.6. Il est à noter que l'image du point $\mathbf{M}_2$ (resp. $\mathbf{M}_1$) obtenue après le mouvement 1 (resp. mouvement 2) se trouve sur la même droite épipolaire que le point $\mathbf{m}_1$ (resp. $\mathbf{m}_2$).

Une autre manière, plus rapide mais faisant apparâtre moins clairement le rôle des surfaces critiques, d'aboutir à l'équation (6.10) est de substituer dans l'équation de Longuet-Higgins pour le mouvement 1, $\mathbf{M}_1'^T\mathbf{E}_1\mathbf{M} = 0$, l'action du mouvement 2, $\mathbf{M}_2' = \mathbf{R}_2\mathbf{M} + \mathbf{t}_2$ en utilisant le fait qu'à cause de l'ambiguïté, $\mathbf{M}_1' = \lambda\mathbf{M}_2'$.

Les surfaces critiques étaient connues des photogrammètres autrichiens du début du siècle, sous le nom de "*gefährliche Flächen*". Elles ont ensuite été très étudiées en vision par ordinateur, d'un point de vue théorique, aussi bien dans le cas du flot optique [101, 63, 123] que dans le cas général de deux images obtenues par un déplacement fini [91, 105, 64, 106]. Nous nous proposons, dans un cadre de détermination de la matrice fondamentale, de mettre en évidence quelques conséquences *pratiques* de l'existence de ces surfaces.

6.2.2 La transformation quadratique entre deux images ambigües

Considérons deux images ambigües. D'après ce que nous venons de voir, il existe deux matrices fondamentales distinctes $\mathbf{F}_1$ et $\mathbf{F}_2$ telles que pour tous les couples $(\mathbf{q}, \mathbf{q}')$ de points correspondants,

$$\mathbf{q}'^T\mathbf{F}_1\mathbf{q} = 0 \quad \text{et} \quad \mathbf{q}'^T\mathbf{F}_2\mathbf{q} = 0$$

cette équation exprime projectivement que le point $\mathbf{q}'$ appartient aux deux droites $\mathbf{F}_1\mathbf{q}$ et $\mathbf{F}_2\mathbf{q}$, donc il est égal à leur intersection:

$$\mathbf{q}' = \mathbf{F}_1\mathbf{q} \wedge \mathbf{F}_2\mathbf{q} \tag{6.12}$$

Cette équation définit dans le cas général une *transformation quadratique* entre les coordonnées des points des deux images. Par définition, ce sont des transformations du plan projectif, polynomiales homogènes de degré 2 en chacune des composantes projectives, qui sont inversibles et dont l'inverse s'exprime également à partir de polynômes de degré 2. Il est

[5] L'interprétation de l'ambiguïté selon laquelle les deux mouvements conduisent à des images indiscernables n'est pas correcte: si nous regardons les points de $\mathcal{Q}_1$ après avoir effectué le mouvement 2, nous verrons des images différentes de celles obtenues en regardant $\mathcal{Q}_1$ après le mouvement 1.

clair, au vu de l'équation (6.12) que $\mathbf{q}'$ est en général de degré 2 en les coordonnées de $\mathbf{q}$. Une exception peut se produire, le cas dans lequel l'équation (6.12) définit une transformation linéaire en les coordonnées de $\mathbf{q}$, auquel cas nous sommes en présence d'une transformation homographique, la surface critique étant alors un simple plan. Un simple calcul permet de vérifier que la transformation inverse de (6.12) est donnée par:

$$\mathbf{q} = \mathbf{F}_1^T\mathbf{q}' \wedge \mathbf{F}_2^T\mathbf{q}'$$

ce résultat est retrouvé aussi en échangeant les rôles des deux images dans le raisonnement initial.

L'intérêt du calcul de la transformation quadratique est de nous permettre de voir uniquement à partir des images si deux ensembles de points observés dans celles-ci sont proches en projection, d'une surfaces critique de l'espace, de la même manière que l'estimation de l'homographie, au chapitre précédent, nous permet de voir si les points des deux images sont proches, en projection, d'un plan de l'espace.

6.2.3 Calcul direct de la transformation quadratique

Notre première approche a été, étant donné un ensemble d'appariements de points entre deux images $\mathbf{q}_i \leftrightarrow \mathbf{q}'_i$, d'essayer de calculer directement la transformation quadratique Φ telle que les points $\Phi(\mathbf{q}_i)$ soient les plus proches possible des points $\mathbf{q}'_i$. Notre problème s'apparente donc à une minimisation portant sur toutes les transformations quadratiques possibles. Par conséquent, nous commençons par étudier les paramétrisations d'une transformation quadratique générale.

Paramétrisations d'une transformation quadratique générale L'exemple le plus simple de transformation quadratique est la transformation réciproque, définie par:

$$\Phi_0(\mathbf{x}) = (x_2x_3, x_3x_1, x_1x_2)^T$$

A partir de cette définition, nous pouvons voir d'une part que Φ_0 est définie en tout point de $\mathcal{P}^2$, à l'exception des points de la base canonique $\mathbf{e}_1 = (1,0,0)^T$, $\mathbf{e}_2 = (0,1,0)^T$ et $\mathbf{e}_3 = (0,0,1)^T$, qui sont dits points fondamentaux de Φ_0, d'autre part que sur son ensemble de définition, elle est bien inversible et birationnelle puisqu'elle est sa propre inverse. Les points fondamentaux de Φ_0^{-1} sont donc confondus avec ceux de Φ_0, ce qui ne se produit pas en général: une transformation quadratique Φ générale est en effet caractérisée par le choix des trois points fondamentaux directs $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$, et par celui des trois points fondamentaux inverses $\mathbf{f}'_1$, $\mathbf{f}'_2$, $\mathbf{f}'_3$, qui sont ceux de Φ^{-1}. Elle peut donc se décomposer sous la forme:

$$\Phi = \mathbf{A}\Phi_0\mathbf{B} \tag{6.13}$$

où $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices d'homographies, qui peuvent être interprétées en termes de changement de base:

$$\begin{array}{llll} \mathbf{A}\mathbf{e}_1 = \mathbf{f}'_1 & \mathbf{A}\mathbf{e}_2 = \mathbf{f}'_2 & \mathbf{A}\mathbf{e}_3 = \mathbf{f}'_3 & \mathbf{A}\mathbf{e}_4 = \mathbf{f}'_4 \\ \mathbf{B}\mathbf{f}_1 = \mathbf{e}_1 & \mathbf{A}\mathbf{f}_2 = \mathbf{e}_2 & \mathbf{A}\mathbf{f}_3 = \mathbf{e}_3 & \mathbf{A}\mathbf{f}_4 = \mathbf{e}_4 \end{array} \tag{6.14}$$

S'agissant d'un changement de base projective, les coordonnées de l'image du quatrième point de la base est spécifié arbitrairement dans chacune des deux images par les points $\mathbf{f}_4$ et $\mathbf{f}'_4$. Une fois que ceux-ci sont choisis, les matrices $\mathbf{A}$ et $\mathbf{B}$ sont obtenues par:

$$\mathbf{B}^{-1} = [\rho_1\mathbf{f}_1\ \rho_2\mathbf{f}_2\ \rho_3\mathbf{f}_3] \qquad \text{avec} \quad \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} = [\mathbf{f}_1\ \mathbf{f}_2\ \mathbf{f}_3]^{-1}\mathbf{f}_4 \tag{6.15}$$

$$\mathbf{A} = [\rho'_1\mathbf{f}'_1\ \rho'_2\mathbf{f}'_2\ \rho'_3\mathbf{f}'_3] \qquad \text{avec} \quad \begin{bmatrix} \rho'_1 \\ \rho'_2 \\ \rho'_3 \end{bmatrix} = [\mathbf{f}'_1\ \mathbf{f}'_2\ \mathbf{f}'_3]^{-1}\mathbf{f}'_4 \tag{6.16}$$

La décomposition (6.13) assure alors que Φ ainsi définie est bien une transformation quadratique, d'inverse $\Phi^{-1} = \mathbf{B}^{-1}\Phi\mathbf{A}^{-1}$. Les paramètres dont dépend Φ sont donc au nombre de douze: les six coordonnées projectives des trois points fondamentaux directs et les six coordonnées projectives des trois points fondamentaux réciproques, ce qui fait apparaître que la décomposition (6.13) de Φ n'est pas unique, puisque celle-ci dépend de dix-huit paramètres qui sont les coefficients des deux homographies. En retranchant les deux facteurs d'échelle ainsi que les coordonnées projectives de $\mathbf{f}_4$ et de $\mathbf{f}'_4$, nous retrouvons bien douze paramètres indépendants.

Cette paramétrisation fait apparaître le caractère hautement non-linéaire du problème de l'estimation de la transformation quadratique, que nous formulons en nous limitant aux cas des points fondamentaux finis, et en prenant $\mathbf{f}_4 = \mathbf{f}'_4 = \mathbf{e}_4 = (1,1,1)^T$:

$$\min_{\{f_{jx},f_{jy},f'_{jx},f'_{jy}\}\ j=1,2,3} \sum_i \{d^2(\mathbf{q}'_i, \Phi(\mathbf{q}_i)) + d^2(\mathbf{q}_i, \Phi^{-1}(\mathbf{q}'_i))\} \tag{6.17}$$

Nous minimisons ainsi de manière symétrique une quantité significative dans l'image, en fonction du nombre exact de paramètres indépendants permettant d'exprimer la structure de la relation recherchée. Afin de fournir à ce processus d'initialisation non-linéaire un point de recherche initial, nous allons présenter une méthode d'estimation linéaire.

Une méthode d'estimation linéaire En écrivant juste le fait que Φ est une transformation du plan projectif, polynomiale homogènes de degré 2 en chacune des composantes projectives, il vient la relation:

$$\begin{cases} \Phi(x_1,x_2,x_3)_1 = a^1_{11}x_1^2 + a^1_{22}x_2^2 + a^1_{33}x_3^2 + a^1_{12}x_1x_2 + a^1_{23}x_2x_3 + a^1_{31}x_3x_1 \\ \Phi(x_1,x_2,x_3)_2 = a^2_{11}x_1^2 + a^2_{22}x_2^2 + a^2_{33}x_3^2 + a^2_{12}x_1x_2 + a^2_{23}x_2x_3 + a^2_{31}x_3x_1 \\ \Phi(x_1,x_2,x_3)_3 = a^3_{11}x_1^2 + a^3_{22}x_2^2 + a^3_{33}x_3^2 + a^3_{12}x_1x_2 + a^3_{23}x_2x_3 + a^3_{31}x_3x_1 \end{cases} \tag{6.18}$$

Comme les équations (6.18) sont linéaires en les dix huit coefficients a^k_{ij}, ceux-ci peuvent être estimés par un moindres carrés linéaire, pourvu que nous disposions de plus de neuf appariements, puisque chaque appariement nous fournit deux équations qui sont les relations

de proportionnalité entre $\mathbf{q}$ et $\mathbf{q}'$. Une fois que les coefficients a_{ij}^k sont calculés, les points fondamentaux directs sont obtenus comme solutions projectives du système:

$$\begin{cases} \Phi(x_1, x_2, x_3)_1 = 0 \\ \Phi(x_1, x_2, x_3)_2 = 0 \\ \Phi(x_1, x_2, x_3)_3 = 0 \end{cases} \tag{6.19}$$

Montrons que pour que ce dernier système soit compatible, les coefficients a_{ij}^k doivent satisfaire trois relations polynomiales. Considérons d'abord le cas $x_3 = 0$. Si $x_1 = 0$ ou si $x_2 = 0$, le résultat est évident car il traduit juste l'annulation de trois coefficients. Dans l'alternative, nous sommes en présence de trois équations du second degré, homogènes en x_1 et x_2. La condition de compatibilité équivaut donc à l'annulation de trois résultants. Considérons maintenant le cas $x_3 \neq 0$: nous exprimons le fait que l'intersection de trois coniques doit donner trois solutions réelles. Les deux premières coniques s'intersectent en quatre points. Parmi ceux-ci, trois d'entre eux doivent également vérifier l'équation de la troisième conique, d'où trois relations. Nous pouvons noter qu'en inversant les rôles des deux images, nous calculons de la même manière des coefficients $a_{ij}^{'k}$, et que les conditions de compatibilité du système:

$$\begin{cases} \Phi'(x'_1, x'_2, x'_3)_1 = 0 \\ \Phi'(x'_1, x'_2, x'_3)_2 = 0 \\ \Phi(x'_1, x'_2, x'_3)_3 = 0 \end{cases} \tag{6.20}$$

permettant d'obtenir trois points fondamentaux donnant également trois relations polynomiales, nous retrouvons bien douze paramètres indépendants.

Lorsque nous ne sommes pas en présence de deux images dont l'une est une transformée quadratique exacte de l'autre, les six conditions de compatibilité permettant d'obtenir une solution exacte de (6.19) et (6.20) ne sont pas satisfaites[6]. Voyons le cas des points fondamentaux directs, celui des points fondamentaux réciproques étant analogue. Nous devons estimer une solution approchée du système (6.19). La technique est d'effectuer une optimisation non-linéaire pour affiner chacun des points fondamentaux $\mathbf{x}$ en minimisant:

$$\min_{x_1,x_2,x_3} \{\Phi^2(x_1, x_2, x_3)_1 + \Phi^2(x_1, x_2, x_3)_2 + \Phi^2(x_1, x_2, x_3)_3\}$$

Nous initialisons celle-ci en résolvant algébriquement les trois paires d'équations qui sont extraites du système (6.19), pour obtenir trois points suffisamment distincts qui vérifient chacun partiellement les équations de ce système.

Une fois que nous avons calculé une estimée des points fondamentaux directs et réciproques indépendamment, nous devons les apparier deux à deux, suivant (6.14). Compte tenu de la permutation sur les $\mathbf{e}_i$ qui rend six groupes de configurations équivalents, il nous reste six configurations distinctes, parmi laquelle nous sélectionnons celle qui minimise le critère (6.17). Nous obtenons ainsi les douze paramètres initiaux.

[6] La situation est similaire à celle obtenue lors du calcul de la matrice fondamentale par le critère linéaire, vis-à-vis de l'importante contrainte de rang. La surparamétrisation nous donne une solution linéaire, mais ne permet pas d'exprimer les contraintes liées à l'objet calculé.

Résultats Nous avons trouvé que la méthode d'estimation précédente est assez précise lorsque nous sommes en présence de deux ensembles de points générés de sorte que l'un est bien approximé par une transformée quadratique de l'autre, par exemple en projetant dans les images des points pris au voisinage de surfaces critiques. Des résultats sont regroupés avec ceux de la méthode suivante. Cependant, si nous ne nous plaçons pas dans des cas favorables, la méthode ne converge pas en général. Ceci n'est pas très surprenant, dans la mesure où elle possède des caractéristiques proches de celle de la méthode d'estimation de la matrice fondamentale, avec une paramétrisation encore plus complexe et davantage de paramètres. Il pourrait sembler qu'estimer la transformation quadratique pour caractériser la stabilité de la matrice fondamentale soit irréaliste, puisque l'estimation de la transformation quadratique est elle-même plus instable que celle de la matrice fondamentale. Cependant, nous pouvons utiliser la matrice fondamentale elle-même à cette fin.

6.2.4 Calcul de la transformation quadratique à partir d'une matrice fondamentale

L'idée est de partir d'une matrice fondamentale $\mathbf{F}_1$, et d'estimer la matrice fondamentale $\mathbf{F}_2$ telle que $\mathbf{F}_1$ et $\mathbf{F}_2$ ensemble définissent une transformation quadratique suivant (6.12). Bien que nous n'ayons pas d'unicité de la décomposition sous la forme (6.12), puisque celle-ci dépend de quatorze paramètres (sept pour chaque matrice fondamentale), elle nous assure cependant d'obtenir une transformation quadratique. Si, étant donné $\mathbf{F}_1$, nous trouvons une matrice $\mathbf{F}_2$ minimisant le critère:

$$\min_{\mathbf{F}_2} \sum_i \{d^2(\mathbf{q}'_i, \mathbf{F}_1\mathbf{q}_i \wedge \mathbf{F}_2\mathbf{q}_i) + d^2(\mathbf{q}_i, \mathbf{F}_1^T\mathbf{q}'_i \wedge \mathbf{F}_2^T\mathbf{q}'_i)\} \tag{6.21}$$

et donnant le résidu $r_{\mathbf{F}_1}$, alors nous savons que ce résidu constitue un majorant théorique de celui obtenu minimisant le critère (6.17). Autrement dit, étant donné $\mathbf{F}_1$, la minimisation de (6.21) nous permet d'obtenir un majorant de la distance projetée des ensembles de points $(\mathbf{q}, \mathbf{q}')$ à la surface critique. Cette méthode de calcul, contrairement à la précédente, se révèle efficace, le résidu $r_{\mathbf{F}_1}$ étant presque toujours meilleur que celui obtenu pratiquement à partir de (6.17), même en utilisant les deux matrices $\mathbf{F}_1$ et $\mathbf{F}_2$ pour initialiser la recherche des points fondamentaux. Elle est très proche de celle permettant de calculer la première matrice fondamentale. Nous avons par conséquent essayé d'utiliser les deux paramétrisations introduites en 4.3. L'initialisation du critère non-linéaire (6.21) est effectuée par une recherche de $\mathbf{F}_2$ au moyen des équations linéaires:

$$\begin{cases} \mathbf{W}_1\mathbf{f} = 0 \\ \mathbf{W}_2\mathbf{f} = 0 \end{cases} \tag{6.22}$$

dans lesquelles nous avons posé

$$\mathbf{W}_1 = \begin{pmatrix} F_{21}^{(1)}u^2u' + F_{22}^{(1)}vuu' + F_{23}^{(1)}uu' \\ F_{23}^{(1)}vu' + F_{21}^{(1)}uvu' + F_{22}^{(1)}v^2u' \\ F_{23}^{(1)}u' + F_{21}^{(1)}uu' + F_{22}^{(1)}vu' \\ -F_{33}^{(1)}u - F_{31}^{(1)}u^2 - F_{11}^{(1)}u^2u' - F_{12}^{(1)}vuu' - F_{13}^{(1)}uu' - F_{32}^{(1)}vu \\ -F_{32}^{(1)}v^2 - F_{33}^{(1)}v - F_{12}^{(1)}v^2u' - F_{11}^{(1)}uvu' - F_{13}^{(1)}vu' - F_{31}^{(1)}uv \\ -F_{32}^{(1)}v - F_{31}^{(1)}u - F_{33}^{(1)} - F_{11}^{(1)}uu' - F_{12}^{(1)}vu' - F_{13}^{(1)}u' \\ F_{21}^{(1)}u^2 + F_{23}^{(1)}u + F_{22}^{(1)}vu \\ F_{23}^{(1)}v + F_{22}^{(1)}v^2 + F_{21}^{(1)}uv \\ F_{21}^{(1)}u + F_{22}^{(1)}v + F_{23}^{(1)} \end{pmatrix}^T$$

$$\mathbf{W}_2 = \begin{pmatrix} F_{31}^{(1)}u^2 + F_{33}^{(1)}u + F_{32}^{(1)}vu + F_{21}^{(1)}u^2v' + F_{22}^{(1)}vuv' + F_{23}^{(1)}uv' \\ F_{32}^{(1)}v^2 + F_{33}^{(1)}v + F_{31}^{(1)}uv + F_{21}^{(1)}uvv' + F_{22}^{(1)}v^2v' + F_{23}^{(1)}vv' \\ F_{32}^{(1)}v + F_{23}^{(1)}v' + F_{31}^{(1)}u + F_{33}^{(1)} + F_{21}^{(1)}uv' + F_{22}^{(1)}vv' \\ -F_{11}^{(1)}u^2v' - F_{12}^{(1)}vuv' - F_{13}^{(1)}uv' \\ -F_{11}^{(1)}uvv' - F_{12}^{(1)}v^2v' - F_{13}^{(1)}vv' \\ -F_{13}^{(1)}v' - F_{11}^{(1)}uv' - F_{12}^{(1)}vv' \\ -F_{11}^{(1)}u^2 - F_{13}^{(1)}u - F_{12}^{(1)}vu \\ -F_{13}^{(1)}v - F_{12}^{(1)}v^2 - F_{11}^{(1)}uv \\ -F_{12}^{(1)}v - F_{13}^{(1)} - F_{11}^{(1)}u \end{pmatrix}^T$$

$$\mathbf{f} = [F_{11}^{(2)}, F_{12}^{(2)}, F_{13}^{(2)}, F_{21}^{(2)}, F_{22}^{(2)}, F_{23}^{(2)}, F_{31}^{(2)}, F_{32}^{(2)}, F_{33}^{(2)}]^T$$

Notons un détail associé à la méthode des valeurs propres pour résoudre le moindres carrés linéaire: nous devons considérer non pas le vecteur propre associé à la plus petite valeur propre, qui représente toujours $\mathbf{F}_1$, mais le suivant.

6.2.5 Exemples numériques

Dans le premier exemple, nous partons d'une transformation quadratique définie en fixant les six points fondamentaux. Nous avons ensuite choisi choisi des points aléatoirement dans la première image, et leur avons appliqué cette transformation quadratique afin de d'obtenir les coordonnées des points correspondants dans la seconde image. Les appariements ayant été générés, il est possible de calculer une matrice fondamentale à partir de chaque paire de points fondamentaux en déterminant les trois paramètres de l'homographie par la méthode de moindres carrés présentée en 2.2.2, les deux épipoles étant les points fondamentaux, nous obtenons ainsi les matrices $\mathbf{F}_1$, $\mathbf{F}_2$ et $\mathbf{F}_3$ qui seront utilisées par la suite. La table 6.1 montre les points fondamentaux directs obtenus par différentes méthodes, à deux niveaux de bruit. Ces résultats sont typiques.

- **TQLIN** désigne la méthode linéaire d'estimation des points fondamentaux décrite par (6.18) et (6.19). En ce qui concerne la localisation des points fondamentaux, elle

bruit (pixel)	méthode	coordonnées des points fondamentaux								résidu
		$\mathbf{f}_1$		$\mathbf{f}_2$		$\mathbf{f}_3$		$\mathbf{f}'_1$		
0		1000	500	0	400	600	0	900	800	
0.2	**TQLIN**	987.3	494.5	-0.3	400.3	598.3	1.4	911.3	810.4	1.36
	TQ	996.4	498.0	-0.8	400.6	599.6	0.4	885.4	787.0	0.30
	F2LIN-INIT			442.1	32.5	429.2	14.0			1.5 0.6
	F2-INIT	1016.2	502.7	-0.8	400.6	614.6	-15.8	905.3	805.7	0.32 0.3
	F2-F			-0.8	400.6	591.7	6.4			0.52
	TQ-F2-F	996.4	498.0	-0.8	400.6	599.6	0.4	885.4	787.0	0.30
1.0	**TQLIN**	766.4	422.1	6.0	399.1	640.4	-71.3	876.6	775.5	124.92
	TQ	890.5	451.7	43.4	394.1	958.0	-326.5	16099.6	13642.2	27.38
	F2LIN-INIT	605.9	23.1			415.3	-19.8	282.2	147.5	13 18
	F2-INIT	979.4	490.5	-4.3	403.3	603.5	-3.3	856.1	757.9	1.6 1.6
	F2-F			-2.6	402.3	588.47	6.8			14
	TQ-F2-F	981.9	490.0	-4.3	403.4	598.2	1.9	833.0	739.9	1.54

Tableau 6.1: Résultat de détermination de transformation quadratique: le cas exact

donne des résultats corrects, même s'ils ne sont pas les plus précis. Ceci est dû au fait que dans cet exemple, les données sont *effectivement* obtenues par transformation quadratique. En revanche la distance aux surfaces critiques, mesurée par le critère (6.17), significatif dans l'espace des mesures, est très grande avec 1 pixel de bruit.

- **TQ** est la méthode consistant à minimiser (6.17) en partant des résultats de **TQ-LIN**. Cette méthode affine un peu la précision lorsqu'il y a peu de bruit, mais diverge complètement dans le cas contraire, comme le montre surtout les coordonnées de $\mathbf{f}'_1$.

- **F2LIN-INIT** est la méthode consistant à calculer une seconde matrice fondamentale par les équations linéaires (6.22), en partant d'une matrices fondamentale exacte. En partant de $\mathbf{F}_1$, (resp. $\mathbf{F}_2$, $\mathbf{F}_3$), les matrices fondamentales obtenues sont proches de $\mathbf{F}_2$ (resp. $\mathbf{F}_3$, $\mathbf{F}_2$) à 0.2 pixel, et de $\mathbf{F}_3$ (resp. $\mathbf{F}_3$, $\mathbf{F}_1$) à 1 pixel de bruit. L'une des trois matrices fondamentales n'est donc pas obtenue[7]. Cette méthode donne des résultats peu précis et une distance résiduelle (6.21) élevée, dès que le bruit significatif.

- **F2-INIT** consiste à partir des résultats de **F2LIN-INIT**, puis à minimiser le critère non-linéaire (6.21). La distance obtenue est celle qui était attendue, compte tenu du bruit, ce qui confirme l'intérêt de cette méthode, qui nécessite cependant la détermination préalable d'une matrice fondamentale.

- **F2-F** est identique à la méthode précédente, à ceci près que la première matrice fondamentale est obtenue par calcul direct (particulièrement instable en présence de bruit,

[7]Aussi, pour évaluer la précision de **F2-INIT**, nous avons initialisé la recherche non-linéaire par la matrice-cible, dans ce cas.

vu la configuration ...) à partir des appariements, qui sont les données mesurables. Les résultats, s'ils ne sont pas très précis, sont néanmoins utiles comme point de départ.

- **TQ-F2-F** consiste à minimiser (6.17) en partant des résultats de **F2-F**. Les dix-huit coefficients de (6.18) sont obtenus par l'équation (6.12). Les résultats en termes de précision de localisation, et surtout de distance résiduelle, sont satisfaisants: dans le cas du bruit faible, les résultats sont identiques à ceux obtenus par **TQ**, et dans l'autre cas, ils sont comparables à ceux obtenus par **F2-INIT**, alors que **TQ** avait divergé.

6.3 Caractérisation qualitative de l'instabilité: l'influence du type de mouvement

Dans cette section, nous analysons et mettons en évidence à l'aide de simulations l'instabilité liée à la *nature* de certains mouvements particuliers. Ces explications sont obtenues de manière synthétique, en contrôlant soigneusement les causes d'instabilité autres que celle étudiée.

La matrice fondamentale dépend du déplacement entre les images, et des paramètres intrinsèques. La stabilité de sa détermination doit donc être fonction, entre autres, du déplacement. Nous avions vu en 2.2.2 que la précision de détermination de la matrice fondamentale et celle des épipoles sont équivalents. Or, les relations présentées en 2.1.4 montrent que la position d'un des épipoles s'exprime directement à partir de la translation. Par conséquent, en nous penchant sur les facteurs qui influencent la stabilité de détermination de la translation, nous nous attendons à avoir une bonne idée de la stabilité de détermination de la matrice fondamentale, ce qui fait le lien avec des travaux existants [172]. Nous identifions trois types de mouvements conduisant à des résultats instables:

- les mouvements de faible amplitude
- les mouvements dont la translation est parallèle au plan image
- les translations pures

6.3.1 L'amplitude de la translation

Il est clair que si la translation est nulle, la matrice fondamentale n'est pas définie. De même, tous les vecteurs-déplacement induits dans l'image ont une norme inférieure au bruit moyen dans l'image, il n'est pas possible de déduire d'information des images. La norme de ces vecteurs croît avec l'amplitude du mouvement tridimensionnel, ce qui explique pourquoi les mouvements de faible amplitude conduisent à des estimations instables. Il est assez intuitif que plus ces vecteurs sont grands, plus leur erreur relative est faible, à niveau de bruit-image

fixé, et donc plus l'estimation du mouvement est fiable[8]. Cependant nous n'avons pas trouvé dans la littérature de démonstration générale de ce fait, pourtant considéré comme classique [37, 172, 72, 144, 176] qui n'est pas simple à établir analytiquement. Nous donnons donc une explication dans un cas simplifié, et des simulations statistiques pour monter sa réalité dans le cas général

Dans le cas d'une translation pure, l'estimation des épipoles est équivalente à celle du foyer d'expansion (FOE) pour laquelle Jezouin [73], en autres, a montré que l'erreur est une fonction décroissante de l'amplitude des déplacements observés dans l'image, à nombre de correspondances égal. Nous reprenons un raisonnement différent [33]. Au cours de la translation $\mathbf{t}$, nous avons les projections du point $M = (X, Y, Z)^T$ en coordonnées normalisées:

$$\begin{cases} x = \frac{X}{Z} \\ y = \frac{Y}{Z} \end{cases} \qquad \begin{cases} x' = \frac{X + t_x}{Z + t_z} \\ y = \frac{Y + t_y}{Z + t_z} \end{cases}$$

Le vecteur-déplacement observé dans l'image au point $(x, y)^T$ s'écrit par conséquent:

$$\Delta x = \frac{t_x - x t_z}{Z + t_z} \quad , \quad \Delta y = \frac{t_y - y t_z}{Z + t_z}$$

Dans le cas exact, l'épipole $\mathbf{e} = (\frac{t_x}{t_z}, \frac{t_y}{t_z})^T$ est obtenu comme solution de $\Delta x = \Delta y = 0$. Supposons à présent que les coordonnées du vecteur-déplacement $(\Delta x, \Delta y)^T$ soient perturbées par le bruit image ε. Le vecteur-déplacement perturbé est:

$$\Delta x + \varepsilon = \frac{(t_x + (Z + t_z)\varepsilon) - x t_z}{Z + t_z} \quad , \quad \Delta y + \varepsilon = \frac{(t_y + (Z + t_z)\varepsilon) - y t_z}{Z + t_z}$$

Si, de plus, tous les points observés se trouvent à la même distance Z, nous trouvons que les coordonnées de l'épipole sont celles de $\mathbf{e}$, perturbées par la quantité $\varepsilon(\frac{Z}{t_z} + 1)$. *L'erreur sur l'épipole est donc une fonction croissante de Z.* Au cours de l'estimation du mouvement à partir de deux vues, nous savons qu'à cause de l'ambiguïté échelle-vitesse, il est équivalent d'augmenter Z, ou de diminuer la norme de la translation, ce qui suggère que l'erreur sur l'épipole est inversement proportionnelle à la norme de la translation, puisqu'en général Zt_z est beaucoup plus grand que 1. Ceci signifie que les cas défavorables sont ceux où la translation est petite *devant la distance aux objets observés*. Il est bien sûr facile de détecter ces situations à partir des seules images, grâce à une simple mesure des vecteurs-déplacement.

Dans nos simulations, nous avons utilisé 100 séries de mouvements. Au sein de chaque série, la composante rotationnelle a été fixée, ainsi que la direction de la translation. Les épipoles obtenus sont donc identiques. Nous avons fait croire la norme de la translation, puis calculé la covariance sur les épipoles par la méthode présentée en 6.1.1, ainsi que les deux plus petites valeurs propres. Les résultats sont entièrement constants, aussi bien sur les covariances

[8] Notons qu'il revient au même de rapprocher les objets, ce qui illustre l'ambiguïté échelle-vitesse bien connue.

prédites, que sur les valeurs propres, dont nous observons que la plus petite valeur demeure à peu près constante, alors que la seconde croît systématiquement. Ceci s'explique bien par le fait que tous les autres facteurs sont inchangés, Nous donnons figure 6.7 un exemple dans lequel nous avons représenté l'écart-type en fonction de la norme de la translation, à laquelle nous avons fait prendre toutes les valeurs entières de 1 à 100. Afin d'illustrer les résultats de la simulation complète, pour chaque série, le logarithme du rapport des covariances obtenues avec les différentes normes, est calculé. La moyenne et l'écart-type de ces valeurs sont représentées à gauche de la figure 6.8, et une représentation graphique de tous ces rapports sous forme de niveaux de gris est donnée à droite de cette figure, sous forme d'un diagramme dans lequel chaque ligne horizontale représente une série de déplacements, l'angle croissant de la gauche vers la droite. Les valeurs les plus claires représentent les rapports les plus élevés.

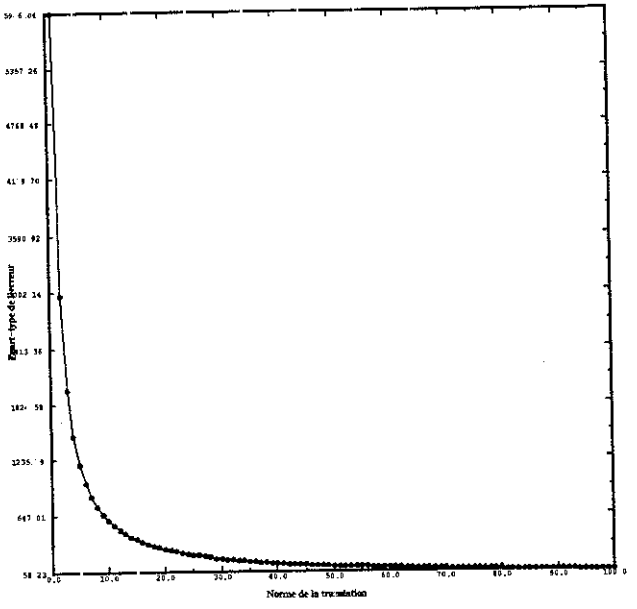


Fig. 6.7: Exemple: effet de la variation de la norme de la translation sur la stabilité

Il faut cependant noter que le nombre de correspondances qu'il est possible d'obtenir en pratique dépend également de l'amplitude du mouvement: lorsque celui-ci augmente le champ de vue commun aux caméras se réduit. Pour des mouvements de grande amplitude, il est difficile d'obtenir des correspondances nombreuses. Or il est évident que l'erreur est une fonction décroissante du nombre n de correspondances utilisées. En nous fondant sur la remarque effectuée en 5.2.1, nous pouvons dire que l'erreur moyenne, qui est donnée par l'écart-type, est à peu près inversement proportionnelle à $\sqrt{n}$, ce qui est classique, et consistant avec les résultats obtenus par de nombreux auteurs[9].

[9]La fonction $\sqrt{n}$ est par exemple mentionnée par [144]

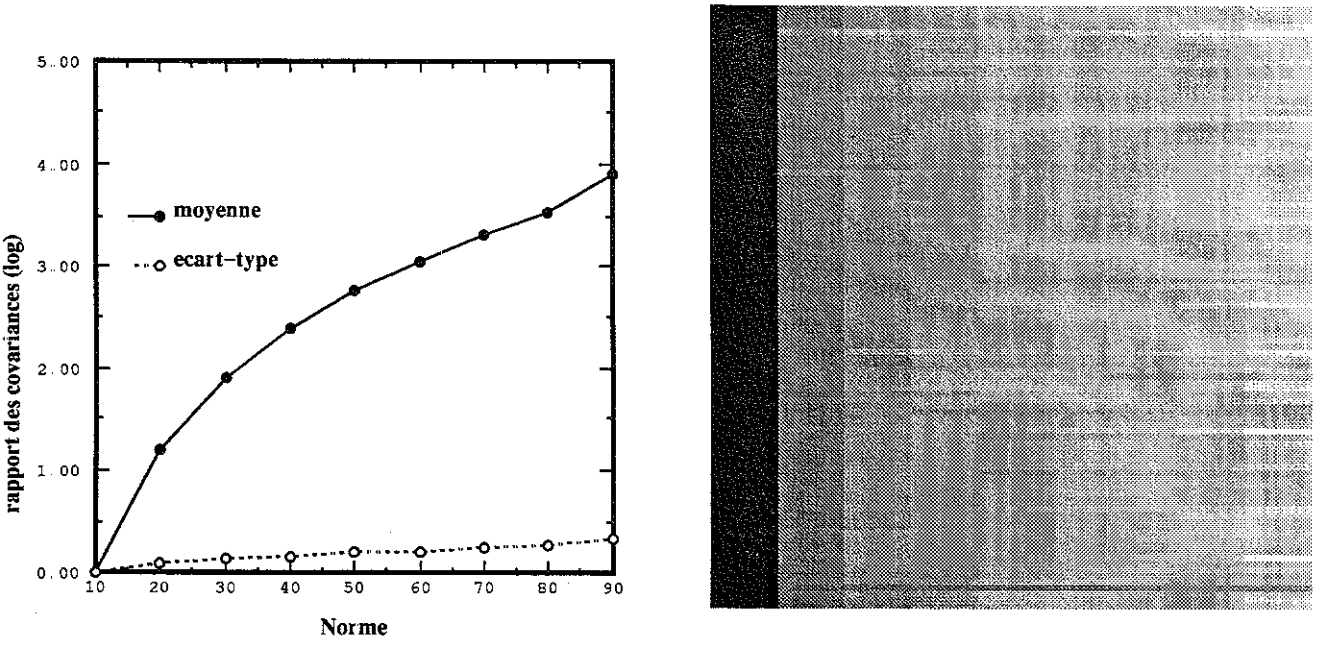


Fig. 6.8: Statistiques: effet de la variation de la direction de la translation sur la stabilité

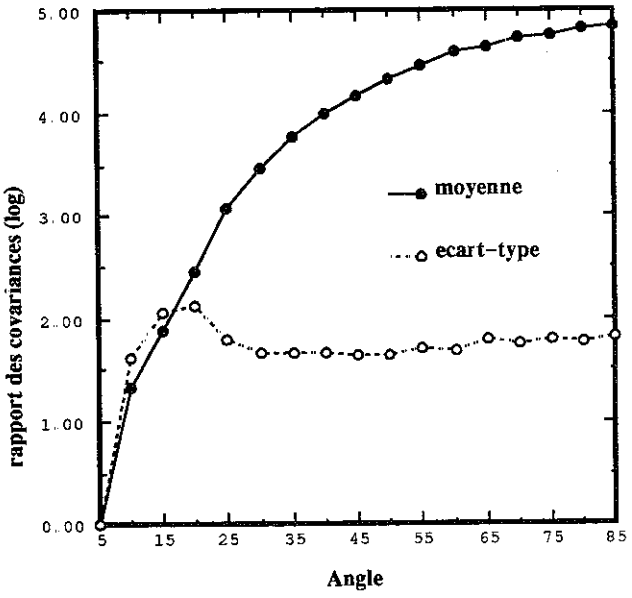
6.3.2 La direction de la translation

L'influence de la direction de la translation sur la stabilité du calcul du mouvement a été remarquée expérimentalement par plusieurs auteurs [37, 111, 3, 172], qui ont montré que les résultats les plus stables sont obtenus avec des translations perpendiculaires au plan image. Plusieurs arguments géométriques [65, 172] ont été invoqués pour expliquer ce fait, qui a également été retrouvé de manière analytique dans un cas particulier [25].

L'idée la plus simple et la plus convaincante est que la composante t_z de la translation perpendiculaire au plan image est plus sensible, puisqu'une translation perpendiculaire à l'image conduit à des déplacements image moins importants qu'une translation parallèle au plan image. Par conséquent, la direction d'une translation t est plus sensible si la composante t_z est faible, puisque l'erreur relative sur t_z sera plus importante. Tout comme pour la propriété mise en évidence précédemment, la démonstration analytique dans le cas général n'est pas connue, cependant, elle apparaît clairement dans le cadre simplifié déjà utilisé: l'erreur sur l'épipole $\varepsilon(\frac{Z}{t_z}+1)$ est inversement proportionnelle à t_z si nous fixons Z, et par conséquent la norme de la translation.

Dans le cadre du calcul de la matrice fondamentale, nous pouvons remarquer que la distance des épipoles au centre de l'image est déterminée par les rapports $\frac{t_x}{t_z}$ et $\frac{t_y}{t_z}$ des composantes de la translation parallèles au plan image par la composante perpendiculaire. Dans le chapitre 3 consacré à la détermination des épipoles par la transformation épipolaire, nous avons déjà constaté et expliqué que la détermination est plus précise lorsque les épipoles sont proches du centre de l'image que lorsqu'ils en sont éloignés, ce qui recoupe l'observation sur la direction de translation.

Nous présentons une simulation pour valider ce résultat dans le cas général. Nous générons aléatoirement 100 séries de déplacements au cours desquelles nous faisons varier l'angle entre le plan image et la direction de translation de façon identique, de 5^o à 85^o. Des résultats très constants sont trouvés. Le principe des graphes présentés figure 6.9 est identique à celui des graphes 6.8, la quantité variée cette fois-ci étant l'angle. Les résultats sont encore tout à fait convaincants, même s'ils sont moins constants que les précédents, ce qui est dû au fait qu'en variant la direction de la translation nous créons un déplacement différent (pour lequel des configurations critiques peuvent apparaître).



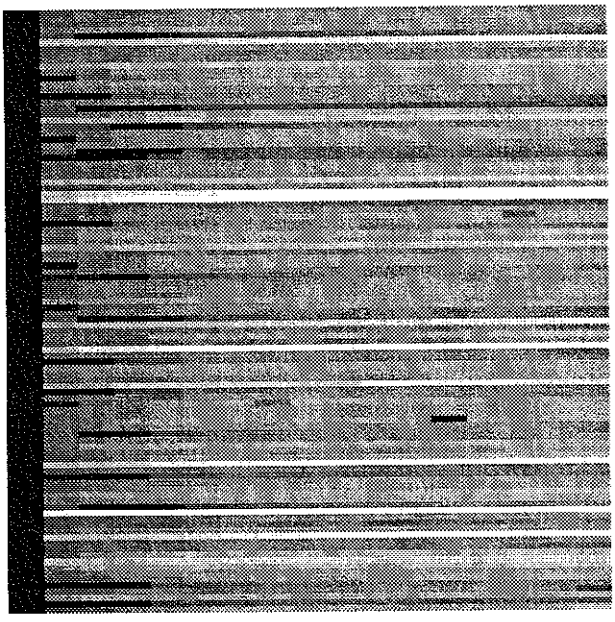


Fig. 6.9: Statistiques: effet de la variation de la direction de la translation sur la stabilité

6.3.3 L'angle de rotation: le cas des translations pures

Nous n'avons pas remarqué d'influence de la valeur de l'angle de rotation sur la stabilité des résultats, à l'exception du cas d'une classe de mouvements particuliers, les translations pures. Celles-ci conduisent, *avec le critère linéaire*, à des estimations instables, fait qui n'a pas été rapporté explicitement dans la littérature, au contraire des deux précédents.

Nous donnons d'abord un exemple pour illustrer cette observation. Nous sommes partis de la translation pure définie par $\mathbf{t} = [-436.43,\ 109.10,\ -218.21]^T$. Une composante rotationnelle a été ajoutée, suivant l'axe: $\mathbf{r} = [-0.3434,\ 0.6203,\ 0.7051]^T$, en faisant varier l'angle de la rotation de 0 à 0.5, pour obtenir un déplacement qui s'apparente de moins en moins à une translation pure. Le choix des valeurs des vecteurs rotation et translation permet de contrôler les autres facteurs d'instabilité: le déplacement est assez important, l'épipole, situé en $(1526.34, -216.19)$ n'est pas très éloigné. Nous avons aussi calculé la distance à la surface critique pour chaque déplacement, et trouvé dans tous les cas une valeur supérieure

à 30 pixels. Dans la table 6.2, nous avons fait figurer, pour chacune des valeurs de l'angle, l'écart-type sur les coordonnées de l'épipole fixe, ainsi que la moyenne des deux plus petites valeurs propres de $\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}$. Pour chaque déplacement, 100 essais ont été effectués en variant le bruit, dont la variance a été fixée à $\sigma = 0.2$ pixel. Il est à noter que pour des valeurs de bruit très faibles, le phénomène que nous cherchons à mettre en évidence ne se produit pas, de même que pour des bruits élevés (supérieurs à 2 pixels). Nous avons ensuite validé notre

angle	$\sigma_{\mathbf{e}'_x}$	$\sigma_{\mathbf{e}'_y}$	λ_1	λ_2
0	1769	920.7	0.3375	1.423
0.0001	1734	935.1	0.3377	1.424
0.001	1289	1061	0.3335	1.444
0.002	951.5	1049	0.3235	1.492
0.005	1049	777.9	0.2795	1.759
0.01	434	593.8	0.1893	2.637
0.02	141	58.43	0.0830	6.065
0.05	76.96	33.03	0.0165	25.34
0.1	71.88	38.04	0.0039	111.2
0.2	89.2	36.04	0.0008	725.2
0.5	124.1	69.3	7.10^{-5}	1100

Tableau 6.2: Exemple: influence de l'angle de la rotation sur la stabilité du critère linéaire

observation au moyen d'essais statistiques, en variant aléatoirement l'axe de rotation, ainsi que la translation au cours de 100 essais. Les valeurs de l'angle essayées sont celles de la table précédente, à l'exception de 0.0001. Les résultats obtenus illustrent clairement le fait qu'en général, plus le mouvement se rapproche d'une translation pure, plus les résultats sont instables. Le cas des translations pure apparaît particulier, car au delà d'une certaine valeur, il n'y a pas d'influence de la valeur de l'angle sur la stabilité, fait par contre déjà remarqué par [172]. Cette observation ne concerne pas les critères non-linéaires normalisés, dont le minimum est largement insensible à ce type d'instabilité. Cependant, comme nous utilisons les résultats du critère linéaire pour initialiser la recherche, celle-ci a de grandes probabilités d'aboutir à un mauvais minimum local, puisque l'estimée initiale peut être très imprécise.

6.4 Caractérisation qualitative de l'instabilité: l'ambiguïté

Nous allons voir que le problème de l'ambiguïté a des conséquences pratiques importantes, l'existence de configurations à forte *instabilité*. Nous donnons d'abord une explication théorique à ce fait, puis nous l'illustrons à l'aide d'un exemple numérique. Nous finissons par une large simulation, dans laquelle nous combinons les explications relatives à la nature du

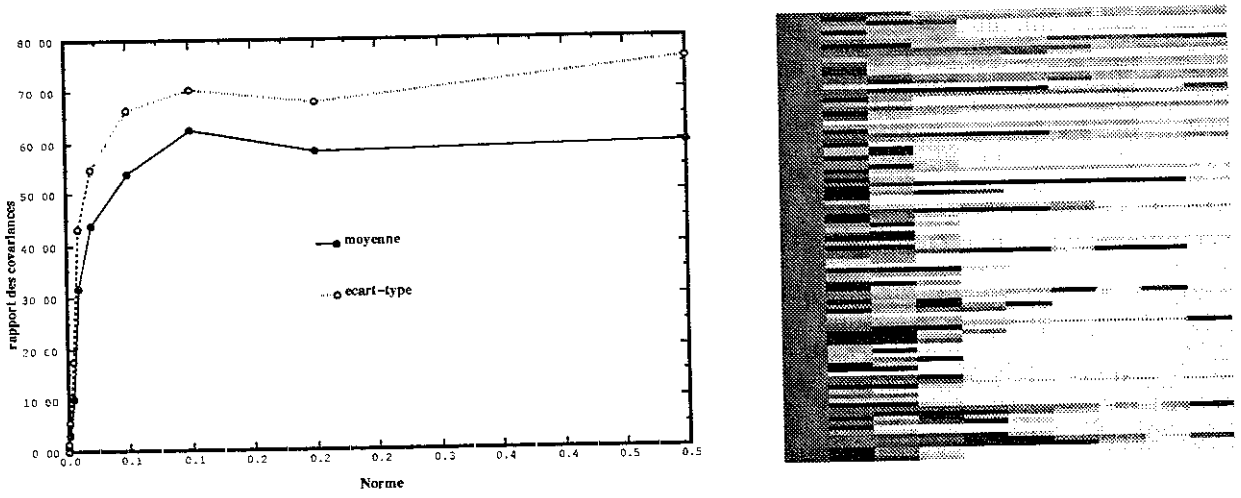


Fig. 6.10: Statistiques: effet de la variation de la norme de la rotation sur la stabilité

mouvement obtenues précédemment, et celles relatives à l'interaction entre le mouvement et la structure tridimensionnelle sous forme de proximité à une surface critique, afin de montrer qu'elles rendent parfaitement compte de l'instabilité observée. En particulier, l'aspect statistique de l'expérimentation et le fait que des points aléatoires soient utilisées montre que les surfaces critiques sont effectivement rencontrées dans des situations vraisemblables.

6.4.1 Lien théorique entre instabilité et ambiguïté

Les surface critiques ont éte présentées en (6.2.1) comme des ensembles de points rendant l'interprétation du mouvement ambigüe. Maybank [105] a montré qu'une configuration tridimensionnelle de points instable par reconstruction est proche d'une surface critique. Nous allons établir le lien inverse, qui n'est pas si évident a priori: il est en effet tout à fait concevable qu'une situation ambigüe soit telle que chacune des solutions soit parfaitement stable. Nous pouvons par exemple imaginer un critère qui posséderait plusieurs minima globaux parfaitement séparés localisés

L'instabilité apparaît très clairement dans la formulation de Horn [64], qui définit les surfaces critiques d'une manière différente de la nôtre. Elles sont en effet introduites comme le lieu de points $\mathbf{M}$ pour lesquels la variation du terme $\mathbf{m}'T\mathbf{E}\mathbf{m}$ est seulement du second ordre en les paramètres du mouvement $\mathbf{r}$, $\mathbf{t}$. La localisation du minimum est donc très imprécise, puisque si nous effectuons le calcul aux moindres carrés, le critère devient une fonction d'ordre au minimum quatre en fonction des variations infinitésimales $\delta\mathbf{r}$, $\delta\mathbf{t}$. Horn obtient des surfaces critiques dont l'équation prend une forme apparemment différente de 6.10, mais l'étude qu'il leur consacre montre qu'elles ont exactement les même propriétés mathématiques que celles apparaissant dans la formulation de Maybank [104]. Nous allons montrer que les deux formes sont équivalentes, ce qui prouvera qu'une situation ambigüe est intrinsèquement instable.

Nous nous plaçons en coordonnées normalisées, le déplacement de la caméra étant $\mathbf{R}$, $\mathbf{t}$, les centres optiques initial et final étant notés respectivement $\mathbf{C}$ et $\mathbf{C}'$. Supposons que la

translation soit perturbée par le vecteur infinitésimal $\delta\mathbf{t}$ et la rotation par le vecteur rotation $\delta\mathbf{r}$. Lorsque nous recherchons une solution approchée, nous utilisons l'équation de Longuet-Higgins pour obtenir un terme d'erreur. A l'aide de produit mixtes, la différence de ceux-ci, obtenus avec et sans perturbation, s'écrit, dans le repère final de la caméra:

$$\Delta = [(\mathbf{t} + \delta\mathbf{t}), \mathbf{C'M}, \mathbf{CM} + \delta\mathbf{r} \wedge \mathbf{CM}] - [\mathbf{t}, \mathbf{C'M}, \mathbf{CM}] \tag{6.23}$$

En effet, une rotation infinitésimale $\delta\mathbf{R}$ peut s'écrire en fonction de son vecteur rotation associé $\delta\mathbf{r}$ grâce à la formule de Rodrigues, en tenant compte du fait que $\theta = \|\delta\mathbf{r}\|$ est un infiniment petit:

$$\delta\mathbf{R} = e^{\delta\mathbf{r}} = \mathbf{I} + \frac{\sin\theta}{\theta}\widetilde{\delta\mathbf{r}} + \frac{1-\cos\theta}{\theta}\widetilde{\delta\mathbf{r}}^2 \sim \mathbf{I} + \widetilde{\delta\mathbf{r}}$$

La différence Δ définie par (6.23) est normalement du premier ordre en fonction de $\delta\mathbf{r}$ et de $\delta\mathbf{t}$. Les situations instables se produisent lorsque cette différence est d'un ordre supérieur: alors une variation des paramètres du mouvement dans la direction $(\delta\mathbf{r}, \delta\mathbf{t})$ n'a qu'un effet faible sur Δ. En négligeant le terme $[\delta\mathbf{t}, \mathbf{C'M}, \delta\mathbf{r} \wedge \mathbf{CM}]$ qui est du second ordre, nous obtenons, en distribuant les produits:

$$\Delta = [\mathbf{t}, \mathbf{C'M}, \delta\mathbf{r} \wedge \mathbf{CM}] + [\delta\mathbf{t}, \mathbf{C'M}, \mathbf{CM}]$$

Nous pouvons réécrire chacun des deux termes, en utilisant $\mathbf{t} = \mathbf{C'C}$ et les propriétés du produit mixte:

$$[\mathbf{t}, \mathbf{C'M}, \delta\mathbf{r} \wedge \mathbf{CM}] = [\delta\mathbf{r} \wedge (\mathbf{C'M} - \mathbf{t}), \mathbf{t}, \mathbf{C'M}] = [(\mathbf{I} + \delta\mathbf{r}) \wedge \mathbf{C'M} - \delta\mathbf{r} \wedge \mathbf{t}, \mathbf{t}, \mathbf{C'M}]$$

$$[\delta\mathbf{t}, \mathbf{C'M}, \mathbf{CM}] = [\delta\mathbf{t}, \mathbf{C'M}, \mathbf{C'M} - \mathbf{t}] = [\delta\mathbf{t}, \mathbf{t}, \mathbf{C'M}]$$

Il vient finalement, en additionnant les deux termes, l'expression:

$$\Delta = [(\mathbf{I} + \delta\mathbf{r}) \wedge \mathbf{C'M} - \delta\mathbf{r} \wedge \mathbf{t} + \delta\mathbf{t}, \mathbf{t}, \mathbf{C'M}]$$

En développant celle-ci, nous trouvons une surface critique du type de celles présentées par Horn. Si nous faisons intervenir les coordonnées de M dans le repère initial de la caméra, nous obtenons, en écrivant la nullité du produit mixte:

$$((\mathbf{I} + \delta\mathbf{r}) \wedge \mathbf{RM} - \delta\mathbf{r} \wedge \mathbf{t} + \delta\mathbf{t})^T \mathbf{t} \wedge \mathbf{RM} = 0 \tag{6.24}$$

Considérons alors une surface critique donnée par 6.10, que nous pouvons écrire dans le repère initial de la caméra:

$$(\mathbf{R}_2\mathbf{M} + \mathbf{t}_2)^T \mathbf{t} \wedge \mathbf{RM} = 0$$

Une telle surface peut aussi s'écrire avec une équation de la forme (6.24), si nous posons:

$$\begin{aligned} \mathbf{I} + \widetilde{\delta\mathbf{r}} &= \mathbf{R}_2\mathbf{R}^{-1} \\ \delta\mathbf{t} &= \mathbf{t}_2 + \delta\mathbf{r} \wedge \mathbf{t} = \mathbf{t}_2 - \mathbf{t} + \mathbf{R}_2\mathbf{R}^{-1}\mathbf{t} \end{aligned}$$

Il faut noter que nous faisons un lien entre des quantités finies et des infiniments petits, ce qui nécessite quelques précautions: $\mathbf{I}+\widetilde{\delta \mathbf{r}}$ n'est une matrice de rotation que si $\delta \mathbf{r}$ est infinitésimal, $\delta \mathbf{t}$, défini seulement à un facteur d'échelle près, est également un vecteur dont la direction est donnée par la seconde expression.

Nous avons donc montré qu'une surface critique définie comme surface ambigüe l'est également au sens où elle conduit à une détermination instable du mouvement: le minimum du critère d'erreur ne peut être localisé précisément, car le terme d'erreur a une variation d'ordre élevé en fonction de la variation des paramètres du mouvement. De plus, nous avons établi la direction infinitésimale d'instabilité, et conjecturons que celle-ci pourrait servir à caractériser l'ensemble des solutions compatibles avec un certain niveau d'approximation.

6.4.2 Une expérience à partir d'une surface critique

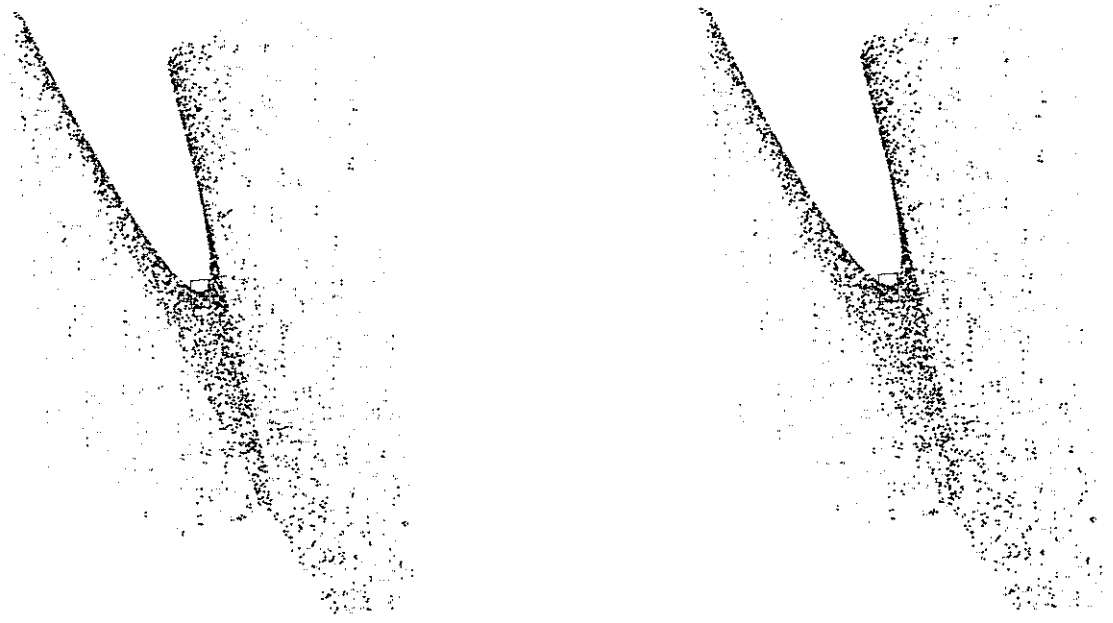


Fig. 6.11: La surface critique utilisée dans les expériences (stéréogramme pour vue croisée)

Nous générons pour commencer une surface critique $\mathcal{S}_{12}$, obtenue à partir des deux mouvements 1 et 2 introduits en (3.1) et déjà largement utilisés, puis nous répartissons aléatoirement un ensemble de points sur cette surface. Un stéréogramme représentant les points tridimensionnels obtenus est présenté figure 6.11.

Les appariements sont ensuite obtenus en reprojetant ces points sur les deux rétines placées à la position initiale et la position obtenue après le mouvement 1. Ils sont utilisés pour déterminer les épipoles par l'intermédiaire des matrices fondamentales calculées avec le

critère non-linéaire **DIST-L**. Lorsque nous ne bruitons pas les points, nous observons une instabilité qui se manifeste par le fait qu'en ajoutant ou en retranchant un seul appariement, nous faisons converger le calcul vers une solution différente. Lorsque les points sont bruités, nous notons une tendance systématique à la divergence, avec également une très grande sensibilité aux appariements choisis ainsi qu'à la variation du bruit.

L'étape suivante consiste à répartir les points de l'espace non plus exactement sur la surface critique, mais à proximité de celle-ci. Plus précisément, chaque point $\mathbf{M}_i(d)$ est obtenu à partir du point $\mathbf{M}_i$ de la surface $\mathcal{S}_{12}$ comme $\mathbf{M}_i \pm d\mathbf{n}_i$, ou $\mathbf{n}_i$ est la normale à la surface $\mathcal{S}_{12}$ au point $\mathbf{M}_i$, et d un réel fixé, qui représente donc la distance à la surface des nouveaux points. Augmenter la valeur de d consiste donc à éloigner les points de la surface critique, ou encore, de manière imagée, à "bruiter" celle-ci. Il est alors intéressant de voir dans quelle mesure l'ambiguïté demeure robuste au "bruit". Dans une première série d'expériences, dont les résultats figurent table 6.3, nous avons fixé différents niveaux de bruit image et différentes valeurs de la distance d, puis nous avons généré à partir des mêmes points critiques des ensembles d'appariements. Pour chacun de ces ensembles, nous avons effectué le calcul des épipoles à l'aide de **DIST-L** en variant le nombre de points sélectionnés, puis nous avons calculé l'écart-type résultant de ces essais. Dans la seconde série d'expériences, nous avons fixé le niveau de bruit à 1 pixel afin de faire varier les perturbations elles-mêmes. Pour chaque valeur de la distance d, 50 essais différents ont été effectués. La table 6.4, montre les moyennes et surtout les écarts types obtenus. Nous avons fait également figurer dans cette table les moyennes d_x et d_y des écarts entre l'image des points de la surface critique exacte et l'image des points de la surface critique perturbée. En effet ces quantités sont les seules qui soient déterminables à partir des images elles-mêmes, comme nous le verrons plus tard. Toutes les valeurs présentées concernent la première image, des résultats similaires étant obtenus dans la seconde image.

	$b=0$				$b=0.5$				$b=1$			
	e_x	e_y	σ_{e_x}	σ_{e_y}	e_x	e_y	σ_{e_x}	σ_{e_y}	e_x	e_y	σ_{e_x}	σ_{e_y}
d = 0	-7587	-4291	6140	3639	1356	1006	1466	872	1396	1031	1261	788
d = 5	2033	1409	10^{-7}	10^{-7}	1014	785	2935	1765	3840	2499	3749	2305
d = 10	2033	1409	10^{-7}	10^{-7}	1994	1369	726.	459	1770	1229	822	492
d = 20	2033	1409	10^{-7}	10^{-7}	2007	1391	153	106	2010	1395	280	199
d = 50	2033	1409	10^{-7}	10^{-7}	1924	1328	39	40	1834	1262	65	68

Tableau 6.3: Influence de la distance à la surface critique et du bruit sur la stabilité

Les valeurs d'écart type obtenues montrent assez clairement que plus les points sont éloignés de la surface critique, plus les calculs conduisent à des épipoles stables. Le fait apparemment curieux, qu'avec les mouvements choisis les résultats semblent plus stables avec les points de la surface critique qu'avec les points obtenus pour $d = 5$ s'explique assez bien par le fait qu'à $d = 0$, les épipoles trouvés oscillent surtout entre $(2034, 1409)$ et $(214, 255)$,

d	d_x	d_y	e_x	e_y	σ_{e_x}	σ_{e_y}
0	0	0	987.61	789.46	1108.41	938.34
5	3.89	7.74	3946.67	2556.09	7669.45	4625.23
10	7.60	14.51	2267.57	1545.47	765.85	446.66
20	15.19	29.12	2012.15	1396.76	224.01	135.48
50	89.34	148.53	2020.80	1401.83	170.26	113.78

Tableau 6.4: Influence de la distance à la surface critique sur la stabilité

alors qu'à $d = 5$, un nouveau minimum local fait son apparition, qui est dans la direction d'un troisième épipole virtuel situé très loin de l'origine.

Lorsque les points sont suffisamment éloignés de la surface critique, le fait d'augmenter le bruit dans l'image augmente la variance des épipoles, ce qui est classique, cependant lorsque les points sont sur la surface critique ou très proches, le fait de les bruiter induit une erreur de reconstruction dont l'effet est d'éloigner les points tridimensionnels ainsi obtenus de la surface critique, ce qui explique qu'en augmentant le bruit dans l'image on diminue alors un peu la variance des épipoles calculés.

6.4.3 Une validation statistique globale

Nous avons utilisé une méthode statistique comparable à celle présentée en 6.1.3. 500 déplacements ont été générés aléatoirement. Afin de ne pas fausser le raisonnement sur le mouvement, le nombre de points utilisés a été fixé à 50 dans tous les cas. Nous avons classé ces déplacements de manière croissante en fonction de la covariance prédite. Sur la figure 6.12, dont l'interprétation synthétise les analyses de cette section, l'axe des abscisses représente donc le degré d'instabilité. Nous avons partiellement contrôlé l'instabilité due aux translations pures et faibles translations en fixant un minimum pour les vecteurs rotation et translation.

déplacements (instabilité croissante)	surface critique à moins de 10 pixels	distance des épipoles au centre de l'image
1-100	9%	754.6 pixels
101-200	13%	1164 pixels
201-300	31%	1783 pixels
301-400	40%	2624 pixels
401-500	49%	25280 pixels

Tableau 6.5: Facteurs d'instabilité dans une expérience statistique

Nous nous sommes d'abord intéressés à la distance de l'épipole au centre de l'image, en représentant en ordonnée une fonction croissante de cette distance. Nous remarquons une

certaine corrélation entre cette distance, et l'instabilité, qui est quantifiée table 6.5. L'idée suivante que nous cherchons à vérifier est que dans les cas pour lesquels cette raison n'explique pas bien l'instabilité, celle-ci provient de la proximité d'une surface critique. La distance, en projection, des points à une surface critique a été calculée à partir des appariements par la méthode décrite en (6.2.4) Nous avons marqué en noir les points correspondant aux mouvements pour lesquels cette distance est inférieure à 10 pixels, ce qui est une distance pour laquelle l'instabilité est significative, comme l'a montré l'exemple présenté table 6.4. Le fait que ces points se situent statistiquement *en dessous* des points clairs (correspondant à des distances à la surface critique supérieures à 10 pixels) valide cette explication. En effet, si nous considérons deux points N_1 et N_2 de notre graphe, d'abscisse à peu près égale, mais d'ordonnées différentes $y_1 > y_2$. Les points N_1 et N_2 correspondent à deux déplacements qui ont un degré d'instabilité équivalent, cependant si nous nous référons uniquement à l'explication par l'éloignement de l'épipole, l'instabilité de N_2 devrait être inférieure à celle de N_1. L'instabilité de N_2 doit donc s'expliquer par un autre facteur, et l'observation que nous avons faite montre que la distance à la surface critique est ce facteur. Nous aussi pouvons observer que les points noirs sont beaucoup plus denses sur la droite du graphe, que sur la gauche, ce qui s'interprète par le fait que les mouvements les plus stables sont statistiquement ceux qui sont le plus éloignés de la surface critique. La corrélation est également significative, comme le montre la table 6.5, cependant il y a une interaction entre les deux explications, qui rend leur combinaison tout à fait convaincante.

Une autre conclusion importante qu'il est possible de tirer de ces simulations est l'omniprésence, ou du tout du moins la présence fréquente, de la surface critique, qui se trouve, en projection, à moins de 10 pixels dans 28% des mouvements. Bien que les surfaces critiques n'existent pratiquement jamais *sous forme exacte* dans la nature, elles ont donc une très grande importance pratique car le *volume critique* à l'intérieur duquel les points doivent se trouver pour qu'une certaine instabilité se manifeste est, au vu de ces expériences, vaste. Nous pouvons détecter ces configurations grâce aux algorithmes de calcul de transformation quadratique introduits dans ce chapitre

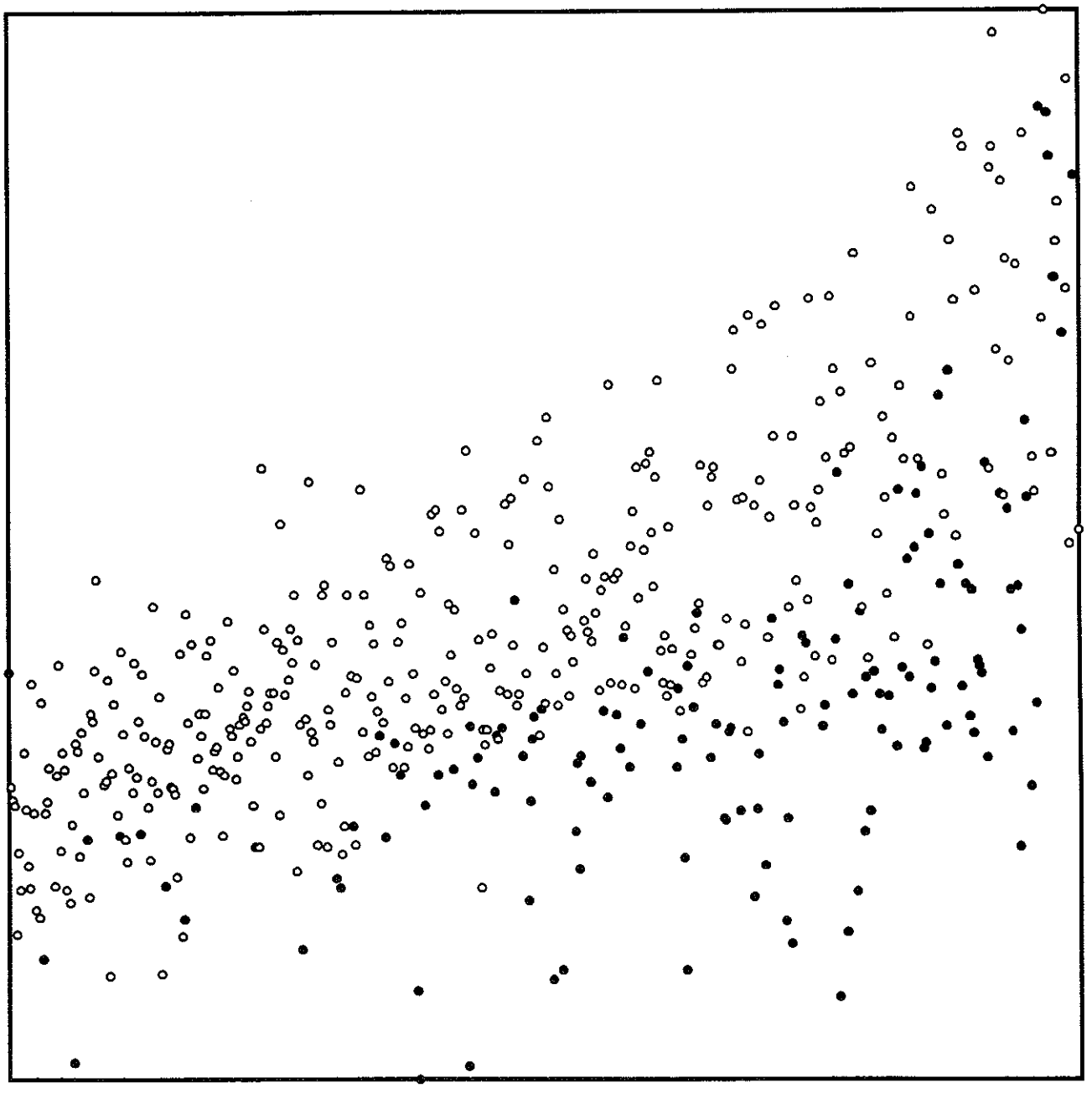


Fig. 6.12: Une caractérisation expérimentale globale de l'instabilité

Chapitre 7

Le calcul des paramètres intrinsèques: théorie

La contrainte géométrique résultant de la rigidité de tout déplacement physique effectué par une caméra se traduit par des équations polynomiales reliant la matrice fondamentale et les paramètres intrinsèques. Cette idée a déjà été présentée sous des formes inachevées que nous commençons par reformuler pour montrer le lien avec notre approche. Bien que nous montrons qu'il est possible d'utiliser la matrice essentielle pour exprimer les contraintes de rigidité, nous préférons faire intervenir la conique absolue et les équations de Kruppa. Nous établissons l'équivalence de celles-ci et des premières expressions, ainsi que deux résultats sur l'indépendance et la singularité des déplacements, qui sont importants en pratique. Nous explicitons ensuite les liens algébriques qui existent entre la conique absolue, les paramètres intrinsèques, et les coefficients obtenus en résolvant les équations de Kruppa. Les contraintes provenant du fait que la conique absolue soit purement imaginaire sont importantes et nous permettent de retrouver le fait que le modèle de caméra que nous considérons est général.

7.1 Calculer les paramètres intrinsèques à l'aide de la contrainte de rigidité

Nous présentons l'idée qui est à la base de nos méthodes de calcul des paramètres intrinsèques: l'utilisation d'une caméra en mouvement, permettant d'obtenir deux équations contraignant ceux-ci. Des travaux antérieurs et voisins sont décrits, dans lesquels l'absence de cette idée a contraint à des hypothèses simplificatrices. Nous exposons ensuite une première approche pour exprimer la contrainte de rigidité permettant d'obtenir les deux équations. Elle fait appel à la matrice essentielle, et nous montrons que les deux formulations existant dans la littérature sont équivalentes.

7.1.1 Le principe: utiliser une caméra en mouvement

L'idée des méthodes de calcul des paramètres intrinsèques que nous présentons dans ce chapitre est d'exploiter le fait que si nous observons une scène rigide par l'intermédiaire d'une caméra mobile unique, dont nous supposons que les paramètres intrinsèques restent inchangés, des contraintes existent du fait de l'invariance de ces paramètres, et du fait de la rigidité du déplacement de la caméra. Ce sont ces contraintes que nous allons expliciter et utiliser pour le calcul des paramètres intrinsèques.

Considérons d'abord le cas de deux vues d'une même scène prises par des caméras quelconques. Sous les hypothèses de projectivité qui sont à la base de notre travail, nous pouvons toujours caractériser la relation entre ces deux vues d'une même scène par une matrice fondamentale $\mathbf{F}$. Nous avons vu qu'une matrice fondamentale comprend sept paramètres indépendants. Ces paramètres dépendent à leur tour, d'une part du mouvement tridimensionnel $\mathbf{D}$ entre les repères intrinsèques attachés à chaque caméra, et d'autre part des paramètres intrinsèques $\mathbf{A}$ et $\mathbf{A}'$ de chaque caméra. Comme nous ne disposons pas de référence métrique, nous ne pouvons considérer la géométrie tridimensionnelle qu'à un facteur d'échelle près. Le mouvement tridimensionnel est alors entièrement caractérisé par la connaissance de la rotation (trois paramètres) et de la direction de la translation (deux paramètres), c'est-à-dire par cinq paramètres. La connaissance de la matrice fondamentale détermine donc exactement deux équations indépendantes reliant les paramètres intrinsèques contenus dans $\mathbf{A}$ et $\mathbf{A}'$.

Dans le cadre que nous décrivons, Trivedi [157] a tenté de déterminer les paramètres d'un modèle comportant uniquement les décalages d'origine, c'est-à-dire tel que

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & u_0 \\ 0 & 1 & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

à l'aide des *trois* équations (7 3), décrites un peu plus bas. Voyons son idée initiale: si nous considérons deux images, prises par des caméras différentes, nous avons deux tels paramètres par caméra, c'est-à-dire quatre paramètres intrinsèques à déterminer, et seulement trois équations. Si nous ajoutons une troisième image de la même scène, prise par une autre

caméra, nous aurons six paramètres intrinsèques à déterminer, et au moins six équations, si bien que la résolution serait envisageable. Cependant Trivedi a montré lui-même que dans le cas particulier qu'il considérait, les trois équations se réduisent en fait à deux équations indépendantes et une tautologie. Aussi conclut-il que le calcul des paramètres intrinsèques n'est pas possible avec cette méthode.

Récemment, Hartley [61] a également utilisé un modèle de caméra simplifié, qui est différent. Il suppose que le seul paramètre intrinsèque inconnu est la distance focale, et donc que la matrice de paramètres intrinsèques se réduit à:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix}$$

Il montre alors, en fournissant en même temps un algorithme, qu'il est possible de décomposer une matrice fondamentale $\mathbf{F}$ en $\mathbf{A}'^{-1T}\mathbf{E}\mathbf{A}$, obtenant ainsi le mouvement tridimensionnel et les distances focales associées à chaque caméra. Ceci est cohérent avec le décompte du nombre d'équations et d'inconnues que nous avons effectué: deux paramètres intrinsèques peuvent être déterminés dans le cas général, si les autres sont connus.

Si nous ne faisons pas d'hypothèse supplémentaire sur les caméras, nous ne pouvons en effet pas obtenir beaucoup d'information sur $\mathbf{A}$ et $\mathbf{A}'$, puisque nous ne pouvons disposer au plus que de deux équations, et que le fait de rajouter des images rajoute autant d'inconnues que d'équations. C'est pourquoi il a été proposé de simplifier le modèle de caméra en réduisant le nombre de paramètres intrinsèques. Il est cependant possible de calculer un modèle avec une matrice de paramètres intrinsèques entièrement générale, si nous nous astreignons à utiliser toujours la même caméra, placée en des points de vue différents. Le fait que nous utilisions toujours la même caméra signifie que nous n'aurons qu'une matrice de paramètres intrinsèques à considérer, puisque nous faisons alors l'hypothèse naturelle que ceux-ci sont constants au cours du temps. Nous avons seulement la supposition supplémentaire à faire que la scène que nous observons est bien rigide, puisque nous sommes amenés à prendre des images à des instants différents, entre les différents déplacements de la caméra. Dans ces conditions, nous voyons que chaque déplacement de la caméra donne lieu à une matrice fondamentale, dont nous devrions pouvoir tirer deux équations contraignant les paramètres intrinsèques $\mathbf{A}$. Comme ceux-ci sont constants d'un mouvement à l'autre, l'accumulation des mouvements permet d'obtenir un nombre d'équations suffisant pour leur calcul.

7.1.2 Les contraintes de rigidité

Dans le cas de deux caméras quelconques, la transformation permettant de passer d'un repère rétinien à l'autre est une transformation linéaire projective quelconque de $\mathcal{P}^3$. Si les deux caméras sont identiques, cette transformation est, dans le repère rétinien d'une des caméras, un déplacement rigide. C'est en exprimant cette rigidité que nous obtenons les deux équations nécessaires au calcul des paramètres intrinsèques.

Un premier ensemble de méthodes permettant d'exprimer la contrainte de rigidité fait intervenir la matrice essentielle. Nous pouvons obtenir celle-ci à partir de la matrice fondamentale et de la matrice de paramètres intrinsèques par

$$\mathbf{E} = \mathbf{A}^T\mathbf{F}\mathbf{A} \tag{7.1}$$

La rigidité du mouvement représenté par F avec les paramètres intrinsèques A est donc équivalente aux conditions de Huang et Faugeras, qui traduisent le fait que E, obtenue par (7.1) est bien une matrice essentielle:

$$\det(\mathbf{E}) = 0 \qquad f(\mathbf{E}) = \frac{1}{2}\text{trace}^2(\mathbf{E}\mathbf{E}^T) - \text{trace}(\mathbf{E}\mathbf{E}^T)^2 = 0$$

La première condition est automatiquement satisfaite, puisque $\det(\mathbf{E}) = 0$. Il nous reste donc une contrainte polynômiale unique, qui est de degré quatre en les coefficients de E. Comme E est quadratique en les coefficients de A et linéaire en ceux de F, nous avons une contrainte qui est de degré huit en les coefficients de A, ce qui n'est pas tout à fait ce que nous attendions. Nous allons cependant voir plus bas que cette contrainte peut se décomposer en *deux* relations polynômiales indépendantes.

Une seconde expression de la contrainte de rigidité, au moyen de la matrice essentielle, a été présentée par Trivedi. Si E est une matrice essentielle, la matrice symétrique $\mathbf{S} = \mathbf{E}\mathbf{E}^T$, qui *a priori* comprend six éléments indépendants, ne dépend que des trois coordonnées de t et a la valeur:

$$\mathbf{E}\mathbf{E}^T = -\mathbf{T}^2 = \begin{bmatrix} t_2^2 + t_3^2 & -t_1t_2 & -t_1t_3 \\ -t_2t_1 & t_3^2 + t_1^2 & -t_2t_3 \\ -t_3t_1 & -t_3t_2 & t_1^2 + t_2^2 \end{bmatrix} \tag{7.2}$$

La matrice $\mathbf{S} = \mathbf{E}\mathbf{E}^T$ a donc une forme particulière dans laquelle les trois éléments diagonaux et les trois éléments non diagonaux sont liés par les trois relations, obtenues avec $1 \leq i < j \leq 3$:

$$4S_{ij} - (\text{trace}(\mathbf{S}) - 2S_{ii})(\text{trace}(\mathbf{S}) - 2S_{jj}) = 0 \qquad (T_{ij}) \tag{7.3}$$

Nous obtenons cette fois-ci *trois* contraintes polynômiales, qui sont également de degré quatre en les coefficients de E, donc de degré huit en les coefficients de A, ce qui serait également paradoxal. Cependant Trivedi a montré que ces trois contraintes polynomiales ne sont pas indépendantes, mais se réduisent en deux équations indépendantes et une identité, si nous avons par ailleurs la nullité du déterminant de E, qui est acquise, comme nous l'avons déjà remarqué.

7.1.3 Equivalence des équations de Trivedi et des conditions de Huang-Faugeras

Nous montrons dans cette section que les deux contraintes algébriques que nous venons de présenter sont équivalentes.

Si les équations de Trivedi sont satisfaites, alors $\mathbf{S}$ a la forme donnée par l'équation (7.2). Cette forme entraîne clairement $\det(\mathbf{E}\mathbf{E}^T) = 0$, et donc $\det(\mathbf{E}) = 0$, ce qui est la première condition de Huang et Faugeras. Un calcul simple montre qu'en additionnant T_{12}, T_{13} et T_{23}, il vient:

$$4(S_{12}^2 + S_{13}^2 + S_{23}^2) + S_{11}^2 + S_{22}^2 + S_{33}^2 - 2(S_{11}S_{22} + S_{22}S_{33} + S_{33}S_{11}) = 0$$

La matrice $\mathbf{S}$ étant symétrique, nous pouvons remplacer le premier terme par l'expression: $4(S_{12}S_{21} + S_{13}S_{31} + S_{23}S_{32})$, ce qui conduit exactement à ce que l'on obtient en développant la seconde condition de Huang-Faugeras:

$$\text{trace}^2(\mathbf{S}) - 2\text{trace}(\mathbf{S}^2) = 0$$

Supposons à présent les conditions de Huang et Faugeras satisfaites, et montrons que nous obtenons alors les équations de Trivedi. Nous savons que les premières conditions sont équivalentes au fait que la matrice $\mathbf{E}$ ait une valeur singulière nulle et deux valeurs singulières σ non nulles égales. Par définition, il existe alors une matrice orthogonale $\mathbf{\Theta}$ telle que:

$$\mathbf{S} = \mathbf{E}\mathbf{E}^T = \mathbf{\Theta} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} \mathbf{\Theta}^T$$

Cette expression se développe en:

$$\mathbf{S} = \sigma^2(\Theta_{i2}\Theta_{j2} + \Theta_{i3}\Theta_{j3})_{1 \leq i,j \leq 3}$$

La matrice $\mathbf{\Theta}$ étant orthogonale, nous avons:

$$\Theta_{i2}\Theta_{j2} + \Theta_{i3}\Theta_{j3} = \begin{cases} -\Theta_{i1}\Theta_{j1} & \text{si } i \neq j \\ 1 - \Theta_{i1}^2 & \text{si } i = j \end{cases}$$

L'élément diagonal $1 - \Theta_{11}^2$ (respectivement $1 - \Theta_{21}^2$, $1 - \Theta_{31}^2$) se réécrit $\Theta_{31}^2 + \Theta_{21}^2$ (respectivement $\Theta_{11}^2 + \Theta_{31}^2$, $\Theta_{21}^2 + \Theta_{11}^2$), ce qui montre que $\mathbf{S}$ a exactement la forme donnée en (7.2).

L'équivalence que nous venons de montrer permet de voir que nous avons effectivement le bon nombre d'équations, la relation $\det(\mathbf{E}) = 0$ étant par ailleurs satisfaite. Cependant les deux équations de Trivedi ne conviennent pas mieux pour le calcul effectif des paramètres intrinsèques que la condition de Huang et Faugeras, dans la mesure où leur degré en les coefficients de $\mathbf{A}$, huit, est tout aussi élevé. C'est pour cette raison que nous allons nous tourner vers une expression géométrique des contraintes de rigidité, qui nous permettront d'obtenir deux équations de degré seulement deux en des coefficients possédant une signification géométrique claire, de degré quatre en les coefficients de $\mathbf{A}$.

7.2 Les équations de Kruppa

Nous introduisons un invariant projectif important, la conique absolue, qui nous permet d'exprimer géométriquement la rigidité d'un mouvement. La traduction algébrique de cette contrainte conduit aux équations dites de Kruppa, que nous établissons. Nous montrons ensuite par des méthodes algébriques, que les équations de Kruppa sont équivalentes à celles obtenues précédemment au moyen de la matrice essentielle. Pour finir, nous examinons deux problèmes liés au choix des mouvements: l'indépendance de ceux-ci, et la singularité de cas particuliers.

7.2.1 La conique absolue

La *conique absolue* Ω est un invariant projectif qui possède des propriétés importantes, liées à la structure euclidienne de l'espace. Elle est définie comme intersection de la quadrique d'équation $x_1^2+x_2^2+x_3^2+x_4^2 = 0$ avec le plan à l'infini Π_∞. Ω est donc définie par les équations:

$$x_1^2 + x_2^2 + x_3^2 = 0 \qquad x_4 = 0 \tag{7.4}$$

La conique Ω ne possède aucun point réel. Elle peut être interprétée comme un cercle de rayon i du plan à l'infini.

Montrons que Ω est laissée invariante par les déplacements rigides de l'espace. Soit $\tilde{\mathbf{M}}$ un point de Ω, il peut s'écrire $\begin{bmatrix} \mathbf{M} \\ 0 \end{bmatrix}$. Soit $\mathbf{D}$ un déplacement de l'espace, qui s'écrit: $\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$, $\mathbf{R}$ étant une matrice de rotation 3×3 et $\mathbf{t}$ un vecteur 3×1. L'image $\tilde{\mathbf{M}}'$ de $\tilde{\mathbf{M}}$ par $\mathbf{D}$ est alors: $\tilde{\mathbf{M}}' = \mathbf{D}\tilde{\mathbf{M}} = \begin{bmatrix} \mathbf{RM} \\ 0 \end{bmatrix}$ La quatrième coordonnée de $\tilde{\mathbf{M}}'$ est donc nulle. Il reste à vérifier que $\mathbf{M}'^T\mathbf{M}' = 0$. On a : $\mathbf{M}'^T\mathbf{M} = \mathbf{M}^T(\mathbf{R}^T\mathbf{R})\mathbf{M}$. Or $\mathbf{R}$ étant une matrice de rotation vérifie $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. Comme par hypothèse $\tilde{\mathbf{M}}$ appartient à Ω, on a $\mathbf{M}'^T\mathbf{M}' = \mathbf{M}^T\mathbf{M} = 0$, ce qui démontre l'égalité visée.

La conique absolue Ω est projetée par la caméra en une conique ω du plan rétinien. Montrons que sa projection ω ne dépend pas de la position ni de l'orientation de la caméra. Considérons une caméra. Les paramètres intrinsèques demeurant constants, nous pouvons considérer qu'elle est déterminée par la donnée de son centre optique $\mathbf{C}$ et de son plan rétinien $\mathcal{R}$. Soit ω l'image de Ω dans $\mathcal{R}$. Après un déplacement rigide D, la nouvelle position de la caméra est $(D(\mathbf{C}), D(\mathcal{R}))$ et la projection ω' observée est l'image de $D(\Omega)$. Or nous venons de voir que $D(\Omega) = \Omega$, d'où $\omega = \omega'$.

7.2.2 Les équations de Kruppa

Une conséquence géométrique de la rigidité L'introduction de la conique absolue nous permet d'obtenir une contrainte géométrique exprimant le fait que le mouvement entre les repères rétiniens des deux caméras est un déplacement rigide, et non une relation

projective linéaire quelconque. Cette contrainte a été inventée par Kruppa [80] au début du siècle, et réintroduite par Faugeras et Maybank lors de leur étude du nombre de solutions au problème du mouvement à partir de cinq appariements [42], puis appliquée au problème de l'auto-calibration dans une étude théorique [107] par ces mêmes auteurs[1].

Montrons que si on considère deux images correspondantes, les lignes épipolaires tangentes à ω doivent se correspondre. Soit Π le plan épipolaire contenant $\mathbf{CC}'$, et tangent à Ω. La tangence est préservée par projection sur une image, aussi la droite épipolaire l obtenue en projetant Π sur la première image est tangente à ω, la projection de Ω dans la première image. De même, la droite épipolaire l' dans la seconde image correspondant à l est tangente à la projection ω de Ω dans la seconde image. Cette situation est illustrée par la figure 7.1.

Il y a en tout deux plans Π_1, Π_2 tangents à Ω, aussi il y a deux paires de tangentes à ω correspondantes, l_1, l_2 dans la première image, l'_1, l'_2 dans la seconde image. Il n'est pas possible de les distinguer, et donc de déterminer si on a $l_1 \overline{\wedge} l'_1$ et $l_2 \overline{\wedge} l'_2$ ou bien $l_1 \overline{\wedge} l'_2$ et $l_2 \overline{\wedge} l'_1$. Cependant nous verrons que cela ne nous empêche pas de tirer les conséquences de ces correspondances.

Expression algébrique Introduisons $\mathbf{K}$, la matrice de la conique duale de ω. On a, par définition,

$$\mathbf{K} = \mathbf{B}^* \tag{7.5}$$

$\mathbf{B}^*$ étant l'adjointe de la matrice $\mathbf{B}$ de ω, c'est-à-dire la matrice définie par $B^*_{ij} = (-1)^{i+j} \det(\mathbf{B}^{ij})$, où $\mathbf{B}^{ij}$ est la matrice obtenue à partir de $\mathbf{B}$ en retranchant la ligne i et la colonne j. Si $\mathbf{B}$ est inversible, alors $\mathbf{B}^* = \det(\mathbf{B})\,\mathbf{B}^{T-1}$.

Nous utilisons la notation de Kruppa pour désigner les coefficients de $\mathbf{K}$, que nous appelons *coefficients de Kruppa*:

$$\mathbf{K} = \begin{bmatrix} -\delta_{23} & \delta_3 & \delta_2 \\ \delta_3 & -\delta_{13} & \delta_1 \\ \delta_2 & \delta_1 & -\delta_{12} \end{bmatrix} \tag{7.6}$$

Nous avons donc les relations:

$$\begin{aligned} \delta_1 &= b_{12}b_{13} - b_{11}b_{23} \\ \delta_2 &= b_{23}b_{21} - b_{22}b_{31} \\ \delta_3 &= b_{31}b_{32} - b_{33}b_{12} \end{aligned} \tag{7.7}$$

et

$$\begin{aligned} \delta_{12} &= b_{12}^2 - b_{11}b_{22} \\ \delta_{23} &= b_{23}^2 - b_{22}b_{33} \\ \delta_{31} &= b_{31}^2 - b_{33}b_{11} \end{aligned} \tag{7.8}$$

[1] Pour préciser ma contribution en ce qui concerne les équations de Kruppa proprement dites: les équations de Kruppa pour l'auto-calibration ont été portées à ma connaissance par Faugeras en Déc. 90 sous la forme où $\tau' = \rho\tau$ nécessitant l'emploi des coniques de Steiner. J'ai travaillé indépendamment de Maybank sur ces bases, sans parvenir à aboutir à un algorithme opérationnel, puis ai développé, également indépendamment, la formulation avec l'homographie générale qui apparaît section 5.4 de [107].

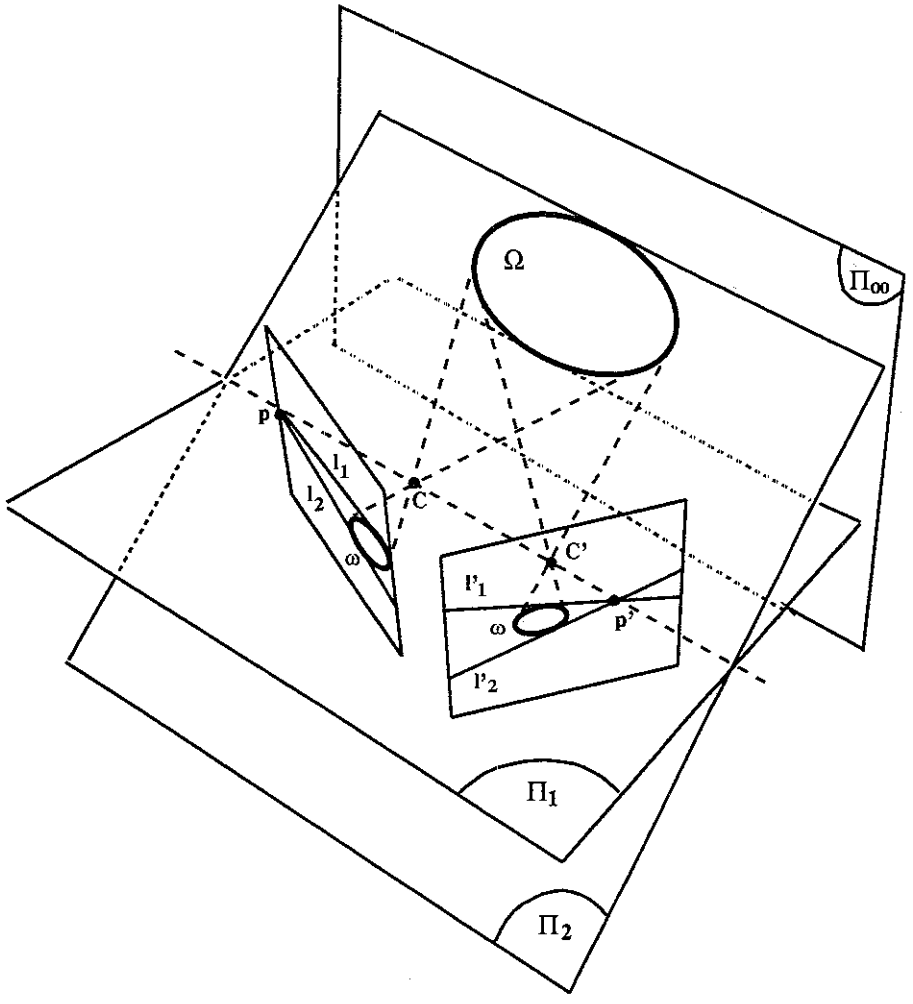


Fig. 7.1: La conique absolue et la transformation épipolaire

Soit $\mathbf{p}$ l'épipole dans la première image et $\mathbf{py}$ une droite épipolaire. La droite $\mathbf{py}$ est tangente à ω si et seulement si son représentant projectif appartient à la conique duale de ω, ce qui s'écrit:

$$(\mathbf{p} \wedge \mathbf{y})^T \mathbf{K} (\mathbf{p} \wedge \mathbf{y}) = 0 \tag{7.9}$$

Considérons un paramétrage des faisceaux d'épipolaires par intersection avec la droite à l'infini $y_3 = 0$. Le point d'intersection $\mathbf{y}_\infty$ a pour coordonnées $(1, \tau, 0)^T$. En substituant ces dernières coordonnées pour $\mathbf{y}$ et les coefficients de Kruppa (7.6) pour $\mathbf{K}$ dans l'équation (7.9), on obtient par des calculs simples une relation de la forme:

$$k_{11} + 2k_{12}\tau + k_{22}\tau^2 = 0 \tag{7.10}$$

où k_{11}, k_{12}, k_{22} sont des coefficients indépendants de $\mathbf{y}$, dont les valeurs sont:

$$\begin{aligned} k_{11} &= -\delta_{13}p_3^2 - \delta_{12}p_2^2 - 2\delta_1 p_2 p_3 \\ k_{12} &= \delta_{12}p_1 p_2 - \delta_3 p_3^2 + \delta_2 p_2 p_3 + \delta_1 p_1 p_3 \\ k_{22} &= -\delta_{23}p_3^2 - \delta_{12}p_1^2 - 2\delta_2 p_1 p_3 \end{aligned} \tag{7.11}$$

L'équation (7.10) traduit le fait que la droite épipolaire l de la première image, de paramètre τ soit tangente à ω. Il y a deux telles droites, comme nous l'avons remarqué plus haut, elles correspondent aux deux solutions de l'équation (7.10). On peut traduire exactement de la même manière le fait que la droite épipolaire l' de la seconde image, de paramètre τ' soit tangente à ω, et on obtient ainsi l'équation:

$$k'_{11} + 2k'_{12}\tau' + k'_{22}\tau'^2 = 0 \tag{7.12}$$

dans laquelle k'_{11}, k'_{12}, k'_{22} sont obtenus par les équations (7.11) en remplaçant les coordonnées p_1, p_2, p_3 de l'épipole $\mathbf{p}$ de la première image par les coordonnées p'_1, p'_2, p'_3 de l'épipole $\mathbf{p}'$ de la seconde image.

Ecrivons à présent que l et l' se correspondent. Ceci signifie que τ et τ' sont reliés par l'homographie épipolaire:

$$\tau' = \frac{a\tau + b}{c\tau + d} \tag{7.13}$$

En substituant cette dernière relation dans (7.12), on obtient l'équation:

$$k''_{11} + 2k''_{12}\tau + k''_{22}\tau^2 = 0 \tag{7.14}$$

où k''_{11}, k''_{12}, k''_{22} sont donnés par:

$$\begin{aligned} k''_{11} &= k'_{22}b^2 + k'_{11}d^2 + 2k'_{12}bd; \\ k''_{12} &= 2k'_{12}ad + 2k'_{22}ab + 2k'_{11}cd + 2k'_{12}bc \\ k''_{22} &= 2k'_{12}ac + k'_{22}a^2 + k'_{11}c^2 \end{aligned} \tag{7.15}$$

La condition que l et l' se correspondent est équivalente au fait que les équations (7.10) et (7.14) admettent des racines identiques. Les trinômes doivent donc être proportionnels. Les équations traduisant cette proportionnalité sont les équations de Kruppa:

$$\begin{aligned} k_{22}k''_{12} - k''_{22}k_{12} &= 0 \\ k_{11}k''_{12} - k''_{11}k_{12} &= 0 \\ k_{11}k''_{22} - k''_{11}k_{22} &= 0 \end{aligned} \tag{7.16}$$

De ces trois équations, seules deux sont indépendantes. Elles sont homogènes, de degré deux en les coefficients de Kruppa δ_i, δ_{ij}, et leurs coefficients ne dépendent que des sept paramètres de la transformation épipolaire entre les deux positions de la caméra.

Notons que des solutions δ_i, δ_{ij} correspondant bien à l'image d'une conique absolue doivent vérifier les contraintes traduisant le fait que la conique représentée par $\mathbf{B}$ ne comprenne aucun point réel. Cela est équivalent au fait que la conique duale représentée par $\mathbf{K}$ ne comprenne aucun point réel, puisque les tangentes sont complexes si, et seulement si la courbe l'est. Les tangentes sont complexes si le paramètre τ de l'équation (7.10) est lui-même complexe, la relation (7.13) garantissant qu'alors τ' est également complexe. Le discriminant réduit de l'équation (7.10) doit donc être strictement négatif, ce qui s'écrit:

$$k_{11}k_{22} - k_{12}^2 > 0$$

En substituant les valeurs (7.11) aux k_{ij}, puis en utilisant les expressions duales de (7.7) et (7.8), cette inégalité s'écrit finalement:

$$b_{11}p_1^2 + b_{22}p_2^2 + b_{33}p_3^2 + 2b_{12}p_1p_2 + 2b_{13}p_1p_3 + 2b_{23}p_2p_3 > 0$$

Ceci est bien cohérent avec le fait que la conique représentée par $\mathbf{B}$ ne comprenne aucun point réel. Nous discuterons plus précisément des conditions requises pour qu'il en soit bien ainsi, dans une section ultérieure.

7.2.3 Lien avec la matrice essentielle et les conditions de Huang-Faugeras

Les équations de Kruppa ont été obtenues à l'aide d'une contrainte géométrique qui est équivalente à la rigidité du déplacement d'une caméra. Or nous avons déjà signalé une autre manière d'exprimer cette contrainte: il s'agit des conditions de décomposabilité de la matrice essentielle comme produit d'une matrice de rotation et d'une matrice antisymétrique. Nous allons utiliser la matrice fondamentale pour montrer l'équivalence de ces conditions et des équations de Kruppa.

Les équations de Kruppa, telles que nous venons de les décrire, ont pour inconnues les paramètres intrinsèques, par l'intermédiaire des coefficients de Kruppa. Leurs coefficients sont obtenus à partir de la transformation épipolaire (épipoles et homographie), dont nous avons vu qu'elle se calcule simplement à partir de la matrice fondamentale. Nous obtenons ainsi

pour chaque déplacement, deux équations dépendant de la matrice fondamentale. D'un autre côté, nous pouvons obtenir la matrice essentielle à partir de la matrice fondamentale et de la matrice de paramètres intrinsèques par $\mathbf{E} = \mathbf{A}^T\mathbf{F}\mathbf{A}$. La rigidité du mouvement représenté par $\mathbf{F}$ avec les paramètres intrinsèques $\mathbf{A}$ est donc équivalente aux deux conditions, qui traduisent le fait que $\mathbf{E}$ est bien une matrice essentielle:

$$det(\mathbf{E}) = 0 \qquad f(\mathbf{E}) = \frac{1}{2}\text{trace}^2(\mathbf{E}\mathbf{E}^T) - \text{trace}(\mathbf{E}\mathbf{E}^T)^2 = 0$$

La première de ces conditions est toujours satisfaite, puisque $\mathbf{F}$, étant une matrice fondamentale, est de déterminant nul. Nous allons à présent prouver que la seconde condition est équivalente *aux deux* équations de Kruppa. Nous effectuons un changement de repère rétinien projectif dans chacun des deux plans rétiniens de sorte que la matrice fondamentale se présente sous forme diagonale. Une manière de voir que ces changements de repère sont toujours possibles est d'invoquer le théorème de décomposition en valeurs singulières: il existe deux matrices orthogonales $\mathbf{\Theta}$ et $\mathbf{\Delta}$ telles que $\mathbf{F} = \mathbf{\Delta}\mathbf{\Lambda}\mathbf{\Theta}^T$. Effectuons les changements de repères projectifs dont les matrices sont $\mathbf{\Theta}$ dans la première rétine et $\mathbf{\Delta}$ dans la seconde rétine. Cela revient, si $\mathbf{A}_0$ est la matrice de paramètres intrinsèques initiale, à considérer les nouvelles matrices de paramètres intrinsèques $\mathbf{A} = \mathbf{A}_0\mathbf{\Theta}$ et $\mathbf{A}' = \mathbf{A}_0\mathbf{\Delta}$ dans la première et la seconde rétine, respectivement. Si l'on avait la relation:

$$\mathbf{p}'^T\mathbf{A}_0^{-1T}\mathbf{F}\mathbf{A}_0^{-1}\mathbf{p} = 0$$

$\mathbf{p}$ et $\mathbf{p}'$ étant les coordonnées normalisées, on a, avec le choix qui vient d'être effectué:

$$\mathbf{p}'^T\mathbf{A}'^{-1T}\mathbf{\Lambda}\mathbf{A}^{-1}\mathbf{p} = 0$$

On peut donc supposer, si l'on considère deux matrices de paramètres intrinsèques différentes pour chacune des rétines, que la matrice fondamentale $\mathbf{F}$ est sous forme diagonalisée:

$$\mathbf{F} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{7.17}$$

avec $\lambda \neq 0$ et $\mu \neq 0$, puisqu'une matrice fondamentale est de rang 2.

Ces changements de repère permettent de simplifier considérablement les calculs, cependant comme ils sont en général différents dans chaque rétine, ils induisent des nouvelles matrices de paramètres intrinsèques qui sont aussi différentes. Il est cependant simple de tenir compte de cette petite difficulté. A cet effet, nous généralisons les équations de Kruppa telles qu'elles ont été présentées précédemment en considérant que la conique absolue peut avoir deux images différentes dans chacune des rétines. Ceci revient à introduire deux matrices de coefficients de Kruppa distinctes, obtenues à partir des deux matrices de paramètres intrinsèques distinctes par $\mathbf{K} = \mathbf{A}\mathbf{A}^T$ et $\mathbf{K}' = \mathbf{A}'\mathbf{A}'^T$, puis à réécrire l'équation (7.14) en remplaçant en plus les coefficients de Kruppa δ_i, δ_{ij} par les coefficients δ'_i, δ'_{ij} dans (7.11). Le reste du calcul reste inchangé. En utilisant la forme (7.17) de la matrice fondamentale,

on obtient les épipoles $\mathbf{e} = \mathbf{e}' = (0,0,1)^T$ et l'homographie $h : \tau \mapsto -\frac{\lambda}{\mu}\tau$, puis, tous calculs faits, les équations de Kruppa:

$$\begin{aligned} \lambda\delta_3\delta'_{23} + \mu\delta_{13}\delta'_3 &= 0 \qquad (E_1) \\ \lambda\delta_{23}\delta'_3 + \mu\delta_3\delta'_{13} &= 0 \qquad (E_2) \\ \lambda^2\delta_{23}\delta'_{23} - \mu^2\delta_{13}\delta'_{13} &= 0 \qquad (E_3) \end{aligned}$$

avec, $\mathbf{l}_1^T$, $\mathbf{l}_2^T$, $\mathbf{l}_3^T$ étant les vecteurs-ligne de $\mathbf{A}$ (des notations similaires sont également adoptées pour la seconde rétine avec des primes):

$$\begin{aligned} \delta_3 &= \langle \mathbf{l}_1, \mathbf{l}_2 \rangle \\ \delta_{13} &= -\|\mathbf{l}_2\|^2 \\ \delta_{23} &= -\|\mathbf{l}_1\|^2 \end{aligned} \tag{7.18}$$

Nous avons présenté les trois équations de Kruppa par commodité, mais comme toujours, il n'y en a que deux qui sont indépendantes. On a dans notre cas la relation:

$$\lambda\delta_{23}E_1 - \mu\delta_{13}E_2 = \delta_3 E_3 \tag{7.19}$$

Exprimons à présent la seconde condition $f(\mathbf{E}) = 0$ sur la matrice essentielle. Cette matrice s'écrit dans notre cas: $\mathbf{E} = \mathbf{A}'^T\mathbf{F}\mathbf{A}$. On trouve finalement, après des calculs effectués en partie grâce au calculateur formel MAPLE, que nous ne détaillons pas:

$$\begin{aligned} f(\mathbf{E}) &= -\frac{1}{2}((\lambda^2\delta_{23}\delta'_{23} - \mu^2\delta_{13}\delta'_{13})^2 + 2\lambda\mu(\lambda\delta_3\delta'_{23} + \mu\delta_{13}\delta'_3)(\lambda\delta_{23}\delta'_3 + \mu\delta_3\delta'_{13})) \\ &= -\frac{1}{2}(E_3^2 + 2\lambda\mu E_1 E_2) \end{aligned}$$

Il donc est clair que les équations de Kruppa entraînent $f(\mathbf{E}) = 0$. Il reste à montrer la réciproque.

Considérons d'abord le cas où $\delta_3 \neq 0$. L'équation précédente peut alors se récrire, compte tenu de (7.19):

$$(\lambda\delta_{23}E_1 - \mu\delta_{13}E_2)^2 + 2\lambda\mu E_1 E_2 \delta_3^2 = 0 \tag{7.20}$$

On en tire:

$$\lambda^2\delta_{23}^2E_1^2 + \mu^2\delta_{13}^2E_2^2 = 2\lambda\mu E_1E_2(\delta_{13}\delta_{23} - \delta_3^2) \tag{7.21}$$

Or, d'après les définitions (7.18) de δ_3, δ_{13}, δ_{23}, l'inégalité de Schwartz implique que le facteur $\delta_{13}\delta_{23} - \delta_3^2$ est toujours positif ou nul. S'il est nul, on déduit de (7.21) que $\delta_{23}E_1 = \delta_{13}E_2 = 0$. Comme alors $\delta_{13}\delta_{23} = \delta_3^2 \neq 0$, on obtient $E_1 = E_2 = 0$. S'il est strictement positif, on en déduit que $2\lambda\mu E_1E_2 \geq 0$. L'équation (7.20) est donc la somme de deux termes positifs, par conséquent ils doivent s'annuler tous les deux, d'où $E_1E_2 = 0$ et $E_3 = 0$.

Il ne nous reste plus qu'à examiner le cas particulier où $\delta_3 = 0$. Les équations de Kruppa se réduisent alors à:

$$\mu\delta_{13}\delta'_3 = \lambda\delta_{23}\delta'_3 = \lambda^2\delta_{23}\delta'_{23} - \mu^2\delta_{13}\delta'_{13} = 0$$

ce qui est équivalent à

$$\begin{cases} \delta'_3 = 0 \\ \lambda^2\delta_{23}\delta'_{23} - \mu^2\delta_{13}\delta'_{13} = 0 \end{cases} \quad \text{ou} \quad \begin{cases} \delta'_3 \neq 0 \\ \delta_{13} = \delta_{23} = 0 \end{cases}$$

Il est immédiat de voir que ce dernier système est également équivalent à:

$$f(\mathbf{E}) = 2\lambda^2\mu^2\delta_{13}\delta_{23}\delta_3'^2 + (\lambda^2\delta_{23}\delta'_{23} - \mu^2\delta_{13}\delta'_{13})^2 = 0$$

ce qui achève la démonstration.

7.2.4 Combien de mouvements sont nécessaires ?

Chaque mouvement permet d'obtenir deux équations algébriques. L'image de la conique absolue dépend de cinq paramètres. Dans le cas général il faut donc un minimum de trois mouvements. Dans le cas du modèle de caméra simplifié à quatre paramètres intrinsèques, deux mouvements doivent être suffisants, puisque l'on a la contrainte additionnelle (7.37). Les expériences confirment ces décomptes. Une étude théorique des équations de Kruppa effectuée par Maybank et Faugeras permet au moyen de techniques de géométrie algébrique d'arriver à ces mêmes conclusions. Nous présenterons un résumé de cette approche plus bas.

Une question intéressante est de savoir s'il est nécessaire de disposer de trois mouvements indépendants, ou si les équations peuvent être résolues même si l'un des mouvements est obtenu par composition des deux autres. Ce détail est important en pratique, puisque dans la première alternative, il est nécessaire d'effectuer trois mouvements et d'utiliser quatre images, alors que dans la seconde alternative, il est suffisant d'utiliser seulement trois images, et de considérer les déplacements 1-2, 2-3, et 1-3. Une manière d'examiner cette question est de considérer deux déplacements $\mathbf{D}_1$ et $\mathbf{D}_2$ et le déplacement composé $\mathbf{D}_3 = \mathbf{D}_1\mathbf{D}_2$. Une condition pour que le système d'équations de Kruppa obtenues avec les trois déplacements $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{D}_3$ admette une solution unique est que les équations de Kruppa obtenues avec $\mathbf{D}_3$ ne puissent être déduites de celles obtenues avec $\mathbf{D}_1$ et de celles obtenues avec $\mathbf{D}_2$. Dans le cas général, cette propriété est vérifiée. Il n'est en effet pas possible de déduire la troisième transformation épipolaire (ou matrice fondamentale) à partir des deux premières, puisque les paramètres intrinsèques sont inconnus. Même si les paramètres intrinsèques sont connus, il demeure l'incertitude sur les normes des translations, qui ne permet pas d'obtenir la direction de la translation composée. Par conséquent l'information contenue dans la troisième transformation épipolaire est indépendante et les équations de Kruppa devraient l'être également. Nous donnons un exemple numérique pour lequel cette propriété est vérifiée avec des valeurs simples. Les deux mouvements de base sont:

$$\mathbf{R}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \mathbf{t}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{R}_2 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{t}_2 = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

Le mouvement obtenu par composition de $\mathbf{D}_1$ et $\mathbf{D}_2$, dans le repère de la première caméra, est:

$$\mathbf{R}_3 = \mathbf{R}_1\mathbf{R}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix} \qquad \mathrm{t}_3 = \mathbf{R}_1\mathrm{t}_2 + \mathrm{t}_1 = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$$

En prenant pour matrice de paramètres intrinsèques $\mathbf{A}$ l'identité, les matrices fondamentales sont simplement les matrices essentielles, et en normalisant les coefficients de Kruppa par $\delta_{12} = 1$, il vient les six équations de Kruppa:

$$\begin{aligned}
E_1 &= 3\,\delta_1 - 2 + 6\,\delta_3\,\delta_{13} + 9\,\delta_3 + 4\,\delta_1\,\delta_3 - 7\,\delta_2\,\delta_{13} + 2\,\delta_2 + 12\,\delta_1\,\delta_2 + 3\,\delta_1\,\delta_{13} + 2\,{\delta_{13}}^2 \\
E'_1 &= 3\,\delta_{23} - 3\,\delta_{13}\,\delta_{23} + 8\,\delta_1\,\delta_{23} + 1 + \delta_1 \\
&\quad - \delta_2\,\delta_{13} - 4\,\delta_2 - 4\,\delta_1\,\delta_2 - 4\,\delta_3\,\delta_{13} - \delta_3 + 4\,\delta_1\,\delta_3 - {\delta_{13}}^2 + \delta_1\,\delta_{13} \\
E_2 &= 2\,\delta_3\,\delta_{23} + 16\,\delta_3 - 8\,\delta_2\,\delta_3 + 4\,\delta_2\,\delta_{23} \\
&\quad + 16\,\delta_2 - 16\,{\delta_2}^2 + 4\,\delta_1\,\delta_{13} + 16\,\delta_1 - 16\,{\delta_1}^2 + 2\,\delta_3\,\delta_{13} - 8\,\delta_1\,\delta_3 \\
E'_2 &= {\delta_{23}}^2 + 4\,\delta_{23} - 4\,\delta_2\,\delta_{23} - {\delta_{13}}^2 - 4\,\delta_{13} + 4\,\delta_1\,\delta_{13} \\
E_3 &= 6\,\delta_{23} + 6\,\delta_3 + 18\,\delta_{23}\,\delta_3 + 12\,\delta_3\,\delta_{13} + 36\,{\delta_3}^2 + 18\,\delta_{23}\,\delta_2 + 36\,\delta_2\,\delta_{13} + 36\,\delta_2\,\delta_3 \\
&\quad - 6\,\delta_{23}\,\delta_1 - 12\,\delta_1\,\delta_{13} + 18\,\delta_2 - 36\,\delta_1\,\delta_2 - 6\,{\delta_{13}}^2 - 18\,\delta_1 + 36\,{\delta_1}^2 \\
E'_3 &= 9\,{\delta_{23}}^2 + 9\,\delta_{23}\,\delta_{13} + 18\,\delta_{23}\,\delta_3 + \delta_{23} - 9\,\delta_{13} + 2\,\delta_3 + 6\,\delta_{23}\,\delta_2 + 6\,\delta_2\,\delta_{13} + 12\,\delta_2\,\delta_3 \\
&\quad - {\delta_{13}}^2 - 4\,\delta_1\,\delta_{13} - 9 + 12\,\delta_1 + 12\,{\delta_1}^2
\end{aligned}$$

Une solution particulière du système formé par les équations E_1, E'_1, E_2, E'_2 provenant des déplacements $\mathbf{D}_1$ et $\mathbf{D}_2$ est:

$$\delta_1 = 0 \quad \delta_2 = -\frac{1}{2} \quad \delta_3 = 1 \quad \delta_{13} = -4 \quad \delta_{23} = 0$$

On vérifie aisément en substituant ces valeurs dans les équations E_3 et E'_3 provenant du déplacement composé $\mathbf{D}_3$ qu'elles ne vérifient pas ces dernières, puisque nous obtenons $E_3 = -27$, $E'_3 = 19$.

7.2.5 Cas dégénérés

Nous examinons les mouvements constitués d'une rotation pure et ceux constitués d'une translation pure et nous montrons que les équations de Kruppa ne donnent aucune information dans ces deux cas.

Considérons d'abord un déplacement constitué par une rotation autour d'un axe passant par le centre optique de la caméra. La translation est le vecteur nul: $\mathrm{t} = (0, 0, 0)^T$, la rotation étant quelconque. Au cours d'un tel déplacement, pour lequel la matrice fondamentale est nulle, il n'y a pas de contrainte épipolaire puisque les deux centres optiques restent confondus. A fortiori, les équations de Kruppa n'ont donc pas de sens.

Considérons à présent le cas des translations pures. La rotation est l'identité $\mathbf{R} = \mathbf{I}_3$, la translation est un vecteur quelconque: $\mathbf{t} = (t_1, t_2, t_3)^T$. Soit $\mathbf{A}_0$, la matrice de paramètres intrinsèques, nous posons pour simplifier les notations $\mathbf{A} = \mathbf{A}_0^{-1}$. A partir de la matrice fondamentale $\mathbf{F} = \mathbf{A}^T\mathbf{T}\mathbf{A}$, qui, dans ce cas, est antisymétrique, nous obtenons des épipoles $\mathbf{e}$ et $\mathbf{e}'$ qui sont identiques entre les deux images, ainsi que l'homographie identité: $\tau \mapsto \tau$. Il est assez simple de voir qu'alors les deux équations de Kruppa sont identiquement nulles. En effet, ces équations sont obtenues en écrivant que les deux tangentes à l'image de la conique absolue sont correspondantes par la transformation épipolaire. Or la transformation épipolaire étant composée d'éléments identiques dans le cas d'une translation, les deux tangentes obtenues dans la première image sont confondues avec celles obtenues dans la seconde image. Par conséquent il n'est pas possible d'en tirer de l'information. Algébriquement, les coefficients k'_{ij} de (7.12) sont égaux aux coefficients k_{ij} de (7.10) à cause de l'égalité des épipoles, les coefficients k''_{ij} de (7.14) sont égaux aux coefficients k'_{ij} à cause de l'identité de l'homographie, si bien que les trinômes définis dans ces équations sont toujours égaux. En écrivant qu'ils sont proportionnels, nous n'obtenons donc qu'une équations identiquement nulle.

Un autre cas qui n'est pas intrinsèquement dégénéré, mais l'est avec le paramétrage des droites épipolaires que nous avons adopté, est celui où l'un des épipoles se trouve à l'infini. Cette considération a également une grande importance pratique car elle signifie que lorsque nous obtenons des épipoles qui sont très éloignés de l'image, si nous utilisons les équations de Kruppa telles qu'elles ont été établies précédemment, nous obtiendrons des résultats extrêmement instables, dus au fait que les tangentes à la conique absolue étant très proches de lignes parallèles, leur paramétrage par la direction conduit à des valeurs toujours très proches.

7.3 Obtention des paramètres intrinsèques

Nous mettons d'abord en évidence le lien algébrique qui existe entre la conique absolue et les paramètres intrinsèques, ce qui nous permet de retrouver le fait que le modèle de caméra que nous considérons est général. Nous examinons ensuite le problème du calcul des paramètres intrinsèques à partir des coefficients obtenus en résolvant les équations de Kruppa, et discutons des contraintes provenant du fait que la conique absolue soit purement imaginaire.

7.3.1 Lien entre la conique absolue et les paramètres intrinsèques

Nous pouvons conclure de l'invariance de ω vis à vis des déplacements de la caméra que ω est déterminé uniquement par Ω et les paramètres intrinsèques de la caméra. En conséquence, nous pouvons choisir la position du repère du monde de manière arbitraire pour effectuer les calculs. Prenons alors pour repère du monde, le repère normalisé de la caméra. La matrice de projection $\mathbf{P}$ se décompose alors en $[\mathbf{A}\ \mathbf{0}]$, $\mathbf{A}$ étant la matrice des paramètres intrinsèques et $\mathbf{0}$ le vecteur 3×1 à coordonnées nulles. Soit $\mathbf{M}$ un point de la conique absolue Ω. Comme $\mathbf{M}$

vérifie l'équation (7.4), il peut s'écrire $\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ 0 \end{bmatrix}$ avec $\mathbf{M}_1^T\mathbf{M}_1 = 0$. Considérons le point $\mathbf{m}$, image de $\mathbf{M}$ par la caméra. On a $\mathbf{m} = \mathbf{PM} = \mathbf{AM}_1$, d'où $\mathbf{M}_1 = \mathbf{A}^{-1}\mathbf{m}$. En reportant cette dernière expression dans $\mathbf{M}_1^T\mathbf{M}_1 = 0$, on obtient l'équation de ω:

$$\mathbf{m}^T\mathbf{A}^{-1T}\mathbf{A}^{-1}\mathbf{m} = 0$$

ω est une conique plane dont la matrice 3×3 est:

$$\mathbf{B} = \mathbf{A}^{-1T}\mathbf{A}^{-1} \tag{7.22}$$

L'équation (7.22) montre que ω est déterminée par les paramètres intrinsèques, et vice-versa.

Une conique plane est déterminée par cinq paramètres indépendants, aussi le modèle de caméra qui est à considérer est celui à cinq paramètres intrinsèques. En utilisant celui-ci, nous obtenons la matrice $\mathbf{B}$ de ω, qui est définie à un facteur d'échelle près:

$$\begin{bmatrix} \alpha_v^2 & \alpha_u\alpha_v\cos\theta & -\alpha_v(\alpha_u cos\theta v_0 + \alpha_v u_0) \\ \alpha_u\alpha_v\cos\theta & \alpha_u^2 & -\alpha_u(\alpha_v cos\theta u_0 + \alpha_u v_0) \\ -\alpha_v(\alpha_u cos\theta v_0 + \alpha_v u_0) & -\alpha_u(\alpha_v cos\theta u_0 + \alpha_u v_0) & 2\alpha_u\alpha_v u_0 v_0\cos\theta + \alpha_v^2 u_0^2 + \alpha_u^2 v_0^2 + \alpha_u^2\alpha_v^2 \end{bmatrix} \tag{7.23}$$

L'équation de l'image ω de la conique absolue en coordonnées rétiniennes, obtenue à partir de cette expression, est:

$$\frac{(u-u_0)^2}{\alpha_u^2} + \frac{(v-v_0)^2}{\alpha_v^2} + 2\frac{(u-u_0)(v-v_0)\cos\theta}{\alpha_u\alpha_v} + 1 = 0 \tag{7.24}$$

Cette dernière égalité nous donne la relation entre les paramètres intrinsèques et l'équation de l'image de la conique absolue ω.

A titre d'image de Ω, qui est une conique entièrement complexe, ω doit être également purement complexe. En posant $u_1 = \frac{(u-u_0)}{\alpha_u}$ et $v_1 = \frac{(v-v_0)}{\alpha_v}$, nous pouvons réécrire (7.24) sous la forme:

$$u_1^2 + v_1^2 + 2cos\theta u_1 v_1 = -1$$

En utilisant $(u_1 - v_1)^2 \geq 0$, il vient immédiatement $2|cos\theta u_1 v_1| \leq 2|u_1 v_1| \leq u_1^2 + v_1^2$, ce qui montre que la conique définie par (7.24) ne comporte effectivement aucun point réel. Il est en outre facile de voir que ω admet pour centre (u_0, v_0), puisque dans le repère centré en ce point, ω ne comprend pas de terme linéaire. En revanche les directions principales n'ont pas d'expression très simple.

Montrons à présent la proposition réciproque: toute conique $\mathcal{C}$ de $\mathcal{P}^2$ ne comportant aucun point réel admet une équation de la forme (7.24). De toutes les coniques, les seules susceptibles de ne pas posséder de points réels sont les ellipses, c'est-à-dire les coniques admettant l'équation:

$$a_1^2x^2 + a_2^2y^2 + a_3^2 = 0 \tag{7.25}$$

dans un repère particulier, celui des directions principales de la conique, avec pour origine le centre de l'ellipse. Si (u_c, v_c) est le centre de l'ellipse dans le repère initial des coordonnées

rétiniennes, et si les directions principales sont obtenues par une rotation d'angle ϕ à partir des deux axes du repère initial, la relation entre les coordonnées (x, y) du repère particulier et les coordonnées rétiniennes (u, v) est:

$$\begin{aligned} x &= (u - u_c)\cos\phi - (v - v_c)\sin\phi \\ y &= (u - u_c)\sin\phi + (v - v_c)\cos\phi \end{aligned} \tag{7.26}$$

Montrons que la forme quadratique $\mathbf{Q}$ associée à $\mathcal{C}$, ainsi obtenue, peut s'écrire sous la forme (7.24). Nous obtenons, après substitution de (7.26) dans (7.25) et identification des coefficients des termes quadratiques provenant de l'équation ainsi obtenue avec les coefficients tirés de (7.24), les trois équations:

$$\begin{aligned} \alpha_v^2 &= Q_{11} = a_1^2\cos^2\phi + a_2^2\sin^2\phi \\ \alpha_u^2 &= Q_{22} = a_1^2\sin^2\phi + a_2^2\cos^2\phi \\ \alpha_u\alpha_v\cos\theta &= Q_{12} = (a_2^2 - a_1^2)\cos\phi\sin\phi \end{aligned} \tag{7.27}$$

Il est clair qu'étant donné a_1, a_2, et ϕ, ce système admet une solution si, et seulement si la substitution des deux premières équations dans la troisième détermine une valeur de $\cos\phi$ qui est comprise entre -1 et 1. Il nous faut donc vérifier que:

$$|(a_2^2 - a_1^2)\cos\phi\sin\phi| \leq \sqrt{(a_1^2\cos^2\phi + a_2^2\sin^2\phi)(a_1^2\sin^2\phi + a_2^2\cos^2\phi)}$$

Nous remarquons que cette relation est équivalente à:

$$det(\mathbf{Q}) = Q_{11}Q_{22} - Q_{12}^2 \geq 0$$

Elle est donc toujours satisfaite, puisque par hypothèse $\mathcal{C}$, est une ellipse. Nous avons donc déterminé des paramètres α_u, α_v, $\cos\theta$ tels que la partie quadratique de $\mathcal{C}$ ait la forme (7.24) souhaitée. Une ellipse est définie par la donnée des cinq paramètres constitués par les trois coefficients de sa forme quadratique associée et les deux coordonnées de son centre. Les trois premiers paramètres sont identifiés grâce à (7.27). Nous n'avons qu'à prendre:

$$u_0 = u_c \qquad v_0 = v_c \tag{7.28}$$

pour achever d'exprimer $\mathcal{C}$ sous la forme (7.24). Une autre manière de voir ceci est de prendre des paramètres intrinsèques satisfaisant (7.27) et (7.28), et de chercher à identifier la forme linéaire associée à $\mathcal{C}$:

$$\begin{bmatrix} -(a_1^2\cos^2\phi + a_2^2\sin^2\phi)u_c - (a_2^2 - a_1^2)\cos\phi\sin\phi v_c \\ -(a_1^2\sin^2\phi + a_2^2\cos^2\phi)v_c - (a_2^2 - a_1^2)\cos\phi\sin\phi u_c \\ 2(a_2^2 - a_1^2)u_cv_c\cos\phi\sin\phi + u_c^2(a_1^2\cos^2\phi + a_2^2\sin^2\phi) + v_c^2(a_1^2\sin^2\phi + a_2^2\cos^2\phi) + a_3^2 \end{bmatrix}^T$$

avec la troisième colonne de (7.23). Il est clair que les deux premiers éléments vérifient l'égalité attendue. La troisième égalité est vérifiée pour autant que la condition $\alpha_u^2\alpha_v^2 = a_3^2$ soit satisfaite. Cette condition définit une relation entre a_1, a_2, et a_3 qui peut toujours

être vérifiée pour un certain choix de a_3, l'équation (7.25) n'étant définie qu'à un facteur multiplicatif près.

Nous savons que l'image d'une conique absolue est nécessairement une ellipse sans point réel. Ayant montré que toute ellipse de ce type peut s'écrire sous la forme (7.24), issue de notre modèle de caméra, nous retrouvons le fait que ce modèle est entièrement général, au sens où il suffit donc pour décrire tous les changements de coordonnées rétiniennes possibles.

7.3.2 Liens entre paramètres intrinsèques et coefficients de Kruppa

Cherchons à présent à exprimer la relation entre les paramètres intrinsèques et les coefficients de Kruppa. En utilisant les relations (7.5) et (7.22), il vient, puisque $\mathbf{K}$ est définie à un facteur d'échelle près seulement: $\mathbf{K} = \mathbf{A}\mathbf{A}^T$. En utilisant la définition (7.6) de $\mathbf{K}$, on obtient alors à un facteur d'échelle λ près:

$$\lambda\delta_1 = v_0 \tag{7.29}$$

$$\lambda\delta_2 = u_0 \tag{7.30}$$

$$\lambda\delta_3 = u_0 v_0 - \alpha_u \alpha_v \frac{\cot\theta}{\sin\theta} \tag{7.31}$$

$$\lambda\delta_{12} = -1 \tag{7.32}$$

$$\lambda\delta_{23} = -u_0^2 - \frac{\alpha_u^2}{\sin^2\theta} \tag{7.33}$$

$$\lambda\delta_{13} = -v_0^2 - \frac{\alpha_v^2}{\sin^2\theta} \tag{7.34}$$

En utilisant (7.30), (7.33), et (7.32) il vient:

$$\alpha_u^2 = \frac{\delta_{23}\delta_{12} - \delta_2^2}{\delta_{12}^2}\sin^2\theta$$

et de même avec (7.29), (7.34), et (7.32):

$$\alpha_v^2 = \frac{\delta_{13}\delta_{12} - \delta_1^2}{\delta_{12}^2}\sin^2\theta$$

En substituant ces deux relations dans (7.31), nous obtenons:

$$\cos^2\theta = \frac{(\delta_3\delta_{12} + \delta_1\delta_2)^2}{(\delta_{13}\delta_{12} - \delta_1^2)(\delta_{23}\delta_{12} - \delta_2^2)}$$

Nous pouvons en déduire que le système constitué des équations (7.29) à (7.34) admet une solution avec des paramètres intrinsèques réels si, et seulement si les trois conditions suivantes sont satisfaites:

$$\begin{cases} \delta_{13}\delta_{12} - \delta_1^2 > 0 \\ \delta_{23}\delta_{12} - \delta_2^2 > 0 \\ (\delta_{13}\delta_{12} - \delta_1^2)(\delta_{23}\delta_{12} - \delta_2^2) \geq (\delta_3\delta_{12} + \delta_1\delta_2)^2 \end{cases} \tag{7.35}$$

De ces trois conditions, seules deux sont indépendantes. En effet, les trois conditions ont la forme:

$$a > 0 \ , \ c > 0 \ , \ ac - b^2 > 0$$

or, la troisième condition entraîne $ac > 0$, et donc $a > 0 \Leftrightarrow c > 0$. Etant donné la troisième condition, les deux premières sont équivalentes. Nous pouvons aussi remarquer que les conditions (7.35) entraînent également:

$$\delta_{13}\delta_{12} > 0 \quad \delta_{23}\delta_{12} > 0$$

qui doivent visiblement être satisfaites au simple regard des équations (7.32), (7.33) et (7.34).

Après calculs, nous trouvons que les paramètres intrinsèques sont obtenus par les relations suivantes :

$$\begin{aligned}
u_0 &= -\frac{\delta_2}{\delta_{12}} \\
v_0 &= -\frac{\delta_1}{\delta_{12}} \\
\alpha_u &= \varepsilon\sqrt{\frac{\delta_{23}\delta_1^2 + \delta_{13}\delta_2^2 + \delta_{12}\delta_3^2 + 2\delta_1\delta_2\delta_3 - \delta_{12}\delta_{23}\delta_{13}}{\delta_{12}(\delta_1^2 - \delta_{13}\delta_{12})}} \\
\alpha_v &= \varepsilon\sqrt{\frac{\delta_{23}\delta_1^2 + \delta_{13}\delta_2^2 + \delta_{12}\delta_3^2 + 2\delta_1\delta_2\delta_3 - \delta_{12}\delta_{23}\delta_{13}}{\delta_{12}(\delta_2^2 - \delta_{23}\delta_{12})}} \\
\cos\theta &= \frac{\delta_3\delta_{12} + \delta_2\delta_1}{\sqrt{(\delta_{23}\delta_{12} - \delta_2^2)(\delta_{13}\delta_{12} - \delta_1^2)}}
\end{aligned} \tag{7.36}$$

Il y a unicité dans leur détermination, excepté pour ce qui est du signe des paramètres α_u et α_v. On peut aussi remarquer à partir de ces dernières relations que les paramètres intrinsèques ne dépendent bien que du rapport des coefficients de Kruppa, obtenus, rappelons-le, comme solutions d'équations homogènes. Une seconde remarque est qu'il est très simple de se restreindre au modèle à quatre paramètres (celui sans correction de la perpendicularité). On a alors $\theta = \frac{\pi}{2}$, ce qui introduit la contrainte supplémentaire

$$\delta_3\delta_{12} + \delta_2\delta_1 = 0 \tag{7.37}$$

et les simplifications:

$$\alpha_u = \varepsilon\frac{\sqrt{\delta_{23}\delta_{12} - \delta_2^2}}{\delta_{12}} \tag{7.38}$$

$$\alpha_v = \varepsilon\frac{\sqrt{\delta_{13}\delta_{12} - \delta_1^2}}{\delta_{12}} \tag{7.39}$$

Montrons pour finir que les conditions (7.35) sont en fait équivalentes au fait que la conique de matrice $\mathbf{B} = \mathbf{K}^*$ soit complexe. Une conique générale C a pour équation:

$$P(x,y) = b_{11}x^2 + 2b_{12}xy + b_{22}y^2 + 2b_{13}x + 2b_{23}y + b_{33} = 0 \tag{7.40}$$

Il est connu que la condition pour que (7.40) définisse une ellipse est:

$$b_{11}b_{22} - b_{12}^2 > 0 \tag{7.41}$$

L'équation (7.40) peut dans se cas se réécrire, sous la forme:

$$P_1(x, y) = \lambda_1 x_1^2 + \lambda_2 y_1^2 + P(x_0, y_0) = 0 \tag{7.42}$$

dans un nouveau repère centré en:

$$x_0 = \frac{b_{12}b_{23} - b_{13}b_{22}}{b_{11}b_{22} - b_{12}^2} \quad y_0 = \frac{b_{12}b_{13} - b_{11}b_{23}}{b_{11}b_{22} - b_{12}^2}$$

et ayant pour axes les directions principales de $\mathcal{C}$, c'est à dire les vecteurs propres de la matrice de la forme quadratique associée à $\mathcal{C}$,:

$$\mathbf{Q} = \begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix}$$

les coefficients λ_i étant les valeurs propres de la matrice $\mathbf{Q}$. La condition (7.41) garantit que $\operatorname{sgn} \lambda_1 = \operatorname{sgn} \lambda_2$, aussi l'équation (7.42) n'admet aucune solution réelle si, et seulement si:

$$\operatorname{sgn} \lambda_i = \operatorname{sgn} P(x_0, y_0) \tag{7.43}$$

Explicitons cette condition. Comme $\operatorname{sgn} \lambda_1 = \operatorname{sgn} \lambda_2$, il vient $\operatorname{sgn} \lambda_i = \operatorname{sgn} (b_{11} + b_{22})$. De plus, la condition (7.41) entraîne $b_{11}b_{22} > 0$, d'où

$$\operatorname{sgn} \lambda_i = \operatorname{sgn} b_{11} = \operatorname{sgn} b_{22}$$

Compte tenu de (7.41), $\operatorname{sgn} P(x_0, y_0) = \operatorname{sgn} P(b_{12}b_{23} - b_{13}b_{22}, b_{12}b_{13} - b_{11}b_{23})$. En substituant dans cette dernière expression les valeurs des b_{ij} obtenues à partir de la relation $\mathbf{B} = \mathbf{K}^*$, il vient, après quelques calculs:

$$\operatorname{sgn} P(x_0, y_0) = \operatorname{sgn} (\delta_1^2\delta_{23} + \delta_2^2\delta_{13} + \delta_3^2\delta_{12} + 2\delta_1\delta_3\delta_2 - \delta_{12}\delta_{23}\delta_{13})^2 = +1$$

La condition (7.43) se simplifie donc en: $b_{11} > 0$, et nous avons vu que compte tenu de (7.41), elle équivaut à $b_{11} > 0$ et $b_{22} > 0$. Compte tenu des valeurs

$$\begin{aligned} b_{11} &= \delta_{13}\delta_{12} - \delta_1^2 \\ b_{22} &= \delta_{23}\delta_{12} - \delta_2^2 \\ b_{12} &= \delta_3\delta_{12} + \delta_1\delta_2 \end{aligned}$$

nous avons donc montré que les conditions (7.35) sont équivalentes aux deux conditions (7.41) et (7.43) traduisant le fait que la conique de matrice $\mathbf{B} = \mathbf{K}^*$ soit complexe.

Chapitre 8

Le calcul des paramètres intrinsèques: expériences

Maybank et Faugeras [107] ont montré, sous forme d'un théorème de géométrie algébrique, que l'utilisation des équations de Kruppa résultant de trois mouvements permet de retrouver les paramètres intrinsèques. Nous commençons par résumer rapidement cette approche. Elle conduit à une méthode presque entièrement analytique de résolution dont nous avons constaté qu'elle n'était pas utilisable autrement que de manière théorique.

Nous examinons alors en premier une approche semi-analytique pour la résolution des équations de Kruppa, par la méthode des continuations, adaptée aux systèmes d'équations polynomiales, que nous présentons rapidement avant d'expliquer la manière dont nous l'utilisons.

Une seconde approche est d'utiliser des techniques itératives purement numériques: minimisation non-linéaire et filtrage de Kalman. Elle permet d'exploiter des longues séquences d'images, ainsi que le calcul d'incertitude et la connaissance *a priori* de caractéristiques de la caméra. Les résultats sont évalués au moyen d'une large étude comparative utilisant des statistiques sur les simulations. Quelques exemples réels obtenus avec des images variées illustrent également l'applicabilité de notre méthode.

8.1 Une étude théorique: approche algébrique

Nous résumons dans cette section l'approche de Maybank et Faugeras [107] Cette approche n'est pas utilisable en pratique Nous l'avons implémentée partiellement en MAPLE et constaté que les calculs doivent nécessairement être effectués avec des représentations sous forme de nombres rationnels pour obtenir une précision adéquate Avec des représentations flottantes des nombres, une précision de 16 digits n'est pas suffisante pour obtenir des résultats cohérents Cependant cette approche présente un intérêt théorique, qui est de d'établir des propriétés mathématiques et d'obtenir des interprétations géométriques

L'idée est que deux mouvements conduisant à quatre contraintes sur l'image ω de la conique absolue, et celle-ci étant déterminée par cinq paramètres, les coniques ω compatibles avec deux mouvements forment une famille à une dimension c_{12} Elle peut être représentée par une courbe algébrique de degré sept de $\mathcal{P}^3$ en δ_1, δ_2, δ_3, δ_{12}, obtenue en écrivant les équations de Kruppa dans un repère particulier Cette courbe est de genre quatre Il est possible de tirer parti de ses propriétés pour en trouver une représentation sous forme d'une courbe algébrique de degré six de $\mathcal{P}^2$, plan dual au plan image, dont la construction est à présent décrite:

Soit $\mathbf{p}_1$, $\mathbf{p}_1'$ les deux épipoles obtenus à partir du premier mouvement La transformation épipolaire est caractérisée par la conique de Steiner s_1 passant par $\mathbf{p}_1$ et $\mathbf{p}_1'$, définie comme suit: deux droites épipolaires l et l' se correspondent si, et seulement si $l \cap l'$ appartient à s_1 Les deux tangentes à ω passant par $\mathbf{p}_1$ intersectent s_1 en $\mathbf{x}_1$ et $\mathbf{x}_2$ comme le montre la figure 8.1 La droite $l = \mathbf{x}_1\mathbf{x}_2$ correspond au point $\mathbf{x}_1 \times \mathbf{x}_2$ dans l'espace dual du plan image Le point $\mathbf{x}_1 \times \mathbf{x}_2$ se trouve sur une courbe g_{12} qui est obtenue à partir de c_{12} par une transformation algébrique g_{12} est de degré six et de genre quatre Le point $\mathbf{p}_1 \times \mathbf{p}_1'$ de g correspondant à la droite $\mathbf{p}_1\mathbf{p}_1'$ dans le plan image en est un point triple La courbe g_{12} a trois autres points doubles, qui sont également calculables par des méthodes analytiques Si l'on dispose d'un troisième mouvement, on construit de manière analogue une courbe g_{13} L'une des intersections réelles de g_{12} et de g_{13} correspond à la calibration correcte Elle peut être trouvée par une méthode itérative prenant en compte l'existence du point triple commun et des points doubles

8.2 Une approche semi-analytique

Dans cette première approche, nous exploitons la structure polynômiale de faible degré des équations de Kruppa en vue d'obtenir une méthode exacte qui ne nécessite en particulier pas d'initialisation La phase finale de la résolution nécessite néanmoins l'emploi de techniques numériques spécifiques, les continuations, que nous commençons par présenter Nous donnons ensuite des détails sur la manière dont nous utilisons la méthode des continuations et sur les résultats obtenus

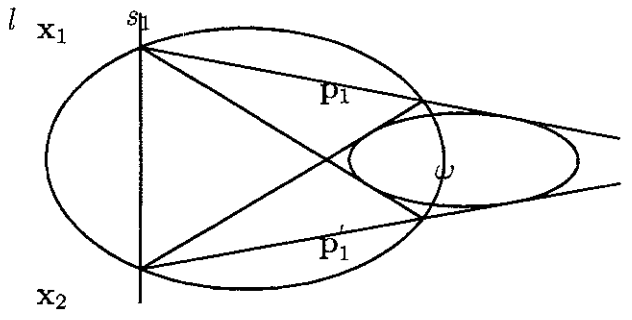


Fig. 8.1: Construction de la courbe duale

8.2.1 La méthode des continuations

L'utilisation de trois déplacements nous permet d'obtenir un système de six équations dont les six inconnues sont les six coefficients de Kruppa. Ces équations sont homogènes, les coefficients de Kruppa ne sont définis qu'à un facteur d'échelle près, aussi il nous reste après normalisation cinq inconnues. Nous avons donc affaire à un problème sur-déterminé, qu'il n'est pas possible de résoudre à l'aide des seules méthodes numériques usuelles. Il faut donc sélectionner cinq équations parmi les six, et essayer de résoudre le système de cinq équations à cinq inconnues ainsi obtenu. Nous avons vu que les équations sont de degré deux en les coefficients de Kruppa, aussi le nombre de solutions dans le cas général sera de $2^5 = 32$. Une fois que toutes ces solutions sont obtenues, il est possible d'utiliser l'équation restante ainsi que les contraintes (7.35) pour éliminer les mauvaises solutions. La difficulté provient du fait qu'en tentant de résoudre un système d'équations polynomiales par une méthode numérique classique, c'est-à-dire en partant d'une estimée initiale puis en utilisant une technique itérative, ne conduit pas en général à toutes les solutions: beaucoup des points de départ vont diverger, alors qu'un certain nombre d'autres convergeront vers la même solution. Or il n'est pas du tout acceptable d'obtenir seulement une partie des solutions, car nous savons que parmi les 32 solutions, il s'en trouve une seule de bonne, et il se pourrait très bien que nous la manquions dans ces conditions

Des méthodes récentes de continuation numérique permettent de calculer de façon assez sûre toutes les solutions des systèmes d'équations polynomiales. Ces méthodes ont bénéficié de multiples perfectionnements ces dix dernières années et sont à présent utilisées de manière assez courante pour résoudre des problèmes complexes de cinématique. Nous ne rentrerons pas dans le détail de ces perfectionnements, et renvoyons pour plus de détails à [118] et à [167] pour une revue. Nous expliquons ici juste le principe de ces méthodes, qui ont constitué un outil important pour la résolution des problèmes abordés dans ce chapitre. L'idée de base de la méthode des continuations appliquée aux systèmes non-linéaires est que des petites perturbations des paramètres du système produisent en général des changements petits dans les solutions. Supposons que nous connaissons les solutions du problème A (système de départ) et que nous voulons les solutions du problème B (système cible). Nous suivons les solutions du problème pendant que nous faisons varier lentement les paramètres de ceux de A à ceux de B. Cette idée est schématisée figure 8.2. Pour un système non-linéaire général, de nombreuses difficultés peuvent surgir, telles que la divergence ou la bifurcation de chemins de solutions. Cependant pour un système polynomial ces difficultés peuvent être évitées.

Le système de départ Il y a trois critères qui guident le choix d'un système de départ: toutes ses solutions doivent être connues, chaque solution doit être non-singulière, et le système doit avoir la même structure homogène que le système cible. L'utilisation de systèmes m-homogènes réduit la complexité des calculs en éliminant des solutions à l'infini, aussi l'homogénéisation est une opération utile. Cependant, pour des raisons de simplicité, nous ne présenterons ici que les systèmes inhomogènes. Aussi, un système de départ acceptable

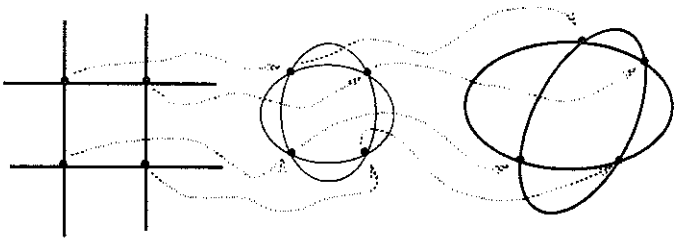


Fig. 8.2: Le principe de la méthode des continuations

est:

$$x_j^{d_j} - 1 = 0, \qquad 1 \leq j \leq n$$

où n est le nombre d'équations, d_j le degré de la $j$ème équation du système cible. Chaque équation donne d_j solutions distinctes pour x_j, et l'ensemble des N solutions, où $N = \prod_{j=1}^{n} d_j$ est obtenu en prenant toutes le combinaisons possibles.

L'Homotopie La contrainte pour le choix de l'homotopie (l'application utilisée pour transformer le système de départ en le système cible) est qu'au cours de la transformation il doive rester un nombre constant de solutions qui décrivent des chemins possédant de bonnes propriétés de régularité et aucune singularité jusqu'à ce que le système cible soit atteint. Des années de pratique ont montré que l'homotopie suivante est suffisante:

$$H(\mathbf{x}, t) = (1 - t)e^{i\theta}G(\mathbf{x}) + tF(\mathbf{x})$$

où $G(\mathbf{x})$ est le système de départ, et $F(\mathbf{x})$ le système cible.

Le suivi des chemins Le suivi des chemins consiste à suivre les solutions de $H(\mathbf{x}, t) = 0$ lorsque t varie de 0 à 1. Ces solutions forment d *chemins de continuations*, d étant le nombre de Bezout du système, qui caractérise le nombre de solutions. Afin de suivre un chemin à partir de la solution connue $(\mathbf{x}^0, t^0)$, nous prédisons d'abord la solution pour $t = t^0 + \Delta t$ en utilisant un développement de Taylor au premier ordre pour H. Il vient $\Delta\mathbf{x} = -J_x^{-1} J_t \Delta t$, expression dans laquelle J_x et J_t sont les jacobiennes de H par rapport à $\mathbf{x}$ et t. Nous corrigeons ensuite la prédiction en fixant t à sa nouvelle valeur et en utilisant la méthode de Newton pour calculer des termes correcteurs $\Delta\mathbf{x} = -J_x^{-1} H(\mathbf{x}, t)$.

8.2.2 Implémentation et résultats

Dans notre utilisation de la méthode des continuations, nous avons cherché à exploiter au maximum la redondance de l'information. Nous avons en effet remarqué que la méthode des continuations n'était pas entièrement fiable: quelquefois plusieurs des 32 solutions ne sont pas trouvées, ou encore les résultats dépendent de l'ordre des équations. C'est pourquoi, lorsque nous disposons de trois mouvements, nous résolvons les six systèmes d'équations qu'il est

possible de former en retranchant successivement l'une des six équations. Ceci nous permet en outre d'obtenir une estimation grossière de la variance des paramètres, qui est néanmoins un indicateur utile.

Nous avons également vérifié l'apport de la contrainte de perpendicularité (7.37) dans deux cas. Le premier est celui où nous ne disposons que de deux mouvements. La résolution est alors possible, puisque nous pouvons générer quatre équations de Kruppa indépendantes, et y joindre la contrainte (7.37), qui constitue également une équation du second degré en les coefficients de Kruppa. Nous obtenons ainsi un système du même type que ceux rencontrés au cours de la méthode générale. Le second cas est celui où nous disposons de trois mouvements. Nous exploitons alors de nouveau en partie la redondance des équations. Nous nous restreignons néanmoins à un sous-ensemble de six systèmes, parmi les $C_6^4 = 15$ systèmes qu'il est possible de générer en prenant pour l'une des équations la contrainte (7.37), et en choisissant les quatre autres équations parmi les six équations de Kruppa disponibles.

calcul des paramètres intrinsèques par les continuations avec trois mouvements

- générer six équations de Kruppa indépendantes à partir de trois matrices fondamentales.
- pour chacun des six systèmes de cinq équations $\mathcal{E}_i$:
  - résoudre $\mathcal{E}_i$ par la méthode des continuations pour obtenir les coefficients de Kruppa
  - éliminer les solutions complexes et les solutions réelles ne vérifiant pas les contraintes (7.35), pour obtenir une liste $\mathcal{S}_i$ de solutions
- pour chaque paire $\mathcal{S}_i$, $\mathcal{S}_j$ de listes de solutions, chercher les deux solutions de $\mathcal{S}_i$ et $\mathcal{S}_j$ les plus proches, au sens de la distance pondérée:

$$d(\mathbf{u}, \mathbf{v}) = \sum_i \frac{\|u_i - v_i\|}{\max(\|u_i\|, \|v_i\|)}$$

- sélectionner dans chaque liste $\mathcal{S}_i$ la solution qui a été trouvée le plus souvent à l'étape précédente
- calculer les paramètres intrinsèques par les formules (7.36).
- effectuer une moyenne des six solutions précédentes

Nous donnons un exemple des résultats obtenus avec chacune des trois méthodes:

- 2 mouvements avec contrainte de perpendicularité

- 3 mouvements avec contrainte de perpendicularité
- 3 mouvements sans contrainte de perpendicularité

Les appariements proviennent de la matrice de projection et des points 3D déjà utilisée ainsi que des trois mouvements donnés ci-dessous. Les matrices fondamentales sont calculées à partir de appariements. La première configuration, qui peut être considérée comme favorable, comprend ces trois mouvements. La seconde configuration comprend deux les mouvements déjà introduits pour illustrer les calculs de matrice fondamentale et le mouvement $\mathbf{r}_2, \mathbf{t}_2$

$$\begin{array}{ll} \mathbf{r}_1 = [0.181796, 0.093026, -0.153898] & \mathbf{t}_1 = [225.839222, 20.333256, 258.742916] \\ \mathbf{r}_2 = [-0.046898, 0.267737, 0.408056] & \mathbf{t}_2 = [51.364504, 112.356859, 241.607041] \\ \mathbf{r}_3 = [-0.029425, 0.043805, -0.390964] & \mathbf{t}_3 = [-265.839530, -210.610522, -295.322816] \end{array}$$

Les résultats sont présentés table 8.1 et table 8.2. Lors des essais effectués avec seulement deux mouvements, nous obtenons parfois plusieurs solutions pour les paramètres intrinsèques. Nous faisons alors figurer la solution donnant l'angle θ le plus proche de $\frac{\pi}{2}$. Pour les essais avec trois mouvements, nous avons aussi fait figurer entre crochets l'estimation de l'écart-type obtenu à partir des différentes solutions. Sur la première série de mouvements, la contrainte de perpendicularité contribue à améliorer les résultats. Ce fait a été constaté sur plusieurs autres exemples, mais n'est pas confirmé par l'étude statistique menée plus bas, laquelle concerne également les résultats obtenus par les méthodes itératives. Il est à noter que les résultats obtenus sur la seconde série sont moins significatifs, du fait de l'existence de nombreux systèmes sans solution dans les calcul, ce que l'estimation de la variance permet de détecter.

Il est intéressant d'examiner la raison des moins bons résultats obtenus dans la seconde série. Des deux nouveaux mouvements introduits, le second conduit à un positionnement nettement moins précis de l'épipole, comme nous l'avons vu abondamment au chapitre précédent. Cependant l'examen détaillé des résultats, par exemple des solutions à deux mouvements, montre que c'est le premier mouvement introduit qui est à la source de l'instabilité. Ceci s'explique assez bien si l'on remarque qu'il est assez proche d'une translation pure, et que celles-ci sont des mouvements dégénérés pour les équations de Kruppa.

L'avantage *considérable* de cette méthode est de ne nécessiter aucune initialisation. Si les points sont connus avec une grande précision, les résultats sont corrects. Un second point positif est qu'il est simple de se rendre compte que la méthode a échoué, soit que les solutions obtenues ne permettent pas de calculer des paramètres réels, soit que la recherche des solutions commune ne donne pas de résultats cohérents. Cependant la méthode souffre de plusieurs inconvénients:

- Le nombre de mouvements utilisés est le nombre minimal, et il n'est pas simple d'exploiter l'information provenant d'une plus longue séquence de mouvements sans multiplier les calculs.
- La méthode des continuations souffre parfois de problèmes de précision et nécessite un temps de calcul qui, s'il n'est pas prohibitif, reste relativement important: il faut une

bruit (pixels)	méthode orth.	méthode mouv.	paramètres obtenus α_u	α_v	u_0	v_0	$\theta - \frac{\pi}{2}$
0.1	oui	1,2	642.32	947.37	245.82	253.94	10^{-13}
	oui	2,3	639.44	944.36	246.04	258.59	10^{-13}
	oui	1,3	641.62	945.73	248.97	255.56	10^{-12}
	oui	1,2,3	641.69 [2.0]	947.49 [3.7]	247.03 [1.2]	256.55 [1.7]	10^{-13} [10^{-12}]
	non	1,2,3	644.40 [2.3]	952.29 [3.8]	237.45 [4.2]	254.61 [1.9]	6.10^{-3} [10^{-3}]
0.5	oui	1,2	651.39	962.47	244.73	246.56	10^{-12}
	oui	2,3	636.54	946.72	245.82	270.46	10^{-12}
	oui	1,3	647.41	953.67	260.91	255.11	10^{-13}
	oui	1,2,3	648.39 [11.1]	963.69 [20.3]	250.84 [6.7]	260.26 [9.1]	10^{-13} [10^{-12}]
	non	1,2,3	664.19 [11.2]	996.03 [20.6]	190.91 [23.8]	248.71 [9.1]	4.10^{-2} [2.10^{-2}]
1.0	oui	1,2	-	-	-	-	-
	oui	2,3	632.49	948.90	245.50	285.45	10^{-13}
	oui	1,3	74.85	455.95	733.93	434.07	10^{-10}
	oui	1,2,3	658.00 [24.8]	986.63 [45.7]	255.61 [14.3]	265.09 [19.7]	10^{13} [10^{-12}]
	non	1,2,3	681.66 [25.7]	1109.05 [75.6]	31.10 [139.9]	231.99 [20.5]	0.13 [0.08]
1.5	oui	1,2	676.05	1002.37	241.89	223.28	10^{-12}
	oui	2,3	627.92	950.11	245.16	300.56	10^{-12}
	oui	1,3	659.79	971.19	293.82	252.73	10^{-13}
	oui	1,2,3	669.62 [42.6]	1013.85 [79.2]	260.16 [23.3]	270.23 [32.3]	10^{-12} [10^{-12}]
	non	1,2,3	633.02 [73.0]	1223.62 [104.5]	190.46 [231.1]	205.49 [43.9]	0.27 [0.2]

Tableau 8.1: Résultats de la méthode des continuations, configuration 1

demi-minute pour résoudre un système de Kruppa (cinq équations du second degré) en utilisant un réseau d'une dizaine de Sun-4.

- Il n'est pas possible d'exprimer les contraintes (7.35) *au niveau de la résolution* car la méthode des continuations fonctionne dans le plan complexe. La conséquence de ce défaut est qu'en présence de matrices fondamentales fortement bruitées, nous pouvons obtenir des listes de solutions dont aucune ne vérifie ces contraintes, ce qui ne permet pas l'estimation des paramètres intrinsèques.

- Nous ne pouvons pas tenir compte de l'incertitude simplement et rigoureusement. Au niveau des matrices fondamentales, la connaissance de cette incertitude est une information importante, comme l'a montré le chapitre précédent, puisque différents mouvements peuvent donner lieu à des calculs dont la stabilité est très variable. Il serait également souhaitable de caractériser l'incertitude des paramètres obtenus.

8.3 Une approche itérative

Les approches itératives présentent plusieurs attraits pratiques, que nous commençons par expliquer. Nous avons ensuite effectué de nombreuses expérimentations avec deux méthodes:

bruit (pixels)	méthode orth.	méthode mouv.	paramètres obtenus α_u	α_v	u_0	v_0	$\theta - \frac{\pi}{2}$
0.1	oui	1,2	647.56	955.50	245.32	250.58	10^{-13}
	oui	2,3	124.317	947.934	230.705	252.053	10^{-13}
	oui	1,3	639.303	943.591	246.083	257.594	10^{-12}
	oui	1,2,3	640.83 [2.7]	947.59 [2.7]	237.90 [10.2]	252.88 [3.1]	10^{-13} $[10^{-12}]$
	non	1,2,3	636.32 [15.8]	942.45 [5.0]	241.87 [5.9]	251.61 [2.8]	0.018 [0.02]
0.5	oui	1,2	-	-	-	-	-
	oui	2,3	-	-	-	-	-
	oui	1,3	635.76	942.88	246.03	265.70	10^{-12}
	oui	1,2,3	654.01 [24.1]	976.83 [22.4]	214.28 [47.0]	232.66 [20.7]	10^{-13} $[10^{-12}]$
	non	1,2,3	623.63 [78.8]	934.15 [31.4]	240.84 [2.1]	237.95 [14.7]	0.089 [0.09]
1.0	oui	1,2	744.34	1110.38	235.28	187.89	10^{-13}
	oui	2,3	-	-	-	-	-
	oui	1,3	630.86	941.71	245.96	277.04	10^{-13}
	oui	1,2,3	505.94 [248.7]	779.03 [389.9]	179.30 [94.4]	407.68 [317.3]	10^{13} $[10^{-12}]$
	non	1,2,3	628.20 [130.4]	936.94 [68.6]	208.05	217.74 [27.9]	0.15 [0.1]
1.5	oui	1,2	2462.05	3943.05	27.53	-558.13	10^{-14}
	oui	2,3	342.86	875.35	219.38	246.91	10^{-14}
	oui	1,3	604.46	885.15	249.27	260.23	10^{-12}
	oui	1,2,3	688.43 [163.9]	1048.77 [254.3]	161.48 [75.0]	207.08 [38.7]	10^{-12} $[10^{-12}]$
	non	1,2,3	1190.91 [1164.0]	1803.80 [1867.3]	109.39 [149.1]	-109.65 [661.5]	0.13 [0.1]

Tableau 8.2: Résultats de la méthode des continuations, configuration 2

une minisation non-linéaire globale, et le filtrage de Kalman étendu.

8.3.1 Avantages pratiques

Dans cette approche, nous renonçons à tirer parti de la structure polynômiale des équations de Kruppa, pour exploiter celles-ci comme de simples équations de mesure dans lesquelles les paramètres mesurés sont les coefficients des matrices fondamentales et les variables à estimer sont les paramètres intrinsèques. L'idée originale est simplement de substituer les valeurs des coefficients de Kruppa données par les équations (7.29-7.34) dans l'expression (7.16) des équations de Kruppa, et de trouver à l'aide de plusieurs mouvements une solution au sens des moindres carrés. Cette formulation a l'avantage de permettre de prendre en compte de manière naturelle un grand nombre de mouvements, ainsi que de tenir compte de la contrainte de complexité, puisque, nous l'avons vu précédemment, le paramétrage par les cinq paramètres intrinsèques rend compte de toutes les coniques purement imaginaires. Il est également très simple d'exprimer la contrainte d'orthogonalité, puisqu'il suffit de faire $\theta = \frac{\pi}{2}$ dans les équations résultantes, et de faire porter les calculs seulement sur les quatre premiers paramètres intrinsèques. De la même manière, il est aussi facile de tenir compte d'éventuelles connaissances *a priori* sur la caméra (par exemple celle du rapport α_u/α_v qui est assez réaliste), en fixant d'autres paramètres. Le principal inconvénient de cette méthode

est de nécessiter un point d'initialisation. Celui-ci peut néanmoins être fourni par la méthode des continuations, ou, dans la plupart des applications, par une estimation correspondant à des valeurs standard. Nous avons implémenté deux méthodes de minimisation, l'une fondée sur l'utilisation d'un simple moindres carrés non-linéaire, l'autre sur un filtrage de Kalman.

8.3.2 Minimisation globale

Dans cette implémentation, afin d'exploiter le nombre maximal de contraintes, nous utilisons les trois équations de Kruppa provenant de chaque mouvement, bien qu'elles ne soient pas indépendantes. Il est à noter que nous ne minimisons pas les valeurs résiduelles des équations de Kruppa telles qu'elles sont exprimées en (7.16) mais plutôt le critère:

$$\sum_i \{(\frac{k_{11}}{k''_{11}} - \frac{k_{12}}{k''_{12}})^2 + (\frac{k_{12}}{k''_{22}} - \frac{k_{12}}{k''_{22}})^2 + (\frac{k_{11}}{k''_{22}} - \frac{k_{11}}{k''_{22}})^2\} \tag{8.1}$$

dans lequel les coefficients k_{ij} et k''_{ij} sont définis en (7.11) et (7.15). Ce critère semble moins simple que celui obtenu directement à partir des équations de Kruppa, puisque celles-ci conduisent à un critère qui est polynomial, de degré quatre en les paramètres α_u, α_v, u_0, v_0 dans le cas du modèle à quatre paramètres ($\theta = \frac{\pi}{2}$), néanmoins nos expérimentations montrent qu'il donne des résultats nettement plus stables, pour une raison que nous avions déjà évoquée au chapitre 3 pour expliquer la supériorité de la méthode des birapports sur les méthodes polynomiales. Une seconde remarque est que dans le cas du modèle à quatre paramètres, les équations obtenues ne dépendent de α_u et de α_v que par l'intermédiaire de leurs carrés, et donc ne déterminent pas du tout le signe de ces valeurs. Compte tenu des conventions que nous adoptons pour les paramètres intrinsèques, nous imposons un signe positif au résultat final.

Nous avons d'abord comparé cette méthode avec celle des continuations, en testant également l'influence de la contrainte de perpendicularité, au moyen d'une approche statistique utilisant 100 triplets de mouvements. A partir de chacun des triplets, les appariements sont générés, puis trois matrices fondamentales sont calculées par la méthode non-linéaire **DIST-L** présentée dans le chapitre 4. Le programme de minimisation a ensuite été lancé à partir des points de départ constitués par: (1) les valeurs exactes, (2) les valeurs obtenues à partir des continuations, (3) des valeurs standard $\alpha_u = 800$, $\alpha_v = 800$, $u_0 = 255$, $v_0 = 255$, $\theta = \frac{\pi}{2}$, correspondant à une grille orthogonale, un point principal situé exactement au centre de l'image, et des ordres de grandeur raisonnables (mais ne supposant pas connu le rapport α_u/α_v) pour les facteurs d'échelle. Les résultats statistiques, obtenus sous forme de distances relatives, sont détaillés pour quatre niveaux de bruit différents table 8.3.

Les conclusions qu'il est possible de tirer de ces résultats sont:

- la précision sur les paramètres α_u et α_v est meilleure que celle sur les paramètres u_0, v_0, quelle que soit la méthode employée.
- la contrainte de perpendicularité conduit à des résultats moins robustes.

bruit (pixels)	méthode	echecs	modèle général					avec perpendicularité			
			α_u	α_v	u_0	v_0	$\theta - \frac{\pi}{2}$	α_u	α_v	u_0	v_0
0.2	Continu	2	0.056	0.062	0.133	0.136	0.035	0.095	0.105	0.231	0.204
	Mini (1)		0.025	0.028	0.071	0.095	0.034	0.095	0.105	0.231	0.204
	Mini (2)		0.054	0.056	0.106	0.123	0.051	0.100	0.084	0.170	0.166
	Mini (3)		0.065	0.063	0.136	0.153	0.055	0.081	0.076	0.157	0.188
0.6	Continu	10	0.065	0.092	0.158	0.200	0.058	0.149	0.158	0.318	0.345
	Mini (1)		0.052	0.058	0.131	0.192	0.060	0.096	0.074	0.284	0.304
	Mini (2)		0.051	0.063	0.158	0.218	0.062	0.221	0.159	0.379	0.443
	Mini (3)		0.069	0.063	0.165	0.218	0.066	0.156	0.148	0.234	0.274
1.0	Continu	7	0.120	0.141	0.263	0.321	0.080	0.198	0.215	0.443	0.454
	Mini (1)		0.115	0.141	0.261	0.327	0.098	0.128	0.116	0.275	0.328
	Mini (2)		0.138	0.160	0.274	0.370	0.130	0.202	0.216	0.342	0.380
	Mini (3)		0.174	0.184	0.281	0.360	0.135	0.182	0.175	0.306	0.362

Tableau 8.3: Résultats statistiques de la méthode de minimisation globale avec 3 mouvements

- les résultats obtenus avec la méthode des continuations et ceux obtenus avec la méthode de minimisation globale sont à peu près comparables, cependant lorsque les points sont bruités la méthode des continuations échoue quelquefois totalement à trouver des paramètres réels.
- la méthode de minimisation globale est assez sensible au choix du point d'initialisation

Utilisation de connaissances sur les paramètres Par ailleurs, nous avons tenté de fixer différents paramètres dans le processus de minimisation en introduisant des connaissances a priori "raisonnables", ce qui permet de ne calculer qu'un nombre réduit de paramètres:

- 1 paramètre: la distance focale effective (α_u), le rapport α_u/α_v exact est fourni, le point principal est placé au centre de l'image.
- 2 paramètres: les facteurs d'échelle (α_u, α_v), le point principal est placé au centre de l'image.
- 3 paramètres (α_u, u_0, v_0), le rapport α_u/α_v exact est fourni.

Nous avons choisi de voir l'effet de l'introduction de ces connaissances en nous fondant, pour des raisons de commodité, sur la procédure itérative avec contrainte de perpendicularité et trois mouvements. L'initialisation est effectuée avec les valeurs standard. Les paramètres intrinsèques varient d'un essai à l'autre, α_u est pris entre 400 et 1200, α_v est pris comme $1.4742\alpha_u$ avec des fluctuations de 1%, et u_0, v_0 sont placés autour du centre de l'image

avec une variance de 20 pixels. Les résultats obtenus avec 100 essais sont présentés table 8.4. Chaque case contient la moyenne des distances relatives, le pourcentage d'essais pour lesquels l'erreur a été inférieure à 5%, ainsi que le rapport du nombre de cas pour lesquels le calcul avec paramètres fixés s'est avéré significativement meilleur, sur celui pour lesquels le calcul avec 4 paramètres s'est avéré significativement meilleur.

bruit	param. fixés	α_u	α_v	u_0	v_0
0.2 pixel	1	0.1009 59 18/16	0.1004 59 24/18	0.0667 47 42/21	0.0636 49 46/15
	2	0.1715 50 15/31	0.2019 49 14/28	0.0667 47 42/21	0.0636 49 46/15
	3	0.1153 70 16/10	0.1157 71 20/10	0.1972 44 25/18	0.2406 39 25/14
	4	0.1495 68	0.1704 66	0.2418 45	0.2704 41
1.0 pixel	1	0.2468 38 34/22	0.2465 37 37/26	0.0667 47 63/7	0.0636 49 71/5
	2	0.3162 25 24/33	0.3453 29 28/33	0.0667 47 63/7	0.0636 49 71/5
	3	0.2767 37 27/21	0.2767 37 32/22	0.3994 20 26/27	0.5145 15 30/27
	4	0.3098 30	0.3245 31	0.4439 22	0.5286 20

Tableau 8.4: L'apport de connaissances a priori dans la détermination des paramètres intrinsèques

Ces résultats montrent que l'introduction du rapport α_u/α_v améliore toujours un peu les résultats. Le fait de fixer le point principal au centre de l'image dégrade la précision sur les facteurs d'échelle, lorsque leur rapport n'est pas imposé, mais donne un point principal toujours plus précis que celui obtenu sans connaissance a priori. Les résultats les meilleurs sont donc obtenus en fixant le maximum de paramètres.

8.3.3 Filtrage récursif

Le filtre de Kalman étendu itéré Dans le cas où nous disposons d'une longue séquence de mouvements, il est intéressant d'utiliser un filtre de Kalman étendu[1]. Cette méthode récursive d'estimation optimale de paramètres $\mathbf{a}$ à partir de mesures bruitées $\mathbf{x}_i$ a été employée pour résoudre de nombreux problèmes de vision. Elle présente l'avantage d'être incrémentale, et de prendre en compte dans sa formulation la covariance des mesures et des paramètres estimés. Dans notre cas, les données sont:

[1] Les références classiques sont les livres [71] [108]. Une présentation pratique se trouve dans des chapitres de [6], [39], et [176].

vecteur de paramètres d'état	$\mathbf{a} = (\alpha_u, \alpha_v, u_0, v_0)^T$
vecteur de mesures	$\mathbf{x} = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^T$
équation de mesure	$\mathbf{f}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$, f_1 et f_2 sont deux[2]équations de Kruppa (7.16)

Le filtre de Kalman s'applique lorsque la relation entre les mesures et les paramètres d'état est linéaire. Il faut donc effectuer une approximation linéaire de $\mathbf{f}$ au voisinage de l'estimation de $\mathbf{a}$ et de $\mathbf{x}_i$, ce qui suppose bien sûr que nous ayons une assez bonne estimée de ces deux vecteurs. En ce qui concerne le vecteur de mesures, cela signifie que les matrices fondamentales sont calculées avec une précision raisonnable, alors qu'en ce qui concerne le vecteur d'état, cela signifie la connaissance d'une bonne estimée initiale des paramètres intrinsèques. Dans ce cas, au voisinage d'une estimation courante $\hat{\mathbf{a}}_{i-1}$ et de la mesure $\hat{\mathbf{x}}_i$, après un développement de Taylor au premier ordre, l'équation de mesure se réécrit:

$$\mathbf{y}_i = \mathbf{M}_i \mathbf{a} + \mathbf{w}_i \tag{8.2}$$

avec

$$\begin{aligned}
\mathbf{y}_i &= -\mathbf{f}(\hat{\mathbf{x}}_i, \hat{\mathbf{a}}_{i-1}) + \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\hat{\mathbf{x}}_i, \hat{\mathbf{a}}_{i-1})\,\hat{\mathbf{a}}_{i-1} \\
\mathbf{M}_i &= \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\hat{\mathbf{x}}_i, \hat{\mathbf{a}}_{i-1}) \\
\mathbf{\Lambda}_{\mathbf{w}_i} &= \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\hat{\mathbf{x}}_i, \hat{\mathbf{a}}_{i-1}) \mathbf{\Lambda}_{\mathbf{x}_i} \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\hat{\mathbf{x}}_i, \hat{\mathbf{a}}_{i-1})^T
\end{aligned}$$

Le critère minimisé est:

$$\min_{\mathbf{a}} \{(\hat{\mathbf{a}}_0 - \mathbf{a})^t \, \mathbf{\Lambda}_{\mathbf{a}_0}^{-1} \, (\hat{\mathbf{a}}_0 - \mathbf{a}) + \sum_{i=1}^{N} (\mathbf{y}_i - \mathbf{M}_i \mathbf{a})^T \, \mathbf{\Lambda}_{\mathbf{w}_i}^{-1} \, (\mathbf{y}_i - \mathbf{M}_i \mathbf{a})\} \tag{8.3}$$

Son premier terme représente l'écart entre le vecteur d'état $\mathbf{a}$ et son estimée initiale $\hat{\mathbf{a}}_0$ pondérée par la matrice de covariances $\Lambda_{\mathbf{a}_0}$. Il faut noter que le choix de celle-ci a une très grande influence sur le déroulement des calculs. Le second terme représente la somme du carré des erreurs de mesure, également pondérées par leurs covariances. Une solution récursive est fournie par les équations du filtre de Kalman, qui calculent après chaque mesure la nouvelle estimée $(\hat{\mathbf{a}}_i, \Lambda_{\mathbf{a}_i})$ à partir de l'estimée courante $(\hat{\mathbf{a}}_{i-1}, \Lambda_{\mathbf{a}_{i-1}})$

$$\begin{aligned}
\hat{\mathbf{a}}_i &= \hat{\mathbf{a}}_{i-1} + K_i(\mathbf{y}_i - \mathbf{M}_i \hat{\mathbf{a}}_{i-1}) \\
K_i &= \mathbf{\Lambda}_{\mathbf{a}_{i-1}} \mathbf{M}_i^T (\mathbf{\Lambda}_{\mathbf{w}_i} + \mathbf{M}_i \mathbf{\Lambda}_{\mathbf{a}_{i-1}} \mathbf{M}_i^T)^{-1} \\
\mathbf{\Lambda}_{\mathbf{a}_i} &= (I - K_i \mathbf{M}_i) \mathbf{\Lambda}_{\mathbf{a}_{i-1}}
\end{aligned}$$

[1]Nous n'utilisons pas *trois* équations de Kruppa car il est connu que les meilleurs résultats sont obtenus avec des équations de mesure indépendantes

Les difficultés avec cette technique proviennent du fait que les approximations au premier ordre permettant d'aboutir à la forme (8.2) ne sont valables que si les estimations $\hat{a}_i$ sont suffisamment proches de la solution exacte a. Un autre problème est que le filtre de Kalman étendu sous-estime sérieusement la covariance [103]. Pour éviter la convergence vers des minima locaux provenant des non-linéarités, nous itérerons le filtre sur l'ensemble de toutes les mesures plusieurs fois, en réinitialisant la matrice de covariance Λ_a à chaque itération, mais en conservant l'estimée $\hat{a}_N$.

Un exemple Nous commençons par donner un exemple figure 8.3, pour les paramètres intrinsèques précédents ont été utilisés. L'initialisation pour le point principal a été effectuée à (350,350), ce qui est une position assez lointaine. Le bruit est fixé à 0.5 pixel. Nous avons représenté les valeurs estimées par le filtre, ainsi que l'ecart-type estimé sur chacune des valeurs en prenant juste la racine carrée de sa covariance. Nous pouvons observer que la convergence se produit au bout de quelques mouvements.

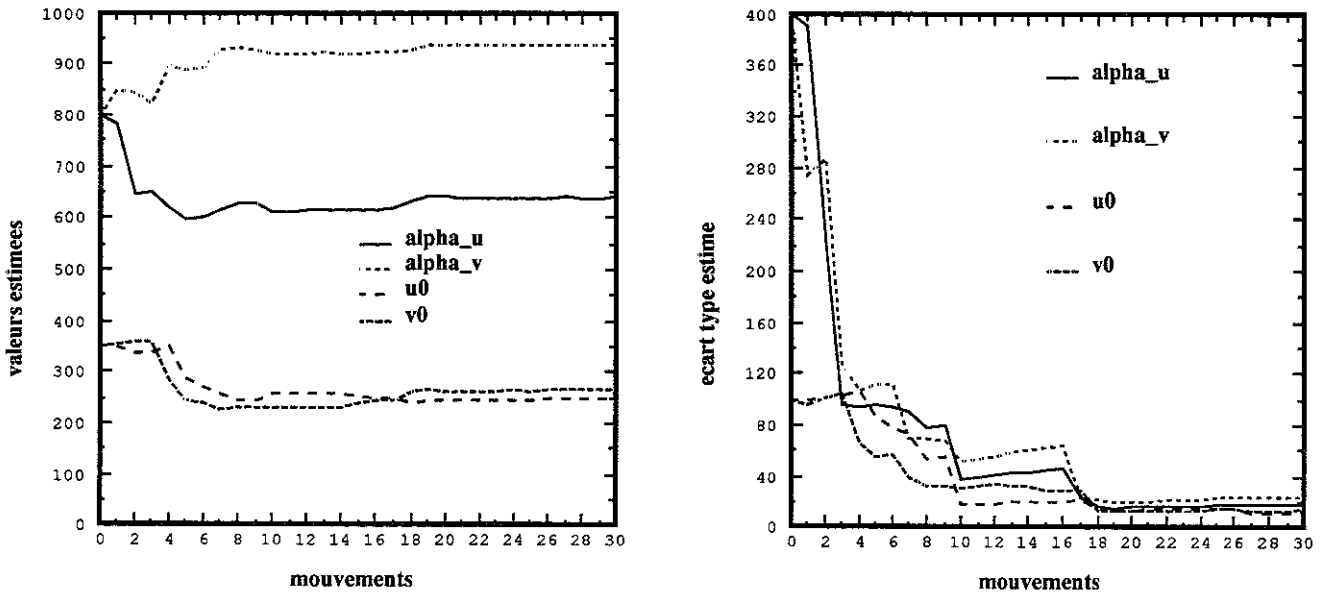


Fig. 8.3: Un exemple de calcul de paramètres intrinsèques par filtrage de Kalman.

Simulations statistiques Nous avons testé le filtre de Kalman au moyen d'une procédure statistique, en générant 100 séries de 15 déplacements, associés chaque fois à des paramètres intrinsèques différents, variés comme précédemment. Les résultats, parmi lesquels nous avons sélectionné, quelques données pour les présenter plus bas, sont relativement complexes à interpréter, car ils dépendent d'un nombre assez élevé de facteurs, qui sont:

- *Le niveau de bruit.* Nous avons effectué systématiquement nos essais avec deux niveaux de bruit gaussien, de variance 0.2 pixel et 1.0 pixel[2] Le comportement global est identique dans les deux cas. Pour quelques valeurs numériques permettant d'avoir une idée des précisions pouvant être atteintes, voir les tables 8.5 et 8.6.

- *Le nombre de mouvements utilisés.* La convergence est assez rapide, et a lieu en général au bout de 5 à 10 mouvements. Nous sommes allés jusqu'à 15 mouvements.

- *Le nombre d'itérations effectuées.* L'itération du filtre accélère la convergence et a à peu près le même effet que l'ajout de mouvements. En cas de convergence vers le bon minimum, des résultats plus précis sont obtenus, dans le cas contraire, les résultats s'éloignent davantage de la bonne valeur. Nous avons effextué jusqu'à cinq itérations. Au delà, les résultats n'évoluent plus.

- *L'initialisation.* Elle est très importante, et détermine en grande partie les résultats du filtre. Nous en avons considéré deux types d'initialisations, qui sont réalistes: l'initialisation standard déjà utilisée, et une initialisation par le résultat de la minimisation avec les trois premiers déplacements. Le choix de la matrice de covariance initiale a également une influence importante sur les résultats finaux. Compte tenu des valeurs courantes des paramètres intrinsèques, nous avons pris: $\sigma_{\alpha_u} = \sigma_{\alpha_u} = 400$, $\sigma_{u_0} = \sigma_{v_0} = 100$ et mis les termes non-diagonaux de la matrice de covariance à 0.

- *La nature des paramètres.* Les paramètres α_u et α_v ont un comportement assez identique, et différent de celui des paramètres u_0 et v_0, comme nous l'avons déjà constaté plus haut.

La mesure constituée par la moyenne des erreurs relatives ne s'est pas révélée adéquate pour caractériser la convergence du filtre dans tous les cas. Nous avons en effet constaté que les valeurs relatives à α_u et α_v demeurent stationnaires ou légèrement croissantes en fonction du nombre de mouvements effectués. Ceci signifie que *statistiquement* le filtre n'améliore pas ces valeurs, compte tenu des cas de convergence vers des faux minima. Cependant, si nous tenons compte de la nature de ces cas, nous observons bien que dans la plupart des essais, le filtre a amélioré les valeurs initiales. Nous prenons donc comme mesure de qualité d'estimation d'une valeur, le pourcentage d'essais pour lequel la différence relative entre l'estimée et la valeur exacte est supérieure à un certain seuil.

La figure 8.4 permet d'observer la convergence, ainsi que l'effet de l'initialisation. Sur ces figures, obtenues avec un bruit de 0.2 pixel, nous avons représenté le pourcentage d'essais pour lequel la différence relative entre l'estimée et la valeur exacte est supérieure à 5%. La figure 8.5, permet de voir l'effet de l'itération du filtre, l'initialisation standard ayant été utilisée.

[2]Ces valeurs correspondent à peu près à la précision atteinte par les détecteurs fondés sur des modèles et les détecteurs fondés sur des opérateurs

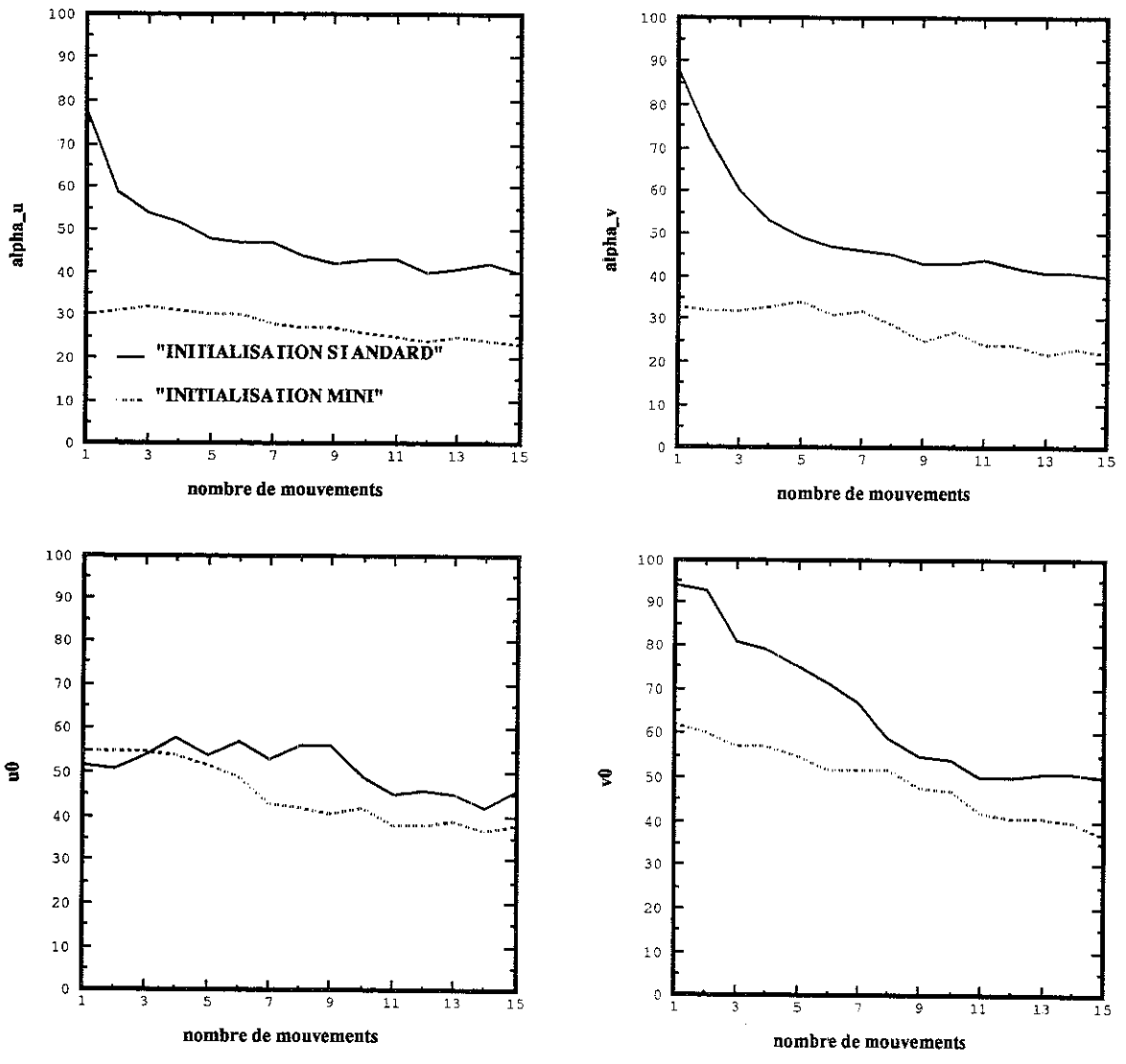


Fig. 8.4: Statistiques de convergence des paramètres intrinsèques avec deux types d'initialisation

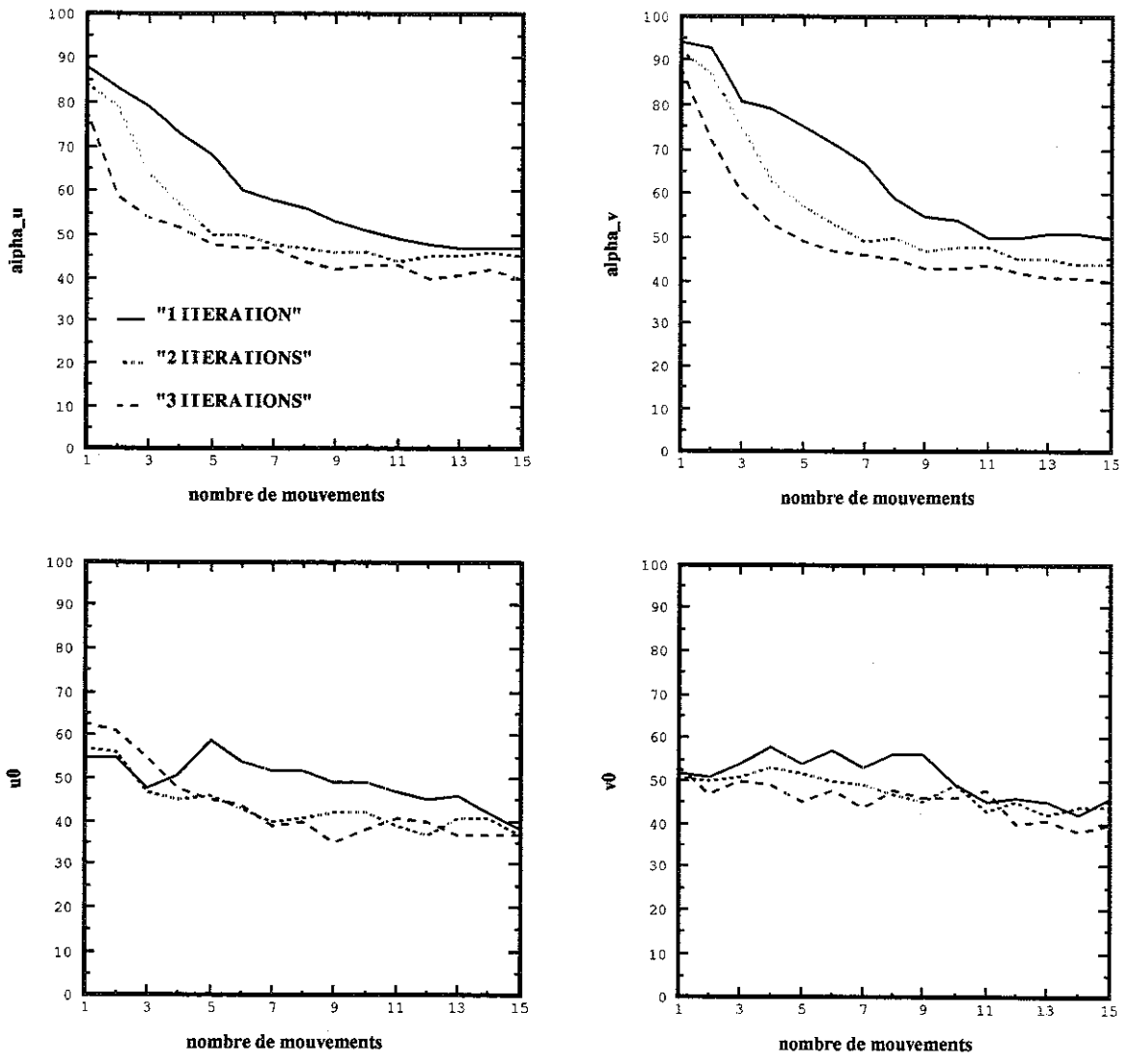


Fig. 8.5: Statistiques de convergence des paramètres intrinsèques avec itérations

Comparaison avec la méthode par minimisation Une idée de la précision numérique de la méthode est fournie par les tables 8.5 et 8.6. La première de ces tables compare les résultats obtenus par la méthode de minimisation globale décrite précédemment, avec ceux obtenus par le filtre de Kalman utilisé de manière réaliste, c'est-à-dire initialisé en utilisant la minimisation sur trois mouvements. Les premières lignes représentent la moyenne des erreurs relatives. Celle-ci diminue, pour la méthode de minimisation, diminue toujours avec le nombre de mouvements employés. En revanche, pour la méthode par filtrage, elle ne diminue que pour les paramètres u_0 et v_0. Cependant ceci ne signifie pas que le filtre n'a pas d'effet, comme permet de le constater les secondes lignes, qui montrent le pourcentage d'essais pour lequel la différence relative entre l'estimée et la valeur exacte est supérieure à 5% (resp 10%) pour 0.2 (resp 1.0) pixel de bruit. Nous trouvons comme [170] que les résultats du filtre de Kalman sont donc *statistiquement* moins bons que ceux obtenus par minimisation globale, à cause du nombre des cas où le premier ne trouve pas le bon minimum. Ce fait a aussi été souligné dans [82]. Nous pouvons remarquer qu'alors les faibles valeurs de covariance obtenues sont trompeuses. L'explication de ce moins bon comportement est que, vu la manière récursive dont nous l'utilisons, en introduisant successivement les données, le filtre de Kalman ne dispose pas de l'ensemble de l'information: si nous nous référons à l'équation (8.3), nous pouvons voir que les éléments intervenant dans la mesure i sont calculés avec l'estimation $\hat{\mathbf{a}}_{i-1}$ *obtenue avec seulement les* i *premières mesures*, et non avec l'estimation $\hat{\mathbf{a}}$ utilisant toutes les mesures. Pour les premières valeurs de i, l'estimation $\hat{\mathbf{a}}_{i-1}$ est assez imprécise, vu que peu de mesures sont utilisées. Les premiers termes du critère (8.3) peuvent donc être suffisamment incorrects pour fausser totalemment l'estimation finale de $\hat{\mathbf{a}}$. En pratique, si les valeurs partent dans de mauvaises directions au début, elles ont tendance à demeurer piégées dans des minima locaux.

La seconde table 8.6 montre l'effet de l'application du filtre de Kalman lorsque l'estimée initiale est suffisament précise. En l'occurence, nous avons comparé les résultats obtenus par la minimisation globale avec 15 mouvements, et ceux obtenus en effectuant un filtrage en partant de ces résultats. Nous constatons alors une précision accrue, qui s'explique par le fait que le filtre prend en compte la covariance des matrices fondamentales dans le calcul.

8.4 Résultats

Nous faisons le point sur les résultats précédents, qui ont été obtenus à l'aide de nombreuses simulations, puis discutons de l'applicabilité de la méthode, avant de présenter plusieurs résultats obtenus sur des images réelles.

8.4.1 Une évaluation des méthodes

Nous avons présenté plusieurs méthodes pratiques permettant d'exploiter les équations de Kruppa. Compte tenu du fait qu'elles possèdent des caractéristiques assez différentes, nous serions tentés penser qu'aucune d'elle ne s'affirme vraiment supérieure aux autres. Avec un

bruit (pixel)	nb mouv	α_u mini	α_u kalman	α_v mini	α_v kalman	u_0 mini	u_0 kalman	v_0 mini	v_0 kalman
0.2	3	0.1494	0.1389	0.1487	0.1398	0.3047	0.2778	0.3091	0.2840
		25	32	28	32	49	55	52	57
	5	0.0952	0.1377	0.0837	0.1462	0.2264	0.2676	0.2591	0.2713
		22	30	20	34	42	52	43	55
	10	0.0652	0.1390	0.0714	0.1333	0.1970	0.2343	0.2401	0.2407
		20	26	23	27	40	42	47	47
	15	0.0464	0.1392	0.0753	0.1201	0.2033	0.2243	0.2042	0.2167
		10	23	18	22	33	38	39	37
1.0	3	0.3886	0.3535	0.3944	0.3647	0.5237	0.5070	0.5783	0.5540
		62	53	68	57	76	71	79	79
	5	0.3335	0.3383	0.3124	0.3664	0.4742	0.4936	0.499	0.5221
		57	51	62	55	70	68	78	76
	10	0.2875	0.3534	0.2913	0.3707	0.4406	0.4644	0.5144	0.4888
		60	49	56	53	70	61	76	65
	15	0.3017	0.3662	0.2710	0.3712	0.4336	0.4348	0.4954	0.4774
		59	50	58	53	67	58	73	68

Tableau 8.5: Comparaison de la minimisation et du filtre de Kalman

bruit (pixel)		α_u mini	α_u kalman	α_v mini	α_v kalman	u_0 mini	u_0 kalman	v_0 mini	v_0 kalman
0.2	err	0.0412	0.0408	0.0737	0.0659	0.1832	0.1448	0.1916	0.1396
	% err > 0.05	22	13	28	15	48	29	58	40
	% meilleur	6	12	5	16	4	27	6	26
1.0	err	0.2337	0.2498	0.2245	0.2796	0.4165	0.3500	0.4854	0.4171
	% err > 0.05	68	49	64	58	78	59	78	72
	% meilleur	21	30	25	29	14	37	17	37

Tableau 8.6: Effet de l'application du filtre de Kalman

nombre minimal de mouvements, il semble cependant que les meilleurs résultats soient obtenus avec la méthode des continuations répétée. Le principal inconvénient est le temps de calcul, qui est de plusieurs minutes. L'utilisation de mouvements multiples se prête bien à l'emploi de méthodes itératives, qui se révèlent toutes rapides. L'utilisation d'une minimisation globale non-linéaire, utilisant toutes les données et d'éventuelles connaissances *a priori*, suivi d'un raffinement par filtrage de Kalman étendu itéré devrait conduire aux résultats les plus précis.

Les difficultés que nous avons rencontrées sont de plusieurs types:

- Les ordres de grandeur de la précision moyenne obtenue avec les diverses méthodes sont en fait assez comparables, mais leur comportement varie suivant les cas et ne s'est pas révélé prévisible: une méthode donnée peut se révéler la meilleure dans certains cas et échouer complètement dans d'autres cas.
- La qualité de détermination du point principal n'est pas satisfaisante: il est préférable de supposer que celui-ci se trouve au centre de l'image, plutôt que de chercher à le calculer par autocalibration. Au contraire, il est avantageux d'utiliser cette connaissance pour éliminer des solutions aberrantes.
- Les simulations semblent montrer que les matrices fondamentales doivent être déterminées à partir d'appariements dont l'erreur de localisation ne dépasse pas le pixel.

Il par conséquent possible de dire sans exagérer que:

> *l'autocalibration est une méthode relativement instable dont les résultats ne sont pas comparables à ceux obtenus par une bonne méthode de calibration classique*

Il s'agit donc d'un problème très difficile, lorsqu'il est abordé dans sa plus grande généralité, ce qui est l'une des conclusions essentielles de ce travail. Il est en fin de compte assez normal, qu'en partant d'informations beaucoup moins riches, nous parvenions en fin de compte à des résultats moins précis.

8.4.2 L'autocalibration est utilisable

Nous voudrions pourtant tempérer ces conclusions apparemment négatives: d'abord, les paramètres peuvent, dans les cas favorables, être obtenus avec une précision absolue tout à fait correcte, ce que plusieurs exemples présentés dans la sous-section suivante montreront. Ces cas favorables sont connus, grâce aux études menées aux chapitres 5 et surtout 6, et il est facile pour un système de vision *actif* d'en tirer parti pour s'autocalibrer. De plus, les simulations précédentes ont montré que les résultats sont plutôt instables *statistiquement*, à cause d'une certaine proportion de cas où l'autocalibration donne des résultats aberrants. Des résultats précis peuvent être néanmoins obtenus dans le nombre relativement important de cas où l'algorithme a convergé. Un second point important, qui fera l'objet du chapitre suivant, est que même avec l'estimation relativement inexacte des paramètres intrinsèques fournie

par l'autocalibration, il est possible d'achever la calibration en calculant des paramètres extrinsèques qui sont suffisamment précis pour aboutir à des informations tridimensionnelles métriques utiles. Nous verrons que lors de ces processus, il est possible d'améliorer encore le calcul des paramètres intrinsèques de manière significative. Nous n'avons pas encore dit notre dernier mot en matière d'autocalibration.

Avons-nous déjà atteint notre objectif initial, qui était de développer une méthode automatique qui ne nécessite aucun modèle des objets observés, aucune connaissance du mouvement de la caméra, et aucune connaissance a priori des paramètres intrinsèques ? Notre réponse serait que la méthode que nous avons présentée a ce potentiel, et ce, de façon pratique. Cependant, les problèmes d'obtention automatique de correspondances très précises n'ayant pas encore été résolus, nous devons introduire quelques contraintes pour limiter l'instabilité. Il s'agit du choix des mouvements effectués, et de l'élimination des solutions aberrantes en utilisant des faibles connaissances sur les paramètres intrinsèques. Ces contraintes demeurent très faibles, et sont sans comparaison avec celles requises par la calibration classique.

8.4.3 Quelques cas réels

Tous les exemples de cette partie ont été obtenus avec la même caméra,

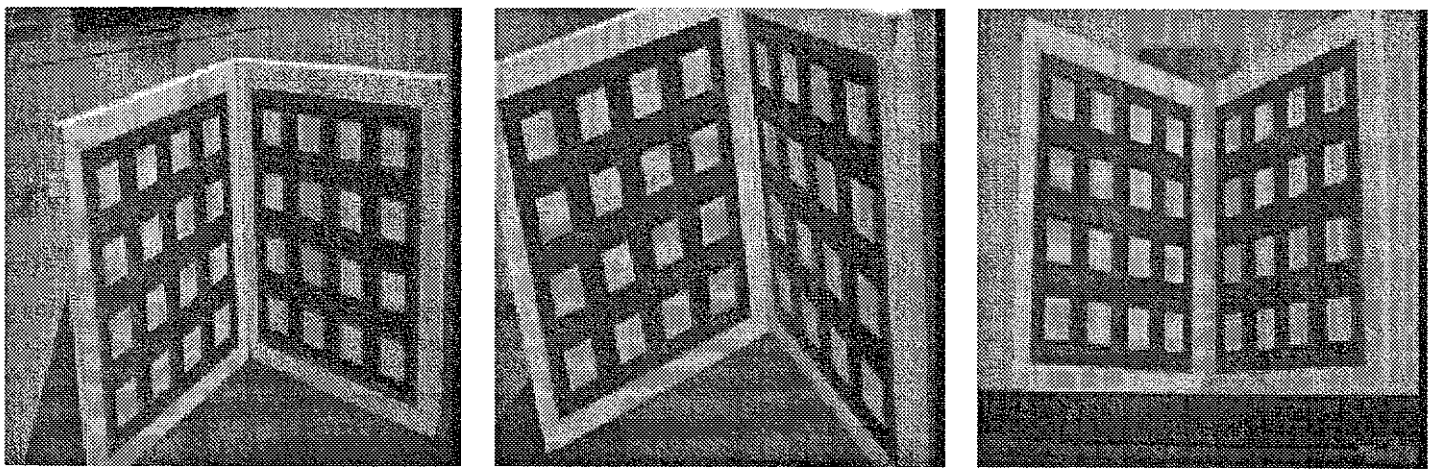


Fig. 8.6: Les images de mire

Une mire de calibration La première série d'image est constituée par la mire couramment utilisée pour effectuer la calibration classique. Les trois images sont montrées figure 8.6. Nous avons utilisé pour l'autocalibration les 128 points d'intérêts constitués par les coins des 32 petits carrés. Ces points ont été extraits par la méthode de Vaillant dans les images 1 et 3, et par la méthode de Blaszka et Deriche dans l'image[3] 2. La précision de détermination est comparable, et très bonne, de l'ordre de 0.2 pixel. Ces points ont aussi servi pour la

[3]La méthode de Vaillant échoue dans cette image car la mire est regardée de manière trop oblique.

calibration classique, qui, effectuée sur chacune des trois images par la méthode de Faugeras et Toscani, a donné les paramètres:

	α_u	α_v	u_0	v_0
image 1	769.30	1171.47	258.52	276.58
image 2	757.80	1153.21	243.20	262.56
image 3	770.09	1172.44	247.12	235.48

Les paramètres intrinsèques obtenus, avec contrainte de perpendicularité, par la méthode des continuations et la méthode itérative sont:

$$\alpha_u = 770.71 \ , \ \alpha_v = 1198.66 \ , \ u_0 = 283.31 \ , \ v_0 = 249.98$$
$$\alpha_u = 781.12 \ , \ \alpha_v = 1219.34 \ , \ u_0 = 292.91 \ , \ v_0 = 240.93$$

Dans cet exemple, si nous n'imposons pas cette contrainte, nous obtenons des résultats très constants, mais dans lesquels le point principal est décentré de manière exagérée:

$$\alpha_u = 833.23 \ , \ \alpha_v = 1313.83 \ , \ u_0 = 335.58 \ , \ v_0 = 230.46 \ , \ c = 0.02$$
$$\alpha_u = 831.14 \ , \ \alpha_v = 1310.02 \ , \ u_0 = 334.00 \ , \ v_0 = 231.33 \ , \ c = 0.02$$

En revanche, en fixant le point principal au centre de l'image, nous trouvons par la méthode itérative: $\alpha_u = 750.20$, $\alpha_v = 1149.96$ qui sont des valeurs assez précises, puisque l'erreur relative n'excède pas 2%.

Il est à noter que dans cet exemple, les points se trouvent dans deux plans, ce qui n'est pas favorable pour la détermination de la matrice fondamentale, comme nous l'avons vu chapitre 5. En revanche l'utilisation de la méthode des plans est très simple et conduit à des calculs plus rapides et des matrices fondamentales un tout petit peu plus précises comme le montrent les épipoles:

	$1 \rightarrow 2$	$2 \rightarrow 3$	$1 \rightarrow 3$
TDIR	[1326 -70] [7873 -1381]	[1361 3982] [-694 -3859]	[-4743 16535] [854 -1144]
GENERAL	[1357 -77] [8634 -1533]	[1330 3854] [-743 -4043]	[-3582 12594] [865 -1157]
PLANS	[1341 -75] [8384 -1492]	[1384 4010] [-680 -3739]	[-4028 14110] [883 -1197]

Si les paramètres intrinsèques obtenus par la méthode des continuations avec les nouvelles matrices fondamentales sont légèrement moins précis, ils varient nettement moins vis-à-vis de la contrainte de perpendicularité:

$$\alpha_u = 815.23 \ , \ \alpha_v = 1241.24 \ , \ u_0 = 310.06 \ , \ v_0 = 219.32 \ , \ c = 4.4.10^{-3}$$
$$\alpha_u = 802.08 \ , \ \alpha_v = 1208.24 \ , \ u_0 = 299.81 \ , \ v_0 = 228.22$$

Le décentrage du point principal est important. En le fixant au centre, nous trouvons de nouveau des valeurs précises: $\alpha_u = 762.26$, $\alpha_v = 1133.13$.

Des fragments de mire Dans cette séquence de six images, nous avons utilisé des ensembles de carrés similaires à ceux formant la mire. Cependant, nous les avons disposés suivant des surfaces courbes, ou occulté une partie. L'autocalibration est facilitée par la présence de ces carrés qui sont une source de points d'intérêts précis. En revanche la calibration classique n'est pas possible telle quelle sur ces images. Trois de ces images sont montrées figure 8.7

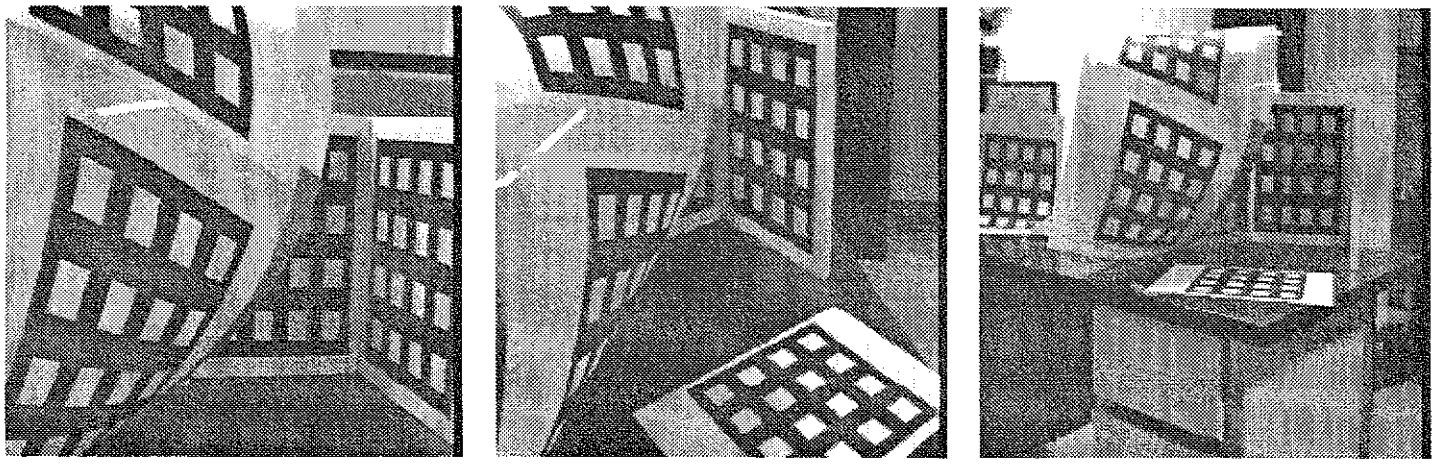


Fig. 8.7: Les images de fragments de mire

Nous avons effectué les cinq mouvements indépendants entre ces six images, et avons varié le nombre de mouvements utilisés de 3 à 5, en essayant à chaque fois d'obtenir tous les paramètres par la méthode itérative. Quelques résultats sont présentés table 8.7. Ils illustrent assez clairement l'amélioration apportée par l'emploi de mouvements multiples: la méthode sur cet exemple devient assez consistente à partir de quatre mouvements, en manifestant une certaine tendance à surévaluer les paramètres.

mouvements	α_u	α_v	u_0	v_0	$\theta - \frac{\pi}{2}$
2 3 4	762.11	1123.21	283.68	197.21	0.06
3 4 5	858.45	1321.54	257.33	282.19	-0.05
4 5 1	640.23	792.84	372.32	312.74	0.35
5 1 2	584.69	639.36	376.39	330.70	0.56
1 2 3	-3.21	-3.20	409.89	268.18	1.57
2 3 4 5	724.65	1231.13	294.11	314.63	-0.00
3 4 5 1	864.70	1308.88	262.32	277.70	-0.01
4 5 1 2	762.96	1125.54	283.99	196.52	0.05
5 1 2 3	688.22	1076.32	320.27	289.32	0.05
1 2 3 4	883.18	1247.74	281.94	260.64	0.15
1 2 3 4 5	809.70	1258.88	292.32	297.70	-0.01

Tableau 8.7: Paramètres intrinsèques calculés sur les fragments de mire

Chapitre 9

Le calcul des paramètres extrinsèques et la reconstruction

Une fois que les paramètres intrinsèques ont été obtenus, nous pouvons déterminer le positionemment relatif, soit de la caméra mobile au cours du temps, soit de celui du système de caméras, si nous en considérons plusieurs.

La première tâche est celle, extrêmement étudiée, du calcul du mouvement tridimensionnel de la caméra. Notre contribution est d'une part, de formuler plusieurs techniques fondées directement ou indirectement sur la matrice fondamentale et d'autre part de mener une vaste étude expérimentale pour étudier la sensibilité des différentes approches à l'erreur sur les paramètres intrinsèques. Nous utilisons ensuite le fait que cette sensibilité n'est pas très grande, si nous décomposons directement les matrices fondamentales calculées, pour proposer une autre approche itérative au problème de l'autocalibration, qui donne des résultats meilleurs que les équations de Kruppa.

Nous introduisons ensuite une nouvelle approche pour la seconde tâche, le calcul des paramètres extrinsèques d'un système de stéréovision. Elle se distingue par le fait qu'elle ne nécessite que les calculs, effectués de manière monoculaire et indépendante pour chacune des caméras, qui ont servi à la calibration de celles-ci. Nous apportons également une solution au problème de la reconstruction à partir de vues multiples, ce qui permet par exemple d'employer des algorithmes de stéréovision trinoculaires.

Ce chapitre s'achève avec des résultats expérimentaux obtenus avec des images réelles, qui illustrent à différentes étapes la faisabilité de notre approche. En particulier nous présentons plusieurs exemples de reconstructions métriques.

9.1 Le calcul du mouvement d'une caméra

Dans la méthode de calibration classique, en général les paramètres extrinsèques sont déterminés en même temps que les paramètres intrinsèques[1]. Il est aussi possible de découpler cette estimation et d'obtenir les paramètres extrinsèques en déterminant directement ce qu'ils représentent: le mouvement tridimensionnel faisant passer du repère d'un objet à celui de la caméra[2]. Dans le cas de l'autocalibration, nous n'utilisons que les repères liés aux caméras. Le calcul des paramètres extrinsèques classiques est donc remplacé par celui de la détermination du mouvement tridimensionnel reliant les positions successives occupées par la caméra au cours du temps[3]. Si les paramètres intrinsèques sont connus, nous pouvons nous placer en coordonnées normalisées, et le problème est alors classique. Nous présentons et comparons très exhaustivement plusieurs méthodes qui sont fondées sur nos techniques d'estimation de la matrice fondamentale. Elles sont analogues à plusieurs techniques existantes [152, 44] [145] [170] [64] [73], notre comparaison expérimentale étant plus systématique que celles disponibles dans la littérature. Nous faisons ensuite également des expérimentations inédites sur leur sensibilité vis-à-vis de la détermination des paramètres intrinsèques.

9.1.1 Méthodes de détermination du mouvement en coordonnées normalisées

Décomposition de la matrice essentielle Une fois que la matrice fondamentale $\mathbf{F}$, ainsi que les paramètres intrinsèques $\mathbf{A}$ sont calculés, nous obtenons immédiatement la matrice essentielle par:

$$\mathbf{E} = \mathbf{A}^T\mathbf{F}\mathbf{A} \tag{9.1}$$

Les paramètres du mouvement peuvent alors être obtenus en décomposant cette matrice. Il s'agit d'un problème classique. Il est connu que la factorisation n'est pas unique, la translation $\mathbf{t}$ ne pouvant être déterminée qu'à un facteur d'échelle algébrique près, et la rotation $\mathbf{R}$ à un angle de π près. L'ambiguïté sur le signe de $\mathbf{t}$ et sur l'angle de $\mathbf{R}$ peut être levée à partir de la donnée d'une seule correspondance: le point reconstruit doit se trouver devant la caméra, aussi bien à l'instant initial qu'à l'instant final. La formulation originale [90] est instable car elle ne prennent pas en compte le fait qu'à cause du bruit, la matrice $\mathbf{E}$ calculée n'est pas exactement une matrice essentielle. Une autre solution explicite élégante [39] souffre du même défaut [134]. Aussi il est préférable de se tourner vers des méthodes permettant de décomposer une matrice vérifiant de manière seulement approchée les conditions de décomposabilité.

La direction de la translation est obtenue de manière approchée de la même façon que

[1] Voir le chapitre 1 pour une présentation de ces méthodes.

[2] *Pose estimation* Il est équivalent de localiser un objet par rapport à la caméra, ou la caméra par rapport à un objet.

[3] Problème dit *Structure from motion*, la reconstruction étant aisée une fois que le mouvement est connu. Plusieurs références ont été données en 4.1.1

les épipoles au (2.15), en résolvant le problème aux valeurs propres:

$$\min_{\|\mathbf{t}\|=1} \|\mathbf{E}^T\mathbf{t}\|^2 \tag{9.2}$$

En effet, comme $\mathbf{Tt} = 0$, nous avons la propriété $\mathbf{E}^T\mathbf{t} = 0$ qui est tout à fait similaire au problème $\mathbf{Fe}' = 0$ déjà rencontré. D'ailleurs $\mathbf{t}$ est obtenu immédiatement à partir de $\mathbf{A}$ par $\mathbf{t} = \mathbf{A}^{-1}\mathbf{e}'$, d'après (2.9).

Le calcul de la rotation est moins direct, et plusieurs solutions ont été présentées [90] [163] [152] [172] [73] [61]. Nous retenons deux méthodes qui se révèlent en fait équivalentes. La première, introduite par Toscani et Faugeras [152, 44, 153] consiste à trouver la rotation $\mathbf{R}$ minimisant l'erreur $C_1 = \|\mathbf{E} - \mathbf{TR}\|^2$. Une solution explicite est obtenue grâce à l'utilisation de quaternions $\mathbf{q}$ pour représenter la rotation. En utilisant leurs propriétés, le critère C_1 s'écrit, après quelques manipulations:

$$C_1 = \sum_{i=1}^{3} \|\mathbf{E}_i - \mathbf{R}^T\mathbf{T}_i\|^2 = \sum_{i=1}^{3} \mid \mathbf{q} \times \mathbf{E}_i - \mathbf{T}_i \times \mathbf{q} \mid^2$$

$\mathbf{E}_i$ (resp. $\mathbf{T}_i$) désignant les trois vecteurs-ligne de $\mathbf{E}$ (resp. $\mathbf{T}$). Le terme $\mathbf{q} \times \mathbf{E}_i - \mathbf{T}_i \times \mathbf{q}$ est linéaire en les quatre coordonnées de $\mathbf{q}$, ce qui permet de formuler le critère précédent comme un problème aux moindres carrés linéaire:

$$\min_{\mathbf{q}} \sum_{i=0}^{3} \mathbf{q}\mathbf{N}_i\mathbf{N}_i^T\mathbf{q}^T \quad \text{avec} \quad \mathbf{N}_i = \begin{pmatrix} 0 & (\mathbf{E}_i - \mathbf{T}_i)^T \\ \mathbf{T}_i - \mathbf{E}_i & \tilde{\mathbf{E}}_i + \tilde{\mathbf{T}}_i \end{pmatrix}$$

La seconde est la classique méthode de Tsai et Huang [163]. Elle consiste à remarquer qu'une décomposition de $\mathbf{E}$ en valeurs singulières $\mathbf{E} = \mathbf{\Delta\Lambda\Theta}^T$, peut aussi s'écrire, compte tenu du fait que $\mathbf{\Delta}$ et $\mathbf{\Theta}$ sont des matrices orthogonales, sous une forme faisant apparaître une matrice antisymétrique et une matrice orthogonale:

$$\mathbf{E} = \underbrace{(\mathbf{\Delta}\mathbf{T}_1\mathbf{\Delta}^T)}_{\mathbf{T}}\underbrace{(\mathbf{\Delta}\mathbf{R}_1\mathbf{\Theta}^T)}_{\mathbf{R}} \tag{9.3}$$

pourvu que $\mathbf{\Delta} = \mathbf{T}_1\mathbf{R}_1$ soit une factorisation en une matrice antisymétrique et une matrice orthogonale. Dans la factorisation précédente, $\mathbf{T}$ et $\mathbf{R}$ ne sont déterminés qu'à un signe près. Ce signe est obtenu[4] en utilisant le fait que la matrice $\mathbf{R}$ représente une rotation, et, est donc de déterminant $+1$. Si $\mathbf{E}$ est une matrice essentielle, d'après (2.6), quitte à permuter des colonnes de $\mathbf{\Delta}$ et $\mathbf{\Theta}$, nous avons $\mathbf{\Delta} = \sigma^2\mathrm{diag}(0,1,1)$, si bien que nous pouvons vérifier facilement que la factorisation:

$$\mathbf{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \qquad \mathbf{R}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

[4]La formulation de [61] est légèrement incorrecte car elle ne tient pas compte de ce problème de signe, qui est pourtant discuté dans [163].

convient. En remplaçant $\mathbf{T}_1$ et $\mathbf{R}_1$ par leur transposées, nous obtenons la seconde factorisation possible. Il est clair que la translation $\mathbf{T}$ obtenue est identique à celle obtenue par la méthode (9.2) décrite précédemment. Par application de la formule (9.3), il vient deux matrices de rotation dont l'angle diffère de π. L'intérêt de cette formulation réside dans le fait, établi par Hartley [61], que la factorisation est optimale au sens où, si nous partons d'une matrice $\mathbf{E}$ qui ne vérifie pas les conditions de décomposabilité, la matrice $\mathbf{E}_1 = \mathbf{TR}$ obtenue par (9.3) est la matrice essentielle minimisant $C_2 = \|\mathbf{E} - \mathbf{E}_1\|^2$. Comme la translation est calculée de manière identique, cette formulation fait apparaître que C_1 et C_2 sont totalement équivalents, ce qui est confirmé de manière indiscutable par les expériences. Nous nous référerons donc à ces méthodes sous l'unique sigle **FACTOR**.

Méthodes itératives La méthode précédente part de l'hypothèse que nous disposons de la matrice fondamentale et est donc fondée *directement* sur son calcul. Nous pouvons aussi exploiter *indirectement* un calcul de matrice fondamentale pour obtenir le mouvement. Les avantages sont une minimisation plus efficace et plus précise car seuls cinq paramètres interviennent. Ceux-ci sont:

- Les trois composantes du vecteur rotation $\mathbf{r}$, dont la direction est celle de l'axe de la rotation, et la norme l'angle de celle-ci.
- Deux des composantes du vecteur translation, la troisième étant fixée à 1. Nous avons en effet trouvé lors de nos premières expérimentations que les paramétrages de $\mathbf{t}$ par sa direction (par exemple les angles θ et ϕ en utilisant des coordonnées sphériques), qui sont théoriquement plus appropriés, donnent de moins bons résultats.

La technique est de minimiser les critères linéaire (4.3), distance aux droites épipolaires (4.9), et critère pondéré par le gradient (4.11), par rapport aux cinq paramètres du mouvement, en substituant $\mathbf{A}^{-1T}\mathbf{TRA}^{-1}$ à $\mathbf{F}$ dans ces critères. Nous nous référerons à ces méthodes sous les sigles respectifs **MIN-LIN**, **MIN-DIST**, et **MIN-GRAD**. La première, bien que dérivée du critère linéaire, est une méthode qui nécessite, comme les deux autres, une minimisation non-linéaire. Les deux dernières reposent sur des distances significatives dans les espace de mesure, qui ont été largement justifiées en 4.2.

Une comparaison à paramètres exacts. Dans la première étude comparative, nous supposons que les paramètres intrinsèques exacts ont été obtenus. Les courbes présentées ont été obtenues à l'aide de statistiques portant sur 200 mouvements. Nous avons représenté pour les composantes rotationnelle et translationnelle du déplacement calculé la moyenne sur les trois composantes, de l'erreur relative moyenne. Signalons néanmoins que les composantes ne sont pas estimées avec une précision identique. En particulier l'erreur sur la composante suivant l'axe Z de la translation est plus grande que celle sur les autres composantes, alors que l'inverse se produit pour la rotation. Un exemple est donné table 9.1.2. La matrice fondamentale servant pour la méthode **FACTOR** a été calculée par le critère non-linéaire **DIST-L**. Si nous nous contentons d'utiliser un critère linéaire, les résultats sont nettement

moins bons. Les méthodes non-linéaires nécessitent une initialisation dont le choix affecte beaucoup les résultats. Nous avons considéré comme mouvement initial:

1. le mouvement exact, afin de tester la précision du minimum,

2. le mouvement obtenu par **FACTOR**, ce qui est l'initialisation réaliste

3. un mouvement fixé arbitraire: $\mathbf{r} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T$, $\mathbf{t} = (0, 0, 1)^T$, afin de tester les propriétés de convergence.

Les résultats obtenus avec l'initialisation exacte sont présentés figure 9.1 et 9.2 respectivement pour les composantes rotationnelle et translationnelle, ceux avec les deux dernières initialisations sont en figure 9.3 et 9.4. Nous pouvons tirer plusieurs conclusions de ces graphes.

- Le calcul est nettement plus stable que celui des épipoles. Il s'agit donc d'un problème beaucoup moins difficile que celui de la détermination de la matrice fondamentale. Le comportement des critères est néanmoins assez voisin. En pratique, la précision la plus élevée qui puisse être obtenue est un peu meilleure que celle de **FACTOR**, qui est déjà bonne.

- La détermination de la rotation est plus précise que celle de la direction de translation.

- Les résultats de la méthode **MIN-LIN** dépendent peu de l'initialisation. Ce sont les moins précis, dès que nous disposons d'une initialisation raisonnable. La convergence demeure toutefois assez bonne dans tous les cas.

- Les résultats de **MIN-DIST** et de **MIN-GRAD** sont pratiquement identiques[5]. Néanmoins **MIN-GRAD** est légèrement plus sujet à divergence.

- Ces résultats sont potentiellement les plus précis, mais sont les plus sensibles à l'initialisation. Ils se dégradent un peu, pour rejoindre de manière proche[6] ceux de **FACTOR** lorsque cette méthode est utilisée pour initialiser, et deviennent très imprécis lorsque l'initialisation se fait avec un mouvement fixé arbitrairement.

[5]Ce qui signifie qu'ils conduisent à une erreur moyenne identique. Les *mouvements* trouvés diffèrent un peu. La différence relative moyenne sur la composante translationnelle est de l'ordre de 0.2% avec l'initialisation 1, presque toujours inférieure à 2% avec l'initialisation 2. En revanche elle vaut 60% avec l'initialisation 3, indépendamment du niveau de bruit.

[6]La différence trouvée par plusieurs auteurs [152] [170] [73] provient du fait qu'ils utilisent un critère *linéaire* pour évaluer **E**, l'algorithme des huit points, obtenant une matrice essentielle bien moins précise que celle obtenue par nos méthodes

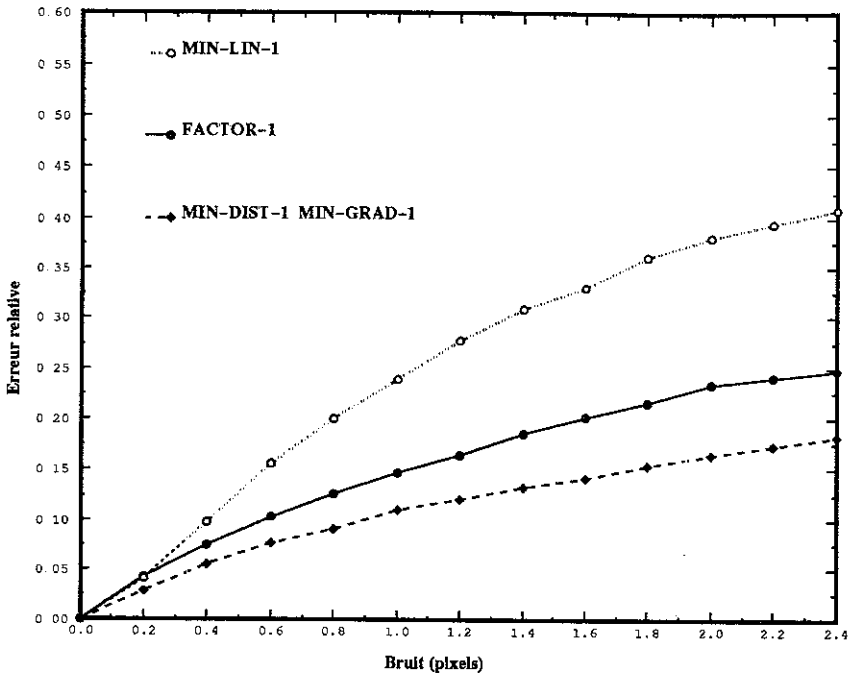


Fig. 9.1: Erreur relative sur la rotation obtenue en initialisant avec les valeurs exactes

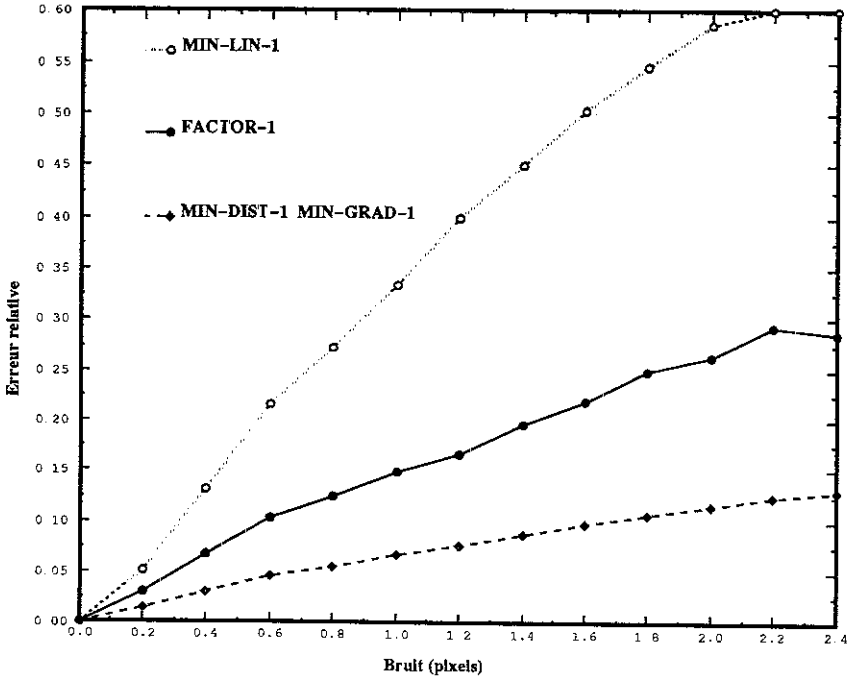


Fig. 9.2: Erreur relative sur la direction de translation obtenue en initialisant avec les valeurs exactes

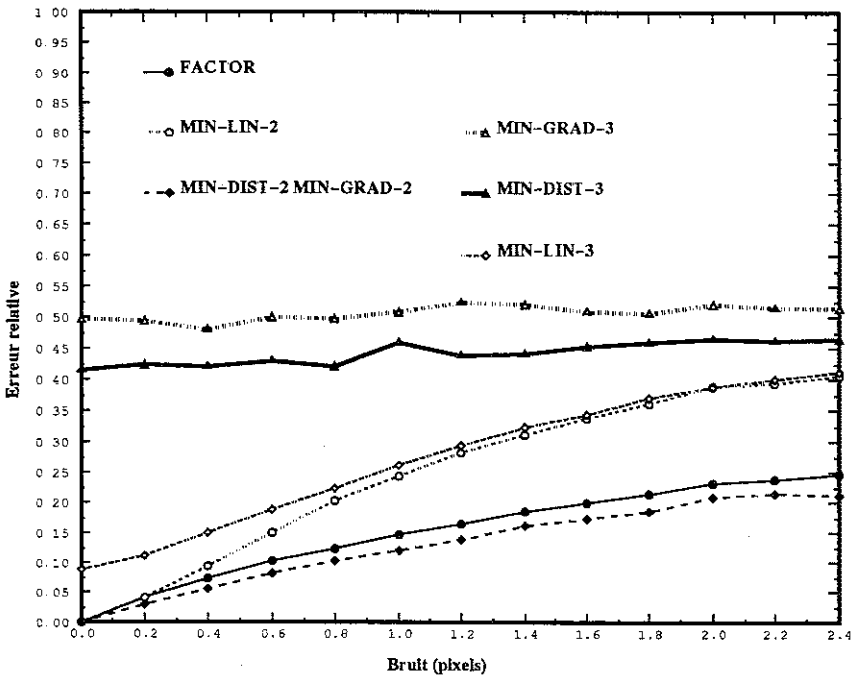


Fig. 9.3: Erreur relative sur la rotation obtenue avec deux initialisations différentes

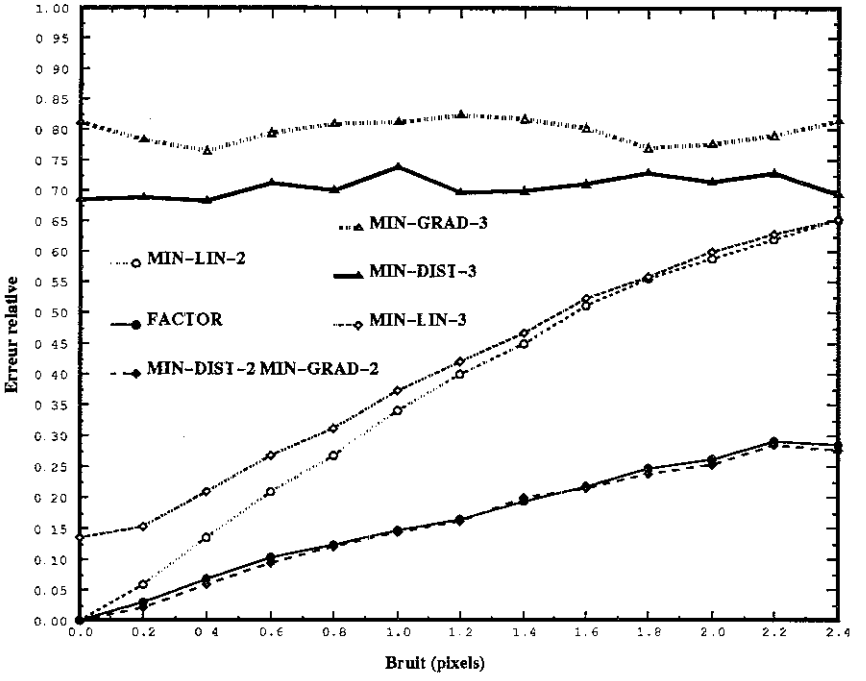


Fig. 9.4: Erreur relative sur la direction de translation obtenue avec deux initialisations différentes

		moyenne des erreurs relatives					
méthode	erreur	r_x	r_y	r_z	t_x	t_y	t_z
FACTOR	sans	0.1370	0.1221	0.0687	0.0889	0.0809	0.1394
	PP 40	0.1964	0.1865	0.0884	0.1375	0.1346	0.1442
	PP 100	0.2598	0.2459	0.1626	0.2530	0.2090	0.1588
	FE 15%	0.2286	0.1976	0.1810	0.1796	0.1724	0.1900
	FE 25%	0.2971	0.2607	0.2477	0.2572	0.2607	0.2550
MIN-LIN	sans	0.1171	0.1173	0.0433	0.0837	0.0785	0.1226
	PP 40	0.2190	0.2004	0.0810	0.1677	0.1505	0.1887
	PP 100	0.3481	0.3374	0.1485	0.3117	0.2743	0.2368
	FE 15%	0.3061	0.2890	0.1797	0.2300	0.2170	0.3202
	FE 25%	0.3922	0.4366	0.3072	0.3214	0.3400	0.4288

Tableau 9.1: Exemples numériques: effet de l'imprécision des paramètres intrinsèques sur le calcul du mouvement

9.1.2 Sensibilité vis-à-vis de l'imprécision sur les paramètres intrinsèques.

Peu d'études ont été effectués sur la sensibilité de différents calculs de mouvement et de structure aux erreurs de détermination des paramètres intrinsèques. Les seuls résultats disponibles sont ceux de Kumar et Hanson [81]. Ils diffèrent de ceux que nous trouvons, le problème étudié n'étant pas identique. C'est pourtant une question importante car elle détermine la précision de calibration qui est nécessaire en fonction de la précision finale souhaitée sur les reconstructions tridimensionnelles qui sont l'objectif final. Nous présentons une comparaison expérimentale qui met en évidence les ordres de grandeur numériques. Nous représentons les erreurs sur la composante rotationnelle (graphes de gauche) et translationnelle (graphes de droite) de la même manière que précédemment. Les graphes supérieurs sont obtenus avec la méthode **FACTOR**, les graphes inférieurs avec la méthode **MIN-DIST**. La méthode **MIN-LIN**, dont nous n'avons pas représenté les résultats, a un comportement intermédiaire vis-à vis de la variation des paramètres intrinsèques, tout en donnant une précision moindre. La figure 9.5 représente l'effet des erreurs du point principal (**PP**). Chaque point représente 100 essais, ils ont été effectué en ajoutant un écart gaussien par rapport au centre de l'image (255, 255), la déviation allant de 20 pixels à 200 pixels. La figure 9.6 représente l'effet des erreurs des facteurs d'échelle (**FE**). Ceux-ci ont subi une déviation standard de 2.5% jusqu'à 25%. Notons que les effets sur les différentes composantes ne sont pas identiques, comme le montre la table 9.1.2, obtenue avec 0.6 pixels de bruit image. Nous n'entrerons pas davantage en détail dans une étude de sensibilité composante par composante, car elle nécessiterait une analyse que nous n'avons pas encore effectuée.

Parmi les nombreuses conclusions qui peuvent être tirés des graphes, nous voudrions souligner les faits suivants:

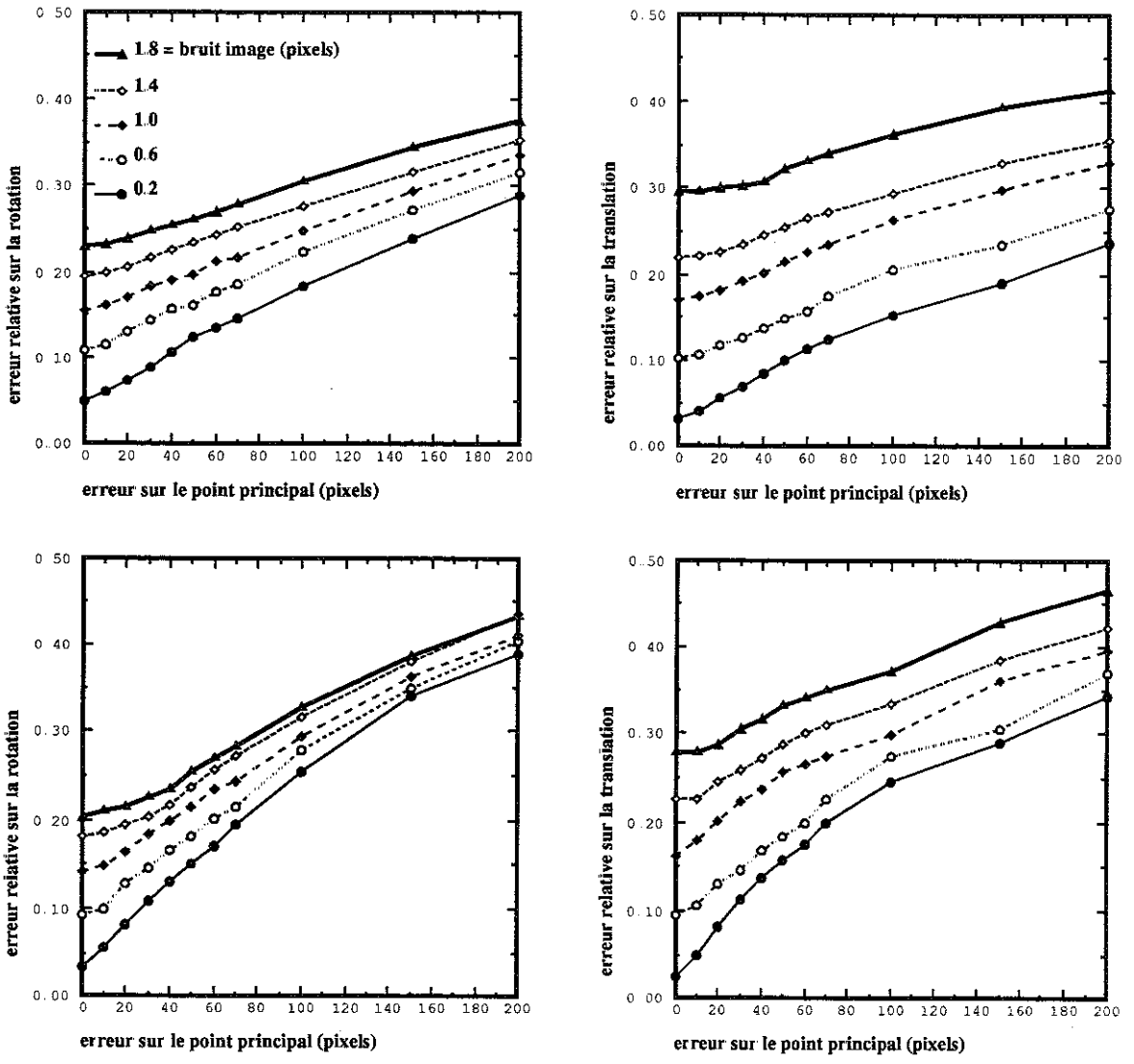


Fig. 9.5: Sensibilité du calcul du mouvement aux erreurs sur le point principal

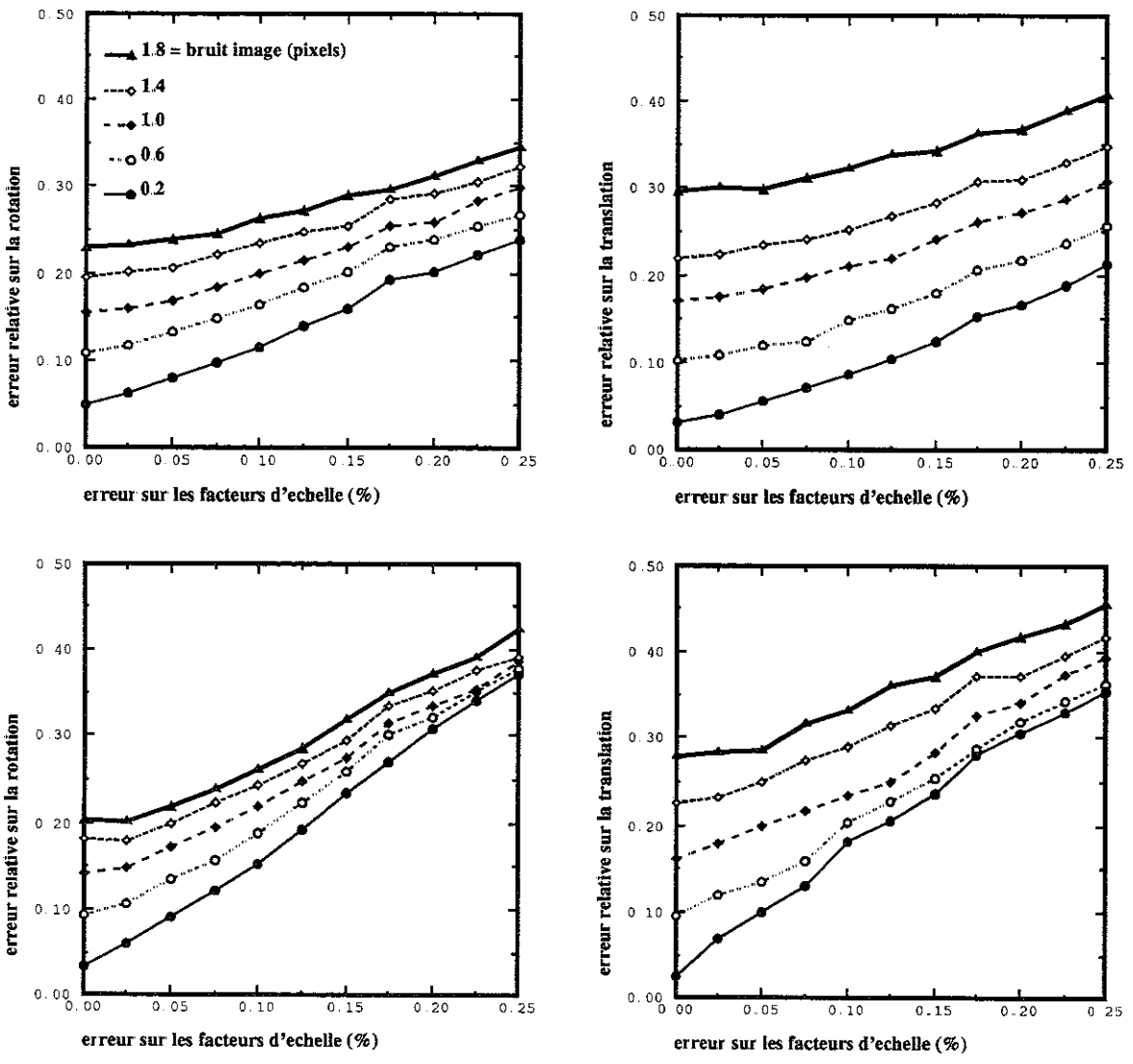


Fig. 9.6: Sensibilité du calcul du mouvement aux erreurs sur le facteur d'échelle

- Les effets de l'imprécision sur les deux types de paramètres sont significatifs, cependant jusqu'à des écarts relativement importants (10% sur les facteurs d'échelle, plusieurs dizaines de pixels de décalage du point principal), ces effets sont moins importants que ceux dus au bruit (par exemple ces imprécisions sur les paramètres ont moins d'effet qu'une augmentation du bruit image de 0.6 à 1 pixel).

- La sensibilité aux erreurs sur les facteurs d'échelle est plus grande que celle aux erreurs de placement du point principal. En effet, en termes d'erreur relative, un déplacement de 120 pixels correspond à 50% d'erreur relative, et a à peu près les mêmes effets que 25% d'erreur sur les facteurs d'échelle.

- Les critères itératifs **MIN-DIST** et **MIN-GRAD** sont davantage sensibles à l'imprécision sur le point principal que la méthode **FACTOR**. Ceci se comprend par le fait que la matrice fondamentale, qui est directement utilisée par **FACTOR**, retient une partie de l'information sur les paramètres intrinsèques exacts. En revanche les critères itératifs compensent entièrement les erreurs sur les paramètres par des erreurs sur le mouvement calculé.

La dernière remarque nous amène à considérer la possibilité de recalculer les paramètres au cours de la minimisation.

9.2 Autocalibration par calcul simultané du mouvement et des paramètres intrinsèques

Nous présentons une généralisation de des méthodes de calcul du mouvement, qui permet d'affiner les paramètres intrinsèques eux-mêmes. Cette méthode se révèle meilleure que celles fondées sur les équations de Kruppa.

9.2.1 Effet de l'ajustement des paramètres avec un seul mouvement

Une extension naturelle des techniques précédentes est de minimiser l'un des critères fondées sur la matrice fondamentale (4.3), (4.9), ou (4.11), par rapport aux cinq paramètres du mouvement, *et aux paramètres intrinsèques* en substituant $\mathbf{A}^{-1T}\mathbf{T}\mathbf{R}\mathbf{A}^{-1}$ à $\mathbf{F}$ dans ces critères. Cela revient en fait à rechercher la matrice fondamentale sous la forme 2.8, en imposant un paramétrage qui tient compte directement des contraintes de rigidité introduites en 7.1.

Bien que nous pouvons disposer d'un nombre de mesures arbitrairement grand, il faut nous souvenir que nous ne pouvons déterminer que sept paramètres, car le calcul s'effectue indirectement en faisant intervenir la matrice fondamentale, qui ne comporte pas davantage de degrés de liberté. Etant donné le décompte d'inconnues qui avait été effectué en 7.1, il est clair qu'il n'est pas possible d'obtenir les quatre paramètres intrinsèques, mais seulement

deux. Le point principal a une importance moindre, qui vient du fait qu'il puisse être fixé au centre de l'image sans erreur très importante, cette erreur elle-même n'affectant pas beaucoup le calcul du mouvement. Aussi, lorsque nous ne disposons que d'un seul mouvement, nous fixons les paramètres $(u_0, v_0)^T$, et faisons varier les facteurs d'échelle α_u et α_v. Nous présentons seulement les résultats obtenus en minimisant le critère **MIN – DIST** à paramétrage modifié pour tenir compte de la variation des facteurs d'échelle. La figure 9.7 montre les erreurs relatives obtenues à différents niveaux de bruit, en fonction l'erreur relative sur les facteurs d'échelle. Elle est à comparer à la partie supérieure de la figure 9.6. A cet effet, nous avons également rappelé sur cette figure l'erreur relative (seulement aux deux niveaux de bruit extrêmes que nous avons considéré) obtenue par la méthode **FACTOR**, laquelle est la moins sensible aux erreurs sur les paramètres intrinsèques. Au vu des graphes, il est tout à fait clair que:

- la nouvelle méthode est très peu sensible à l'effet de l'erreur sur les paramètres intrinsèques, du moins lorsque cette erreur ne dépasse pas 25%,
- plus l'erreur sur les paramètres est grande, et plus le bruit image est faible, plus cette méthode domine la méthode de calcul à paramètres fixés, dans le cas contraire, elle se révèle moins bonne.

La dernière remarque est assez simple à expliquer: si les paramètres dont nous partons sont exacts, en nous autorisant à les varier, nous ne pouvons qu'augmenter l'erreur, si en revanche ils sont très inexacts, le fait de les rapprocher de leur valeur exacte n'est pas très difficile, et améliore le résultat. L'erreur sur les paramètres finaux se trouve ainsi "compensée" par l'erreur sur le mouvement, dans un sens où dans un autre, pour arriver à un niveau à peu près constant, qui ne dépend que du bruit. Ces observations sont confirmées par l'examen des paramètres intrinsèques obtenus, dont l'erreur relative est représentée figure 9.8. Nous pouvons y avoir que l'erreur sur les paramètres finaux ne dépend pratiquement pas de celle sur les paramètres initiaux, mais seulement du niveau de bruit, et ce jusqu'à 25% d'erreur sur les paramètres initiaux. Au delà, des problèmes de convergence apparaissent. Une manière d'interpréter ces résultats est de dire que la précision avec laquelle la méthode peut retrouver les facteurs d'échelle ne dépend, lorsque la méthode converge, que du niveau de bruit image, et que c'est alors cette précision qui détermine celle du calcul du mouvement.

9.2.2 Minimisation globale avec des mouvements multiples

La méthode L'approche précédente ne fait intervenir que deux mouvements, ce qui ne nous permettait que le calcul de deux paramètres intrinsèques. Nous pouvons faire intervenir simultanément plusieurs mouvements, et calculer tous les paramètres intrinsèques. Avec le modèle simplifié à quatre paramètres, deux mouvements sont suffisants pour obtenir ceux-ci, ou seulement les facteurs d'échelle, en surdéterminant le problème. A partir de trois mouvements, le problème est toujours surdéterminé. Etant donné le caractère hautement non-linéaire de la minimisation, et le grand nombre de paramètres sur laquelle elle porte,

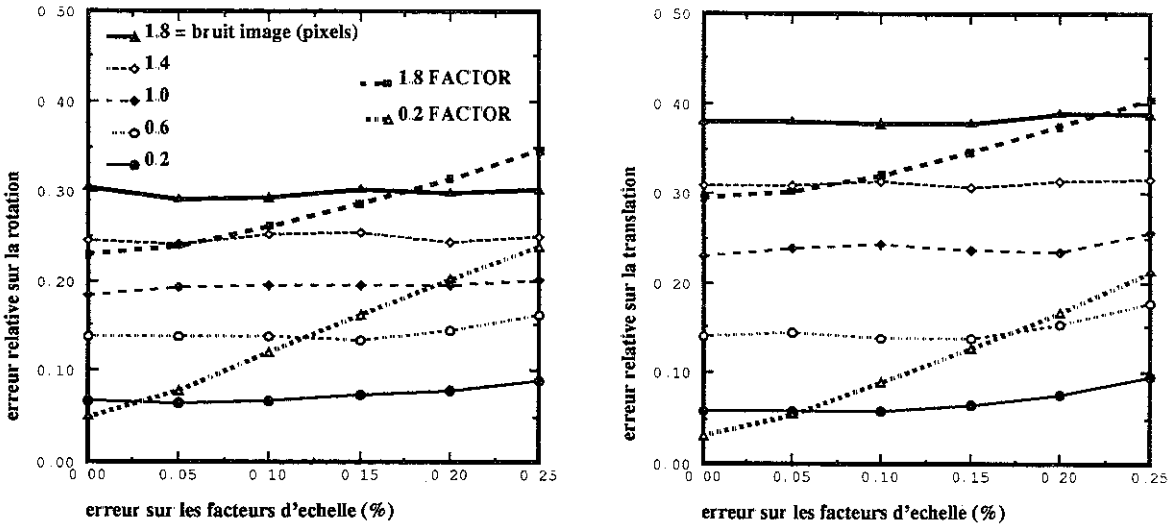


Fig. 9.7: Sensibilité du calcul du mouvement aux erreurs sur les facteurs d'échelle, lors que ceux-ci sont ajustés dans la minimisation

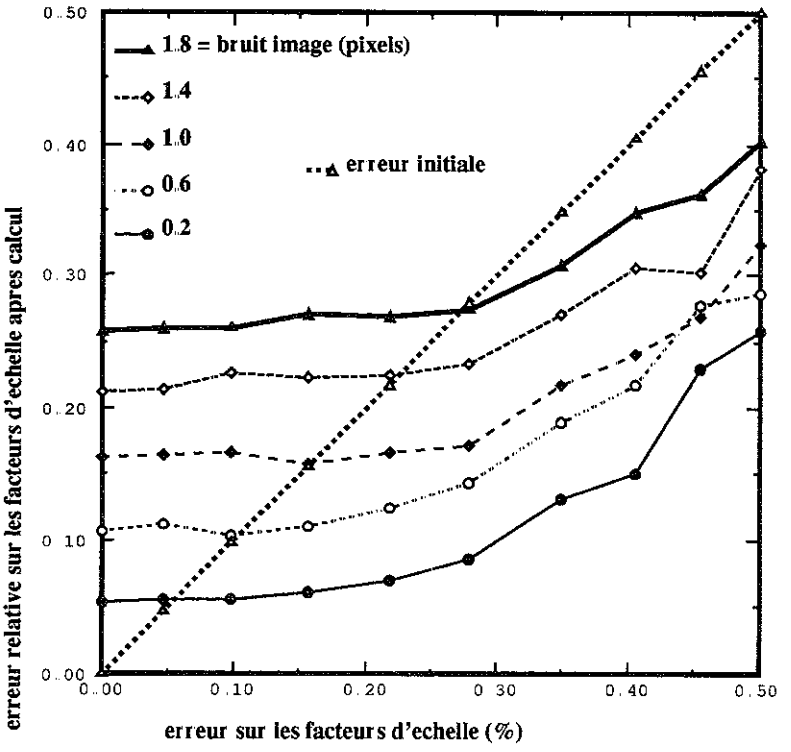


Fig. 9.8: Précision d'ajustement des facteurs d'échelle lors du calcul du mouvement

son succès repose sur l'obtention de bonnes valeurs d'initialisation. Résumons d'abord la méthode utilisant N mouvements ($N \geq 2$), et pour chaque mouvement i, un minimum de huit correspondances $\{(\mathbf{q}_{ij}, \mathbf{q}'_{ij})\}_j$:

Calcul global des paramètres intrinsèques

1. Calculer les N matrices fondamentales $\mathbf{F}_i$: initialisation par le critère linéaire 4.5, calcul itératif par **DIST – L** (4.12,4.9).
2. Partir d'une estimée initiale des paramètres intrinsèques α_u, α_v, u_0, v_0, éventuellement obtenue par les équations de Kruppa (si aucune estimée initiale n'est disponible: 8.2.2, sinon (8.1)).
3. Calculer les N mouvements d'initialisation $(\mathbf{r}_i, \mathbf{t}_i)$ par **FACTOR** (9.2,9.3), à partir de $\mathbf{F}_i$ et des paramètres α_u, α_v, u_0, v_0,
4. Minimiser, par rapport aux $5N+4$ variables (ou $5N+2$ si u_0, v_0 sont fixés au centre de l'image) le critère:

$$\min_{\substack{\alpha_u,\alpha_v,u_0\ v_0 \\ \{\mathbf{r}_i,\frac{t_{ix}}{t_{iz}},\frac{t_{iy}}{t_{iz}}\}_{i=1\ldots N}}} \sum_{i=1}^{N}\sum_j \{d^2(\mathbf{q}'_{ij}, \mathbf{A}^{-1T}\mathbf{T}_i\mathbf{R}_i\mathbf{A}^{-1}\mathbf{q}_{ij}) + d^2(\mathbf{q}_{ij}, \mathbf{A}^{-1T}\mathbf{R}_i^T\mathbf{T}_i\mathbf{A}^{-1}\mathbf{q}'_{ij})\} \tag{9.4}$$

où d est la distance point-droite (4.7), $\mathbf{T}_i = \widetilde{\mathbf{t}}_i$, $\mathbf{R}_i = e^{\mathbf{r}_i}$, et $\mathbf{A} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$

5. Revenir éventuellement à l'étape 3 avec les nouveaux paramètres intrinsèques.

Résultats expérimentaux Nous présentons à présent des résultats qui montrent que la minimisation globale faisant intervenir les mouvements (désignée par le sigle **MOUV**) améliore dans la plupart des cas les résultats obtenus par les équations de Kruppa **KRUPPA**. Dans les deux cas, la même initialisation standard a été utilisée. Nous avons aussi lancé **MOUV** avec les résultats de **KRUPPA**, et utilisé le sigle évident correspondant. Le but est également de montrer l'importance de certains des détails intervenant dans la description précédente. Une comparaison est effectuée au moyen de graphes, présentés figure 9.9. Les graphes de droite sont obtenus en faisant varier les quatre paramètres, ceux de droite en fixant le point principal et en calculant uniquement les deux facteurs d'échelle. Chaque point représente 100 essais, effectués en variant les mouvements et les paramètres intrinsèques comme pour la table 8.4[7]. Les graphes représentent la moyenne des erreurs relatives sur les facteurs d'échelle uniquement, en fonction du bruit image. L'erreur sur le point principal est

[7] Les erreurs sont plus grandes car nous avons utilisé moins de points: 30

		moyenne des erreurs relatives			
mouv	méthode	α_u	α_v	u_0	v_0
2	**KRUPPA**	0.3601	0.3556	0.5247	0.5892
	KRUPPA-PP	0.2863	0.2938	0.0681	0.0589
	MOUV	0.2368	0.2264	0.4433	0.5825
	MOUV-PP	0.1348	0.1680	0.0681	0.0589
	MOUV-KRUPPA	0.3043	0.2996	0.5262	0.5915
	MOUV-KRUPPA-PP	0.1514	0.2047	0.0681	0.0589
5	**KRUPPA**	0.2800	0.2909	0.4673	0.5461
	KRUPPA-PP	0.3105	0.3169	0.0681	0.0589
	MOUV	0.1131	0.0953	0.2794	0.3197
	MOUV-PP	0.1128	0.0998	0.0681	0.0589
	MOUV-KRUPPA	0.0878	0.0751	0.2664	0.2943
	MOUV-KRUPPA-PP	0.1054	0.0991	0.2683	0.2482

Tableau 9.2: Exemples numériques: différentes méthodes de calcul des paramètres intrinsèques

	FACTOR-KRUPPA (3)		**MOUV-KRUPPA** (4)		**FACTOR-MOUV-KRUPPA** (5)	
	r	t	r	t	r	t
2	0.4661	0.3904	0.4678	0.3854	0.4444	0.3639
2 **PP**	0.3431	0.2812	0.3398	0.3082	0.2639	0.2876
5	0.4336	0.3731	0.3693	0.3215	0.2800	0.1953
5 **PP**	0.3549	0.2877	0.3145	0.2782	0.2577	0.2092

Tableau 9.3: Exemples numériques: évolution des erreurs sur le mouvement

assez corrélée, et se révèle toujours plus importante, sauf si nous fixons le point principal, comme le montre la table 9.2.2, qui présente les erreurs sur les paramètres obtenues à 1 pixel de bruit. La table 9.2.2, qui a été calculée au même niveau de bruit, montre que le calcul du mouvement s'améliore effectivement au cours des différentes étapes de l'algorithme.

L'ensemble de ces données permettent de constater que la nouvelle méthode améliore nettement les résultats. Plus nous disposons de mouvements, plus la différence est grande, et plus il devient intéressant d'initialiser tout de même par **KRUPPA**. Les meilleurs résultats sont encore obtenus en fixant le point principal au centre, sauf dans des conditions favorables: bruit faible, nombre de mouvements grand.

Essayons d'expliquer la différence de comportement des méthodes. Les premières étapes du calcul sont identiques pour les deux méthodes, et consistent en la détermination des matrices fondamentales. Lors de la phase de détermination des paramètres intrinsèques, la méthode fondée sur les équations de Kruppa n'utilise plus que ces matrices, la contrainte de

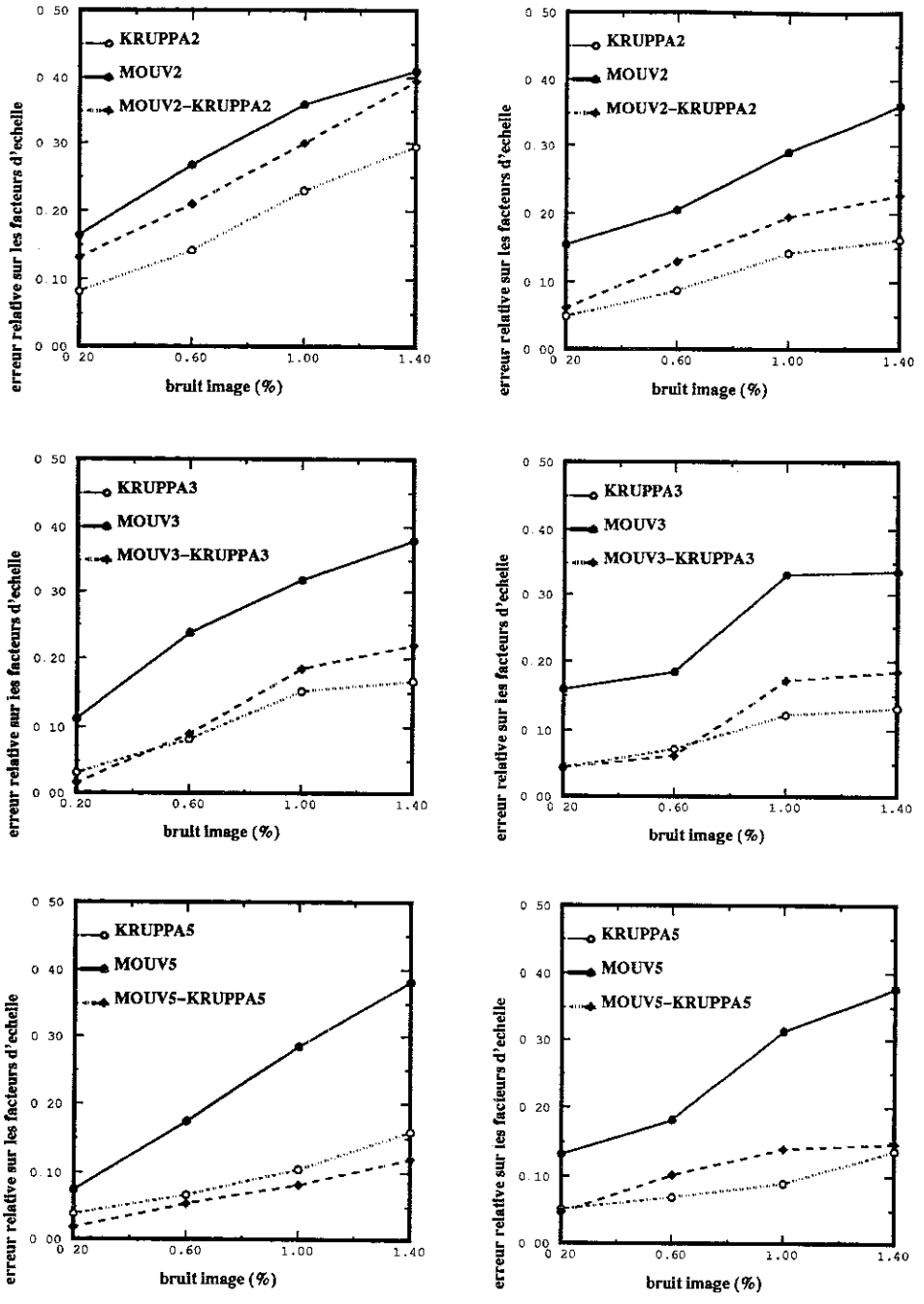


Fig. 9.9: Comparaison des équations de Kruppa et de la méthode avec calcul du mouvement. Gauche 4 paramètres. Droite: 2 paramètres.

rigidité étant employée pour *éliminer* les mouvements inconnus. Il en résulte une méthode qui ne fait intervenir que les inconnues recherchées, d'où la possibilité d'une solution semi-analytique, et la rapidité des formulations itératives. En revanche, dans la méthode fondée directement sur la décomposition de la matrice fondamentale que nous venons de présenter, c'est la *forme de la paramétrisation* qui assure la validité des paramètres recherchés vis-à-vis de toutes les contraintes disponibles. Il est alors nécessaire de faire appel à toutes les variables inconnues en calculant également le mouvement, nécessite de disposer d'une bonne itération, et peut conduire à des calculs itératifs plus lourds. En contrepartie, d'une part nous disposons dans le critère minimisé de toutes les contraintes possibles liées au problème, d'autre part, nous faisons intervenir de nouveau toute l'information disponible, sous forme d'appariements.

9.3 Méthodes pour le calcul des paramètres extrinsèques en stéréovision

Jusqu'à présent, nous n'avons considéré qu'une caméra unique. Si nous utilisons un système de stéréovision, nous pouvons ainsi autocalibrer séparément chacune des caméras. Nous expliquons d'abord en quoi consistent les paramètres extrinsèques du système dans un cadre d'autocalibration, puis nous présentons une méthode qui ne nécessite pas l'établissement de correspondance d'une caméra à l'autre, et peut être conduite à partir des données qui ont été utilisées pour autocalibrer séparément chaque caméra. Nous considérons pour finir le problème posé par les systèmes trinoculaires.

9.3.1 Les paramètres extrinsèques d'un système binoculaire

Nous avons vu en 1.1.1 qu'une caméra est modélisée par 5 paramètres intrinsèques contenus dans la matrice $\mathbf{A}$ et 6 paramètres extrinsèques contenus dans la matrice $\mathbf{D}$. Un système de stéréovision classique[8] est constitué de deux caméras liées de manière rigide. Si nous faisons le décompte des paramètres obtenus dans le cadre de la calibration classique, nous en trouvons vingt-deux: les paramètres intrinsèques de chacune des caméras, $\mathbf{A}_1$ et $\mathbf{A}_2$, ainsi que leur paramètres extrinsèques, $\mathbf{D}_1$ et $\mathbf{D}_2$. Cependant, le choix d'un système d'un système de coordonnées particulières est arbitraire, les propriétés métriques étant par définition invariantes par changement de repère orthonormé. Comme nous n'utilisons pas de repère attaché à un objet de l'environnement, puisque nous ne supposons aucune connaissance de celui-ci, nous devons prendre comme repère de référence celui attaché à l'une des caméras (dite arbitrairement *première caméra*), prise dans la position initiale. Les seuls paramètres extrinsèques significatifs pour un système de stéréovision sont donc ceux exprimant la position et l'orientation *relative* des deux caméras, sous la forme d'une matrice de déplacement $\mathbf{D}$ représentant

[8]Par opposition aux système actifs avec modification de la vergence ou autres mouvements "oculaires" [79] [175]. Ces systèmes étant particulièrement délicats à calibrer, il est également possible d'effectuer de la stéréovision active sans modifier le positionnement relatif des deux caméras [93, 94].

le déplacement du repère de la première caméra vers celui de la seconde caméra, exprimé dans le repère de la première caméra. Les déplacements sont schématisés figure 9.10. Aussi, une fois que nous avons autocalibré chacune des caméras individuellement, nous devons juste calculer cette matrice pour achever la calibration du système de stéréovision.

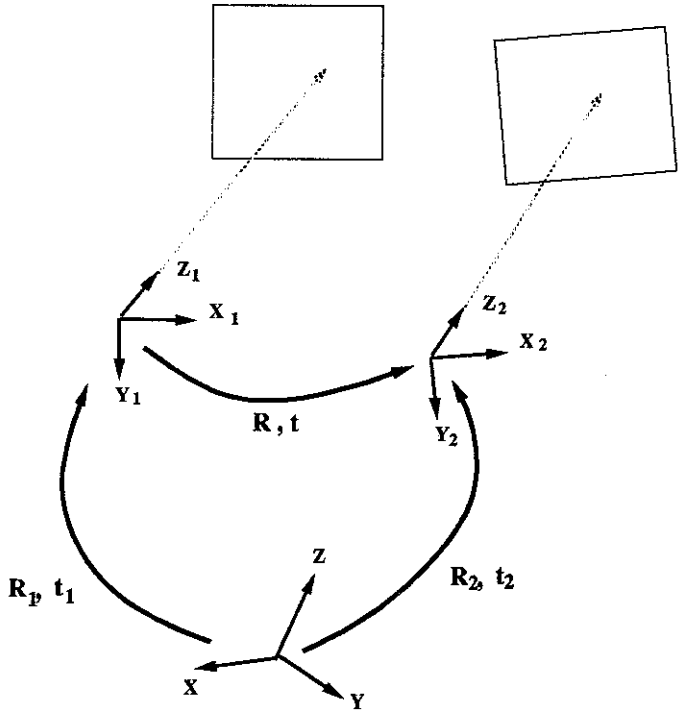


Fig. 9.10: Les paramètres intrinsèques d'un système de stéréovision

Exploiter la nature du problème Une méthode possible serait d'utiliser la technique de calcul du mouvement présentée précédemment, d'une caméra à l'autre. Il suffirait pour cela de considérer des paramètres différents dans chaque image. Cette méthode présente une particularité très attrayante: le mouvement entre les deux caméras étant fixé, *il est possible d'accumuler des appariements entre les deux images pris à des instants différents*, ce qui permet d'effectuer un calcul avec un très grand nombre d'appariements et d'obtenir des résultats très précis.

9.3.2 Une méthode monoculaire pour calibrer un système binoculaire

L'inconvénient de l'approche précédente est cependant d'obliger à établir des correspondances d'une caméra à l'autre, ce qui est moins facile que d'obtenir des correspondances temporelles monoculaires. En effet, le mouvement est fixe, il n'est pas possible d'utiliser des

positions intermédiaires, et de plus les caractéristiques des deux caméras peuvent différer. Un second inconvénient est qu'il est moins facile d'utiliser la méthode de raffinement des paramètres intrinsèques car le nombre de ceux-ci est alors doublé. Nous présentons une méthode pour contourner ces difficultés.

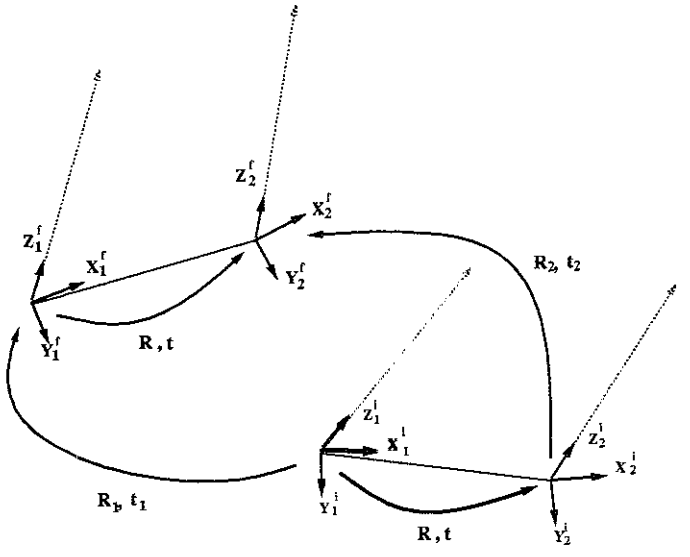


Fig. 9.11: Déplacement d'un système de stéréovision

Tout comme pour l'autocalibration monoculaire, nous effectuons des déplacements inconnus du système stéréo. Nous partons du fait que nous avons pu autocalibrer les caméras, et que nous avons pu calculer, par l'une des méthodes décrites précédemment, le déplacement $\mathbf{D}_1$ de la première caméra, exprimé dans son propre repère, et le déplacement $\mathbf{D}_2$ de la seconde caméra, également exprimé dans son propre repère. La difficulté provient du fait que $\mathbf{D}_1$ and $\mathbf{D}_2$ sont alors connus dans des repères différents. Pour surmonter cette difficulté, l'idée est d'exploiter la situation schématisée figure 9.11 en utilisant la commutativité du diagramme suivant:

$$\begin{array}{ccc} & C_1^f \xrightarrow{\mathbf{D}} C_2^f & \\ \mathbf{D}_1 \uparrow & & \uparrow \mathbf{D}_2 \\ & C_1^i \xrightarrow{\mathbf{D}} C_2^i & \end{array}$$

dans lequel $\mathbf{D}$, désigne le déplacement faisant passer du repère de la première caméra, à celui de la seconde caméra, exprimé dans le repère de la première caméra. Comme par hypothèse, la position relative de la première et de la seconde caméra ne changent pas au cours du temps, de la première position à la seconde, puisque les deux caméras sont fixées rigidement, nous pouvons écrire l'équation matricielle:

$$\mathbf{D}\mathbf{D}_1 = \mathbf{D}_2\mathbf{D} \tag{9.5}$$

où $\mathbf{D}$ est la matrice 4×4 du déplacement inconnu de la première caméra vers la seconde caméra, $\mathbf{D}_1$ et $\mathbf{D}_2$ sont les matrices 4×4 de déplacement de la première et de la seconde ca-

méra, dans leurs repères respectifs. L'équation (9.5) peut se décomposer en les deux équations matricielles suivantes:

$$\mathbf{R}\mathbf{R}_1 = \mathbf{R}_2\mathbf{R} \tag{9.6}$$

$$(\mathbf{I} - \mathbf{R}_2)\mathbf{t} = \mu_2\mathbf{t}_2 - \mu_1\mathbf{R}\mathbf{t}_1 \tag{9.7}$$

où μ_1 et μ_2 sont les facteurs d'échelle inconnus associés à $\mathbf{D}_1$ et $\mathbf{D}_2$, respectivement. Nous montrons plus bas que si deux déplacements du système stéréo sont effectués, alors il est possible de résoudre d'abord les deux équations matricielles (9.6) résultant de ces deux déplacements pour obtenir $\mathbf{R}$, puis les deux équations vectorielles (9.7) pour obtenir $\mathbf{t}$ à un facteur d'échelle près. Des travaux similaires ont été effectués en robotique dans un autre contexte, celui de la calibration "main-oeil" [20] [140], [164], [19], et nous renvoyons le lecteur à ces références pour des études approfondies, en particulier en ce qui concerne l'unicité des solutions et des méthodes permettant de combiner un grand nombre de mouvements. Notons cependant que si 9.6 est un problème bien connu en robotique, l'autre équation ne l'est pas (9.7) car c'est le contexte d'autocalibration qui nous conduit à effectuer les calculs à un facteur d'échelle près.

Calcul de la rotation

Pour résoudre l'équation (9.6), nous utilisons une représentation des rotations sous forme de quaternions:

$$\begin{aligned} \mathbf{R} &= (s, \mathbf{v}) \\ \mathbf{R}_1 &= (s_1, \mathbf{v}_1) \\ \mathbf{R}_2 &= (s_2, \mathbf{v}_2) \end{aligned}$$

Nous esquissons notre formulation de la solution à ce problème, qui peut être considéré comme classique. En écrivant l'équation (9.6) avec cette représentation, on obtient les deux équations:

$$\begin{aligned} s(s_1 - s_2) &= \mathbf{v}\,(\mathbf{v}_1 - \mathbf{v}_2) \\ s(\mathbf{v}_1 - \mathbf{v}_2) + (s_1 - s_2)\mathbf{v} + \mathbf{v} \wedge (\mathbf{v}_1 + \mathbf{v}_2) &= 0 \end{aligned} \tag{9.8}$$

Posons: $\mathbf{v} = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 + \gamma(\mathbf{v}_1 \wedge \mathbf{v}_2)$. Quelques manipulations avec les équations (9.9) en utilisant les propriétés des quaternions donnent:

$$\alpha = \beta$$

$$s_1 = s_2 \tag{9.9}$$

puis:

$$s = \gamma(\mathbf{v}_1^2 + \mathbf{v}_1.\mathbf{v}_2) \tag{9.10}$$

$$\alpha^2\|\mathbf{v}_1 + \mathbf{v}_2\|^2 + \gamma^2(\|\mathbf{v}_1 \wedge \mathbf{v}_2\|^2 + \frac{\|\mathbf{v}_1 + \mathbf{v}_2\|^4}{4}) = 1 \tag{9.11}$$

En utilisant le repère défini par:

$$\begin{aligned} \mathbf{u}_1 &= \frac{\mathbf{v}_1 + \mathbf{v}_2}{\|\mathbf{v}_1 + \mathbf{v}_2\|} \\ \mathbf{u}_2 &= \frac{\mathbf{v}_1 \wedge \mathbf{v}_2}{\|\mathbf{v}_1 \wedge \mathbf{v}_2\|} \\ \mathbf{u}_3 &= \mathbf{u}_1 \wedge \mathbf{u}_2 \end{aligned}$$

l'équation (9.11) peut s'écrire plus simplement:

$$x^2 + (1 + k^2)y^2 = 0 \tag{9.12}$$

où $x = \alpha\|\mathbf{v}_1 + \mathbf{v}_2\|$, $y = \gamma\|\mathbf{v}_1 \wedge \mathbf{v}_2\|$, et $k = \frac{\|\mathbf{v}_1+\mathbf{v}_2\|^2}{2\|\mathbf{v}_1\wedge\mathbf{v}_2\|}$. Cette dernière équation détermine une famille de rotations à un degré de liberté, qui est paramétrisée par une ellipse du plan $(\mathbf{u}_1, \mathbf{u}_2)$.

En utilisant un second mouvement, nous obtenons une ellipse du plan $(\mathbf{u}'_1, \mathbf{u}'_2)$. Quelques calculs montrent que dans le cas général, il y aura une intersection non-vide des deux ellipses, la première définie par l'équation (9.12) et la seconde étant définie de la même manière avec des primes, si la condition suivante est vérifiée:

$$k_{12}(\mathbf{v}_3.\mathbf{v}'_1)^2 = k^2(\mathbf{v}'_3.\mathbf{v}_1)^2 \tag{9.13}$$

Dans ce cas, une solution explicite peut être facilement calculée pour l'intersection des ellipses (9.12), donnant α et γ. Nous obtenons ensuite immédiatement $\mathbf{v}$ et s par les relations (9.10) et (9.11).

Voyons à présent pourquoi lors du déplacement d'un système stéréo, la condition (9.13) est toujours satisfaite. Supposons que nous effectuons deux déplacements au cours desquels nous avons donc:

$$\begin{aligned} \mathbf{R}\mathbf{R}_1 &= \mathbf{R}_2\mathbf{R} \\ \mathbf{R}\mathbf{R}'_1 &= \mathbf{R}'_2\mathbf{R} \end{aligned}$$

De ces deux équations, on peut aussi tirer:

$$\mathbf{R}(\mathbf{R}_1\mathbf{R}'_1) = (\mathbf{R}_2\mathbf{R}'_2)\mathbf{R}$$

En utilisant la contrainte (9.9) à partir de la dernière équation, il vient: $s_1s'_1 - \mathbf{u}_1.\mathbf{u}'_1 = s_2s'_2 - \mathbf{u}_2.\mathbf{u}'_2$. comme on a aussi les contraintes $s_1 = s_2$ et $s'_1 = s'_2$, on obtient finalement la relation:

$$\mathbf{u}_1.\mathbf{u}'_1 = \mathbf{u}_2.\mathbf{u}'_2 \tag{9.14}$$

Quelques calculs simples suffisent à montrer que cette équation est équivalente à la contrainte (9.13).

Calcul de la translation

Supposons à présent que nous avons calculé $\mathbf{R}$, à l'aide de la méthode précédemment décrite. L'image de la transformation $\mathbf{I} - \mathbf{R}_2$ est le plan perpendiculaire à l'axe $\mathbf{u}_2$ de la rotation $\mathbf{R}_2$. Aussi, en partant de la relation (9.7), nous pouvons écrire:

$$\mathbf{u}_2 \cdot (\mu_2 \mathbf{t}_2 - \mu_1 \mathbf{R}\mathbf{t}_1) = 0$$

Ceci nous permet de déterminer le rapport $a = \mu_1/\mu_2$. Il est alors possible de calculer la direction $\mathbf{t}_\perp$ de la composante de $\mathbf{t}$ orthogonale à $\mathbf{u}_2$, ce qui nous donne la contrainte sur $\mathbf{t}$ suivante:

$$\mathbf{t} \in \mathrm{Vect}\langle \mathbf{t}_\perp, \mathbf{u}_2 \rangle \tag{9.15}$$

Si un second mouvement, *pour lequel l'axe* $\mathbf{u}'_2$ *de la rotation est différent*, est effectué, nous pouvons calculer de manière similaire une direction $\mathbf{t}'_\perp$ et en déduire une contrainte:

$$\mathbf{t} \in \mathrm{Vect}\langle \mathbf{t}'_\perp, \mathbf{u}'_2 \rangle \tag{9.16}$$

En combinant les deux contraintes (9.15) et (9.16) nous déterminons $\mathbf{t}$ à un facteur d'échelle par:

$$\mathbf{t} = \lambda(\mathbf{t}_\perp \wedge \mathbf{u}_2) \wedge (\mathbf{t}'_\perp \wedge \mathbf{u}'_2)$$

9.3.3 Détermination des matrices de projection et utilisation de mouvements multiples

Nous montrons comment, à partir de mouvements calculés à un facteur d'échelle près, il est possible d'obtenir des matrices de projection permettant une reconstruction métrique. Un bénéfice de cette étude sera une petite amélioration de la méthode d'autocalibration par calcul global du mouvement qui est possible dans certaines circonstances.

Cas d'un mouvement simple Si nous voulons reconstruire à partir d'un système binoculaire (ou de deux vues), il suffit d'utiliser les matrices de projection:

$$\mathbf{P} = [\mathbf{A} \ , \ 0] \quad \mathbf{P}' = [\mathbf{A}' \ , \ 0]\mathbf{D}^{-1} \tag{9.17}$$

Ne connaissant que la direction de la translation, nous représentons $\mathbf{t}$ par un vecteur de norme 1. La reconstruction s'effectue alors à un facteur d'échelle inconnu près. Une seule information métrique, sur le mouvement ou sur une longueur vue dans une image suffit à obtenir le facteur d'échelle.

Mouvements multiples: le cas trinoculaire Dans l'hypothèse où nous voulons reconstruire à partir de trois vues (numérotées 1, 2, et 3) les déplacements $\mathbf{D}_{12}$ et $\mathbf{D}_{23}$ n'étant connus qu'à un facteur d'échelle près, l'utilisation de la formule (9.17) pour obtenir $\mathbf{P}_2$ et $\mathbf{P}_3$ conduit à un résultat incorrect, dans lequel, en particulier, la contrainte épipolaire n'est pas

respectée entre les images 1 et 3. La raison de cela est que le rapport $\frac{\|\mathbf{t}_{12}\|}{\|\mathbf{t}_{23}\|}$ doit être respecté, ainsi que les signes relatifs. La difficulté provient du fait, remarqué précédemment, que si nous disposons de deux déplacements connus seulement à un facteur d'échelle près, il n'est possible d'obtenir que la rotation:

$$\mathbf{R}_{13} = \mathbf{R}_{23}\mathbf{R}_{12} \tag{9.18}$$

mais pas de déterminer la direction de translation de la composée $\mathbf{D}_{12}\mathbf{D}_{23}$, la seule contrainte étant que celle-ci appartienne au plan $\text{Vect}\langle \mathbf{t}_{23}, \mathbf{R}_{23}\mathbf{t}_{12} \rangle$. Afin de déterminer le rapport algébrique des translations, il faut donc faire intervenir comme précédemment la commutativité, en calculant d'abord directement le déplacement $\mathbf{D}_{13}$ à partir des images. C'est une tâche raisonnable, car si nous voulons reconstruire à partir des images 1, 2, et 3, il doit y avoir une portion de la scène visible dans l'image 1 et dans l'image 3. Soit λ, le rapport *algébrique* de $\mathbf{t}_{12}$ par $\mathbf{t}_{23}$. En exprimant la condition de proportionnalité:

$$\mathbf{t}_{13} \wedge (\mathbf{R}_{23}\mathbf{u}_1 + \lambda \mathbf{u}_2) = 0 \tag{9.19}$$

où $\mathbf{u}_1 = \frac{\mathbf{t}_{12}}{\|\mathbf{t}_{12}\|}$ et $\mathbf{u}_2 = \frac{\mathbf{t}_{23}}{\|\mathbf{t}_{23}\|}$. Nous obtenons:

$$\lambda = -\frac{(\mathbf{t}_{23} \wedge \mathbf{R}_{23}\mathbf{u}_1)_1}{(\mathbf{t}_{23} \wedge \mathbf{u}_2)_1} = -\frac{(\mathbf{t}_{23} \wedge \mathbf{R}_{23}\mathbf{u}_1)_2}{(\mathbf{t}_{23} \wedge \mathbf{u}_2)_2} = -\frac{(\mathbf{t}_{23} \wedge \mathbf{R}_{23}\mathbf{u}_1)_3}{(\mathbf{t}_{23} \wedge \mathbf{u}_2)_3} \tag{9.20}$$

En prenant $\mathbf{t}_{12} = \mathbf{u}_1$, $\mathbf{t}_{23} = \lambda \mathbf{u}_2$, nous obtenons alors à l'aide de (9.17) des matrices de projections mutuellement cohérentes, qui permettent de reconstruire à un facteur d'échelle près.

Le calcul du mouvement revisité Si les mouvements sont exacts, l'équation (9.20) est exactement vérifiée. Lorsque que les mouvements sont calculés de manière approchée, la double égalité n'est pas satisfaite. Le mieux est alors de prendre pour λ la moyenne des trois valeurs, et de considérer la différence de celles-ci comme un indicateur d'erreur très significatif, puisqu'il porte directement sur les mouvements. Nous pouvons cependant modifier l'algorithme introduit au (9.4) pour contraindre les mouvements trouvés à vérifier (9.20). Si nous disposons de correspondances 1-2, 2-3, et 3-1, au lieu de paramétrer le mouvement 3-1 par les cinq valeurs classiques, nous employons un paramètre unique λ. Le mouvement 3-1 peut alors être calculé à partir des mouvements 1-2 et 2-3 par les relations (9.18) et (9.19). L'initialisation de λ se fait par la moyenne. L'avantage de cette technique est, tout en réduisant le nombre de paramètres de quatre, d'augmenter les contraintes sur les déplacements. Elle permet de plus de maximiser la cohérence des géométries épipolaires obtenues à partir des matrices de projection qui en sont tirées, ce qui est important pour l'utilisation d'algorithmes de stéréovision. A cause de cette caractéristique, nous avons employé cette méthode pour obtenir les reconstructions trinoculaires présentées dans la section suivante. Néanmoins, dans les cas réels que nous avons traités, nous avons remarqué que cette approche, si elle semble améliorer le calcul du mouvement, n'a pas d'effet positif sur la précision des paramètres intrinsèques.

9.4 Exemples réels

9.4.1 Calcul des paramètres intrinsèques: une série d'images avec variation de la focale

Pour compléter les expériences effectuées en 8.4.3, nous avons effectué une autre série d'expérimentations en utilisant une autre caméra. Il s'agit de la caméra Tri-CCD munie d'un objectif à focale variable pour laquelle une table de résultats avait été montrée en 1.2. Nous avons travaillé avec quatre focales différentes: neuf, douze, vingt, et trente. Un échantillon d'images correspondant à ces focales respectives est montré figure 9.12. Il s'agit de scènes d'intérieur sans aucun objet pouvant servir à la calibration classique. Pour chacune des séries, nous avons travaillé avec cinq vues, ce qui nous a permis de disposer d'un peu moins d'une dizaine de mouvements. Les points d'intérêt ont été obtenus de manière semi-automatique par le détecteur de Blaszka et Deriche [13], qui s'est avéré très fiable, A partir de ces points, nous avons calculé les matrices fondamentales par la méthode non-linéaire. Les erreurs moyennes obtenues sont presque toujours comprises entre 0.25 pixels et 0.7 pixels. Il est à noter que nous en avons utilisé avec succès la version robuste permettant de rejeter des faux appariements. Un exemple de telle détection est montré figure 9.13, qui est également typique vis-à-vis du type de correspondances qu'il est possible d'obtenir dans ce type de scènes. Les faux appariements consistent en l'inversion des points 10 et 11. Bien que ces points soient apparemment très rapprochés, le critère est assez élevé, à 5.31 pixel. Le M-estimateur élimine ces points en trois itérations, et fait passer le critère à 0.62, avec un changement significatif de la position des épipoles.

La table 9.4 montre les résultats de calcul des paramètres intrinsèques. Ils sont à comparer avec ceux d'une calibration classique effectuée pour référence, mais également avec ceux de la table 1.2, qui, étant des moyennes, sont plus fiables.

Ces résultats confirment l'intérêt de la méthode globale avec calcul des mouvements.

Lorsque la focale est courte, les paramètres obtenus par les deux méthodes ne diffèrent pas beaucoup, et semblent très précis pour les facteurs d'échelle. Nous remarquons que les résultats se dégradent assez nettement avec l'allongement de la focale. Ceci s'explique par la difficulté de faire des mouvements suffisamment grands, tout en maintenant un champ de vue commun entre deux vues, qui est de toutes manière plus restreint, et donne lieu donc à moins d'appariements. De plus, la distance de la caméra à l'objet le plus proche doit augmenter. Bien que l'autocalibration semble surestimer la focale, elle donne néanmoins des résultats qui sont cohérents, en particulier pour le facteur d'échelle, qui est obtenu par la méthode globale avec une variation de 1% lorsque le point principal est pas fixé, et 3% dans le cas contraire.

9.4.2 Reconstructions métriques

Tous les exemples de cette section ont été obtenus avec le programme de Luc Robert [132], qui est un algorithme de stéréovision trinoculaire permettant de reconstruire très efficacement

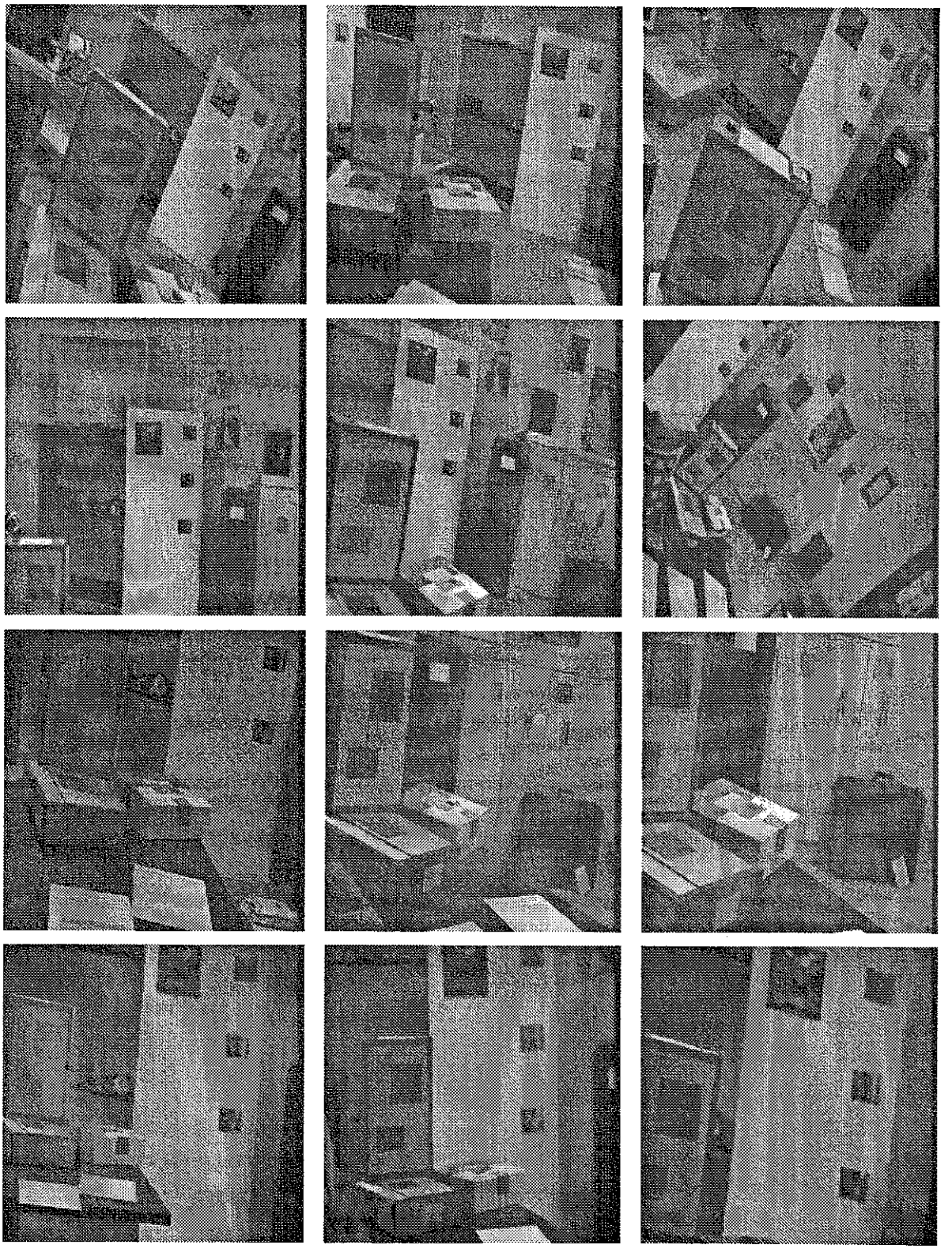


Fig. 9.12: Les images de scène d'intérieur, prises à quatre focales différentes

focale	méthode	α_u	α_v	u_0	v_0	$\theta - \frac{\pi}{2}$	$\frac{\alpha_u}{\alpha_v}$
9	MIRE	481.31	711.54	248.57	260.97	10^{-7}	.6764
	KRUPPA	495.95	725.65	263.04	286.34	0.02	.6834
		488.70	711.39	251.42	254.01		.6869
		472.83	669.47	255.00	255.00		.7062
	MOUV	503.49	760.71	250.24	282.67		.6618
		502.731	756.806	255	255		.6642
12	MIRE	642.45	950.37	248.30	263.31	-5.10^{-7}	.6759
	KRUPPA	648.35	891.50	335.33	53.11	0.35	.7272
		549.77	801.34	182.70	418.16		.6860
		672.19	900.00	255.00	255.00		.7468
	MOUV	636.12	921.36	201.52	338.89		.6904
		682.01	1040.75	255	255		.6553
20	KRUPPA	885.95	1610.78	614.74	-362.90	0.03	.5500
		9.50	20.34	538.23	-1124.19		.4670
		1233.45	1736.28	255.00	255.00		.7103
	MOUV	1208.83	1838.48	251.93	200.58		.6575
		1238.08	1869.84	255	255		.6621
30	MIRE	1573.20	2330.953	207.98	210.35	4.10^{-7}	.6749
	KRUPPA	411.80	1846.64	1651.84	758.30	-0.87	.2229
		452.46	2267.96	1214.65	40.03		.1995
		1242.38	2046.59	255.00	255.00		.6070
	MOUV	2047.61	3063.94	249.678	198.463		.6682
		2024.67	3028.94	255	255		.6684

Tableau 9.4: Paramètres intrinsèques obtenus sur la scène d'intérieur

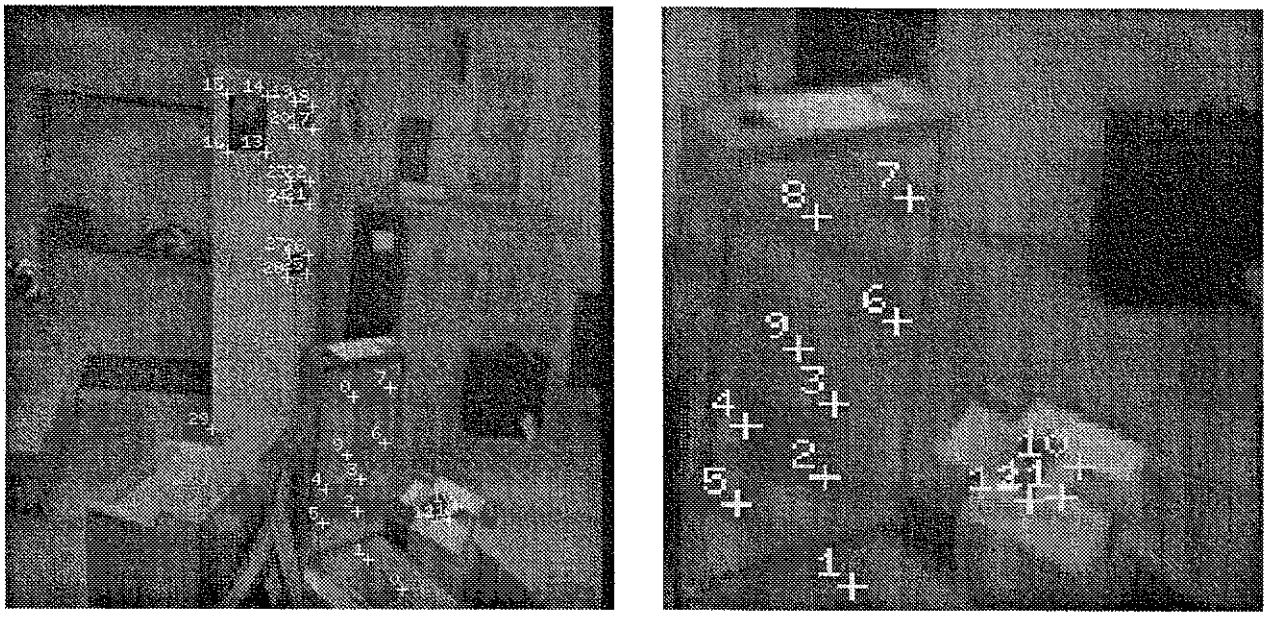


Fig. 9.13: Un exemple de faux appariement détecté

des courbes, même avec les images que nous lui avons fournies, qui ne lui sont pourtant pas très favorables: les déplacements sont grands, et comportent une composante rotationnelle non négligeable, deux conditions nécessaires pour obtenir des résultats d'autocalibration de qualité. Une des particularités de cet algorithme est d'être entièrement formulé en termes de matrices fondamentales, pour la phase primordiale de mise en correspondance. L'ensemble des travaux de cet auteur constituent d'ailleurs la meilleure illustration, originale et importante, de la richesse de la notion de matrice fondamentale, puisqu'il montre que de nombreuses propriétés géométriques peuvent être prédites, d'une vue à l'autre, grâce à celle-ci.

Le succès de l'algorithme à apparier des primitives est une confirmation de la qualité des matrices fondamentales que nous lui avons fournies. Pour la phase de reconstruction, l'information métrique est nécessaire. Nous avons obtenu les trois matrices de projection par la méthode décrite au 9.3.3. Les paramètres intrinsèques ont été déterminés simultanément, en utilisant la méthode décrite section 9.3.3.

La mire de calibration Nous avons d'abord utilisé les images de mire 8.6, pour lesquels les résultats de l'autocalibration étaient bons, malgré le décentrage du point principal. Pour illustrer la qualité des *matrices de projection* obtenues, nous montrons d'abord figure 9.14 des droites épipolaires obtenues avec celles-ci sous un fort agrandissement.

Nous présentons ensuite, figure 9.15 plusieurs résultats de reconstruction, dans lesquels nous montrons la mire reconstruite, sous différents angles. La qualité de la reconstruction est assez comparable à celles obtenues avec les méthodes de calibration classiques, comme en témoigne l'angle entre les deux plans, très proche de 90°, et le fait que ceux-ci soient

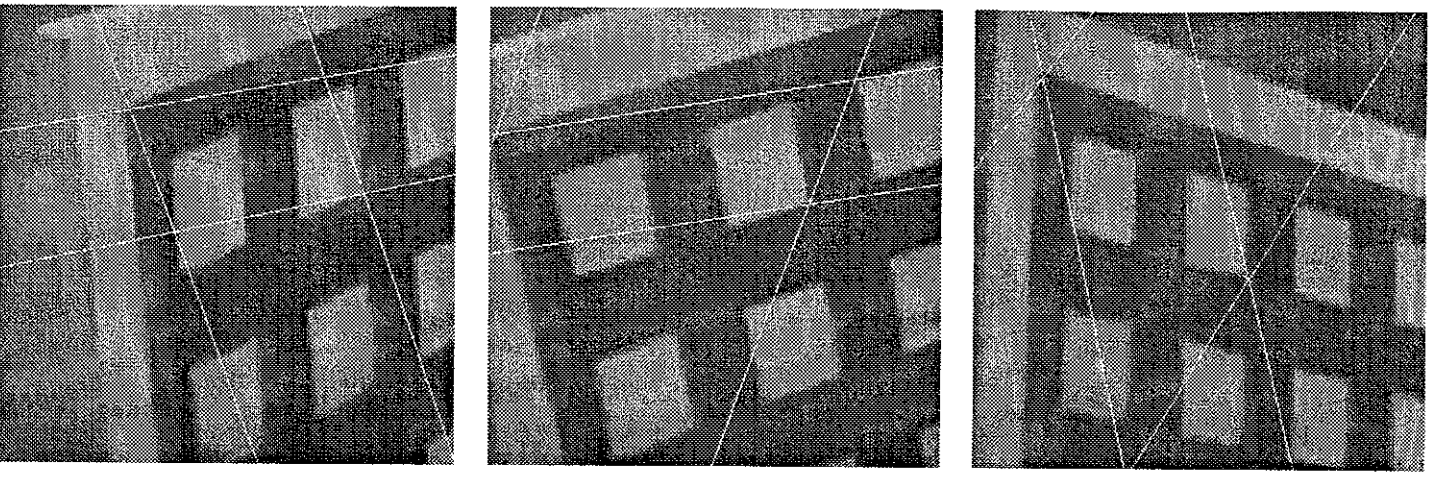


Fig. 9.14: Les images de mire avec droites épipolaires obtenues à partir des matrices de projection obtenues après autocalibration

contiennent très exactement les segments.

Une scène d'intérieur Nous avons ensuite utilisé la même méthode en l'appliquant à un cas plus difficile, celui de la scène ordinaire, qui a servi précédemment. Les résultats présentés sont ceux obtenus à la focale douze. Nous avons également des bonnes reconstructions à la focale neuf, mais il n'a pas été possible d'en effectuer aux longues focales, vingt et trente, plus parce que les images se prêtent alors mal à la stéréovision trinoculaire qu'à cause de l'imprécision de l'autocalibration. Un exemple de triplet utilisé est montré figure 9.16. Nous pouvons voir que les chaînes de pixels pouvant potentiellement être mises en correspondance ne sont pas très nombreuses. Nous pouvons constater sur les figures 9.17 et 9.18 que presque toutes ces chaînes ont été appariées, ce qui témoigne de la qualité de la géométrie épipolaire calculée. Ces figures sont des stéréogrammes pour vue croisée, qui permettent également d'apprécier la reconstruction. Les plans sont bien définis, et les angles respectés. Nous montrons également figure 9.19 deux vues reprojetées, l'une de côté, l'autre de haut, qui montrent que l'échelle des distances est correcte.

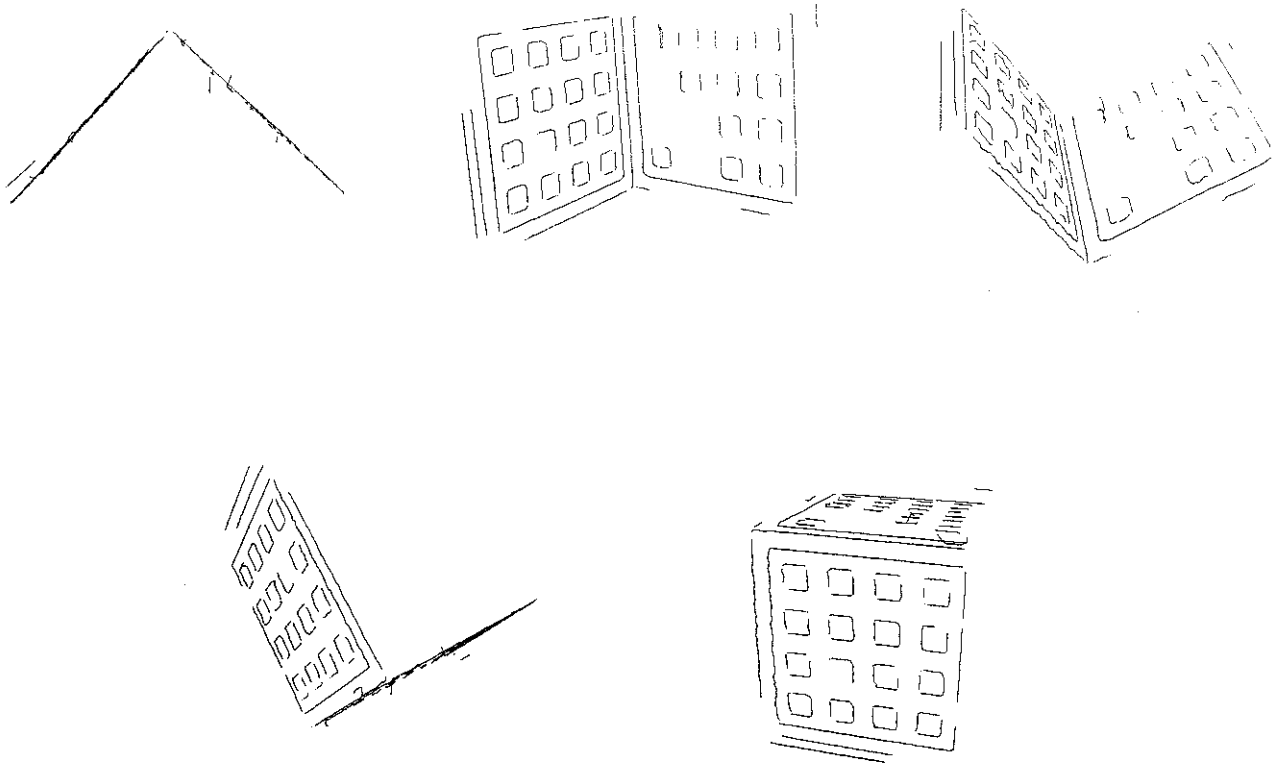


Fig. 9.15: Reconstructions 3D de la mire

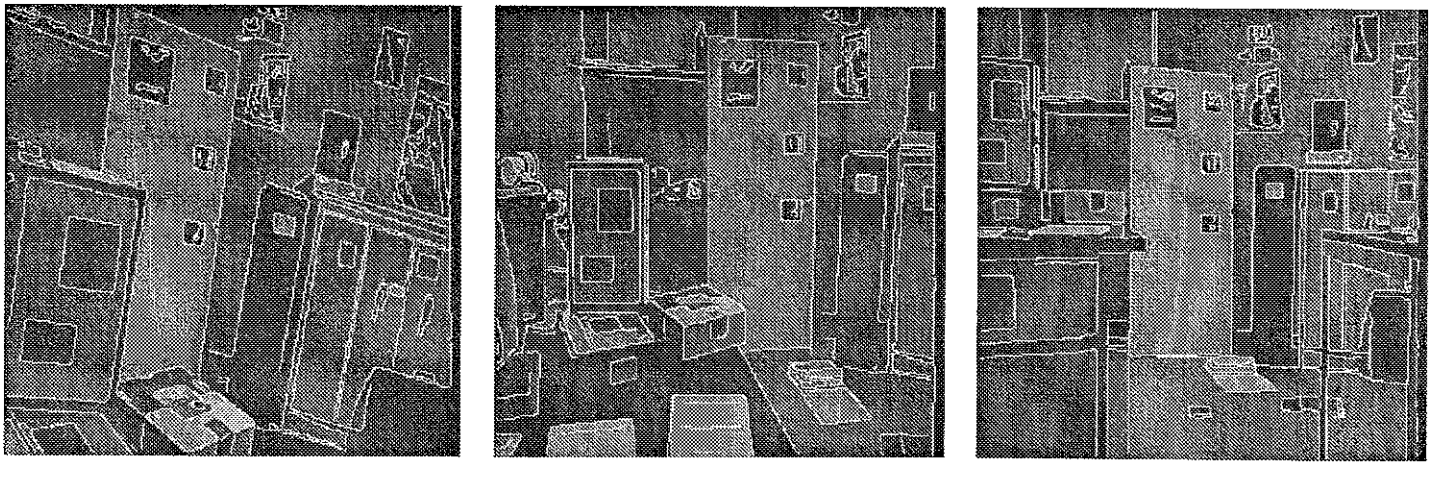


Fig. 9.16: Le triplet de vues ayant servi à reconstruire la scène d'intérieur

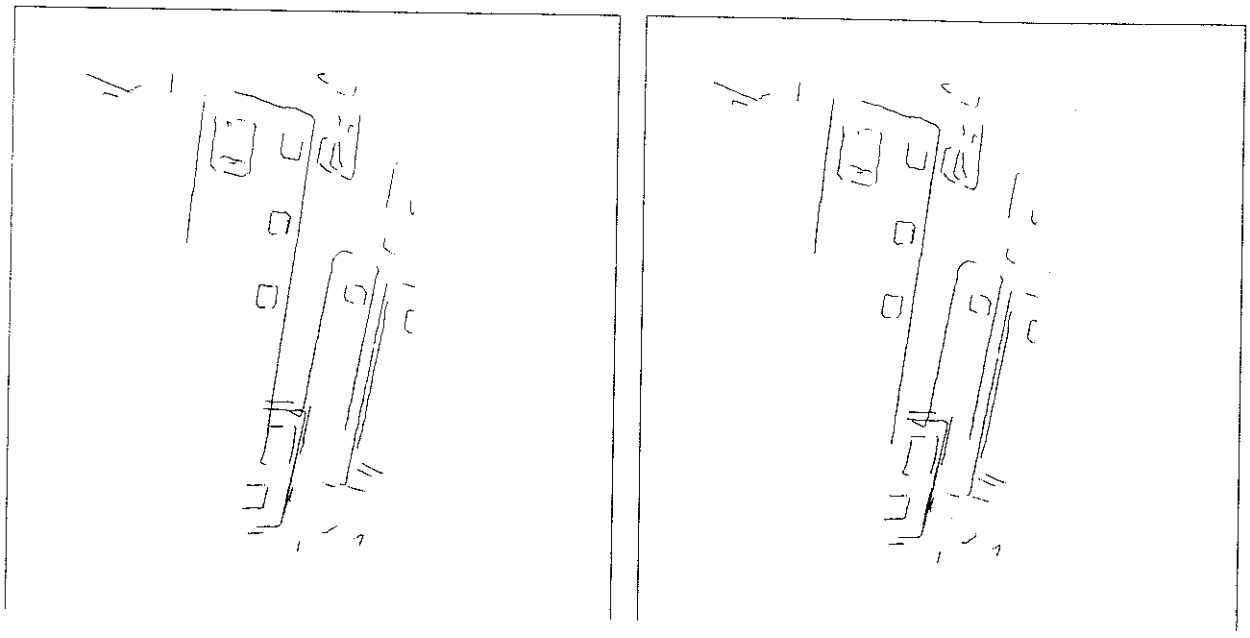


Fig. 9.17: Reconstruction de la scène d'intérieur (stéréogramme pour vue croisée)

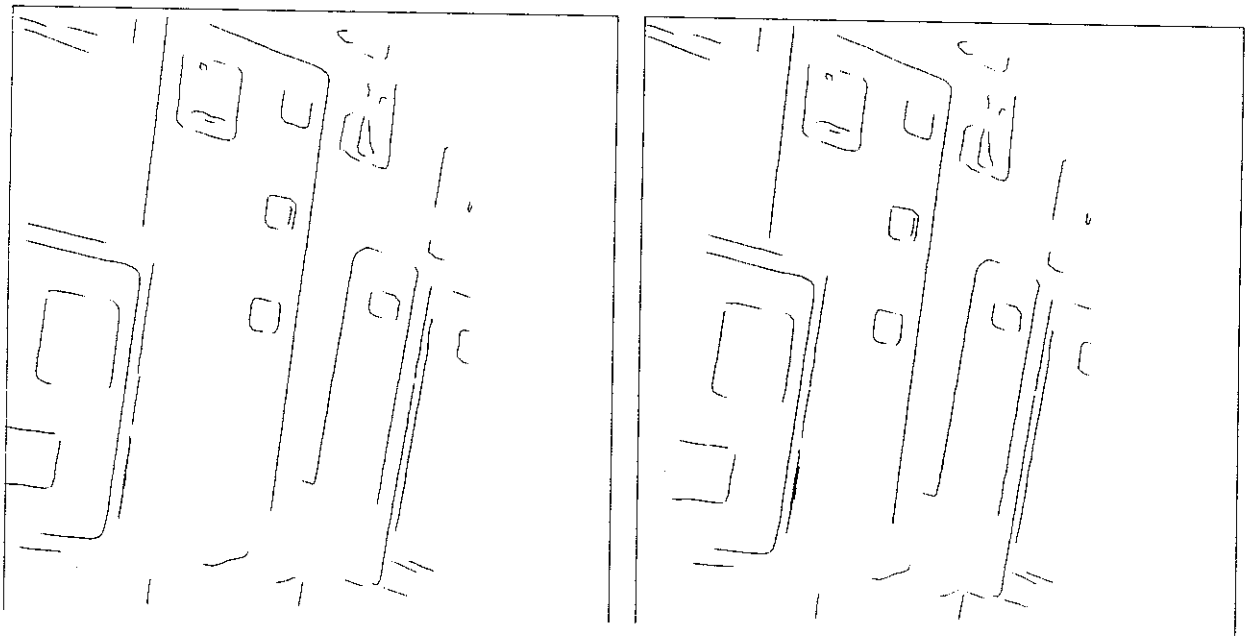


Fig. 9.18: Reconstruction de la scène d'intérieur (stéréogramme pour vue croisée)

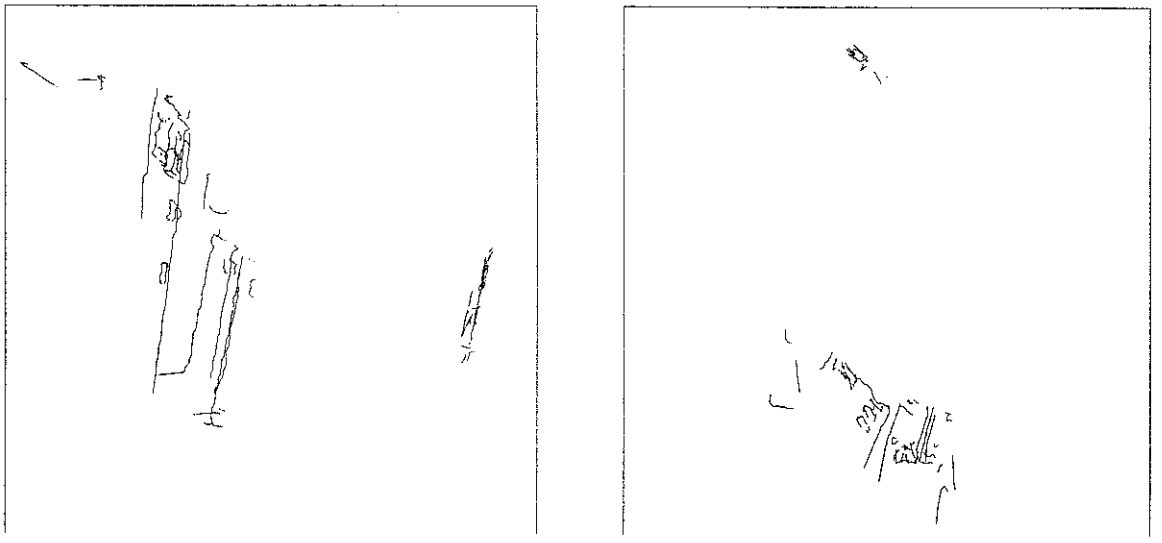


Fig. 9.19: Deux vues reprojetées de la scène d'intérieur

Liste des figures

1.1 Le modèle de la caméra . . . 12
1.2 Les images de mire ayant servi à obtenir des paramètres intrinsèques . . . 21
1.3 Le principe de l'autocalibration . . . 25

2.1 La contrainte épipolaire . . . 34
2.2 La contrainte de Longuet-Higgins . . . 38
2.3 La transformation épipolaire . . . 43
2.4 Statistique: erreur relative sur les coefficients de l'homographie épipolaire . . . 44
2.5 Chaque paire de droites épipolaires partitionne les plans rétiniens en deux demi-plans qui se correspondent . . . 47
2.6 Configurations d'épipolaires permises et configurations interdites . . . 47
2.7 Zones permises pour les paires d'épipoles . . . 49
2.8 Relations d'ordre sur les rétines induites par l'ordre sur les plans épipolaires . . . 49

3.1 La méthode de Sturm: un exemple montrant les cubiques et points remarquables . . . 56
3.2 Champ des déplacements avec les mouvements 1 (gauche) et 2 (droite) . . . 57
3.3 Superposition de cubiques obtenues à différents niveaux de bruit avec le mouvement 1 58
3.4 Superposition de cubiques obtenues à différents niveaux de bruit avec le mouvement 2 58
3.5 Quatre cubiques obtenues avec le mouvement 1, à 0.1, 0.5 et 1 pixel de bruit . . . 70
3.6 Cinq cubiques obtenues avec le mouvement 2, à 0.1, (0.5) et 1 pixel de bruit . . . 71
3.7 Un exemple de cellules trouvées par la méthode des cubiques . . . 74

4.1 Distances des droites épipolaires à l'épipole, dans le cas du critère linéaire . . . 89
4.2 Distances des droites épipolaires à l'épipole, dans le cas du critère linéaire (coupe) . 89
4.3 Distances relatives obtenues en initialisant avec les valeurs exactes . . . 103
4.4 Distances relatives obtenues en initialisant avec le critère linéaire . . . 104
4.5 Distances relatives entre les résultats obtenus avec les deux initialisations . . . 104
4.6 Nombre de faux minima . . . 105

4.7 Epipolaires obtenues avec le critère linéaire (haut), et avec le critère non-linéaire (bas) 107
4.8 Détail montrant l'ajustement avec le critère linéaire (gauche) et le critère non-linéaire (droite) . . . 108
4.9 Intersection des droites épipolaires obtenues avec le critère linéaire . . . 109
4.10 Epipolaires supplémentaires obtenues avec le critère linéaire (haut), et avec le critère non-linéaire (bas) . . . 110
4.11 Un exemple de géométrie épipolaire incohérente obtenue par calcul indépendant sur chaque image . . . 111
4.12 Un exemple de calcul sur des scènes d'extérieur. Linéaire à gauche, non-linéaire (2 faisceaux) à droite . . . 111
4.13 Plusieurs géométries plausibles: Epipolaires estimées (haut) et épipolaires correctes (bas) . . . 112

5.1 Distances relatives des matrices d'homographies obtenues avec différentes méthodes de calcul . . . 116
5.2 Appariements réels (haut), et prédits en utilisant une homographie (bas) . . . 122
5.3 Ambiguïté: quatre exemples de faisceaux d'épipolaires cohérents dans la première image . . . 123
5.4 Droite épipolaire obtenue en observant deux points particuliers . . . 126
5.5 Ensemble d'épipolaires obtenues par homographie . . . 129
5.6 Les quatre images de mire utilisées à titre d'exemple . . . 130

6.1 Ellipse d'incertitude et épipoles bruités (critère linéaire), pour le mouvement 1 . . . 137
6.2 Ellipse d'incertitude et épipoles bruités (critère linéaire), pour le mouvement 2 . . . 138
6.3 Ellipse d'incertitude et épipoles bruités (critère non-linéaire), pour le mouvement 1 . . . 138
6.4 Ellipse d'incertitude et épipoles bruités (critère non-linéaire), pour le mouvement 2 . . . 139
6.5 Corrélation entre covariances calculées et covariances prédites . . . 140
6.6 Construction des surfaces critiques . . . 142
6.7 Exemple: effet de la variation de la norme de la translation sur la stabilité . . . 152
6.8 Statistiques: effet de la variation de la direction de la translation sur la stabilité . . . 153
6.9 Statistiques: effet de la variation de la direction de la translation sur la stabilité . . . 154
6.10 Statistiques: effet de la variation de la norme de la rotation sur la stabilité . . . 156
6.11 La surface critique utilisée dans les expériences (stéréogramme pour vue croisée) . . . 158
6.12 Une caractérisation expérimentale globale de l'instabilité . . . 162

7.1 La conique absolue et la transformation épipolaire . . . 170

8.1 Construction de la courbe duale . . . 185
8.2 Le principe de la méthode des continuations . . . 187
8.3 Un exemple de calcul de paramètres intrinsèques par filtrage de Kalman . . . 196
8.4 Statistiques de convergence des paramètres intrinsèques avec deux types d'initialisation 198
8.5 Statistiques de convergence des paramètres intrinsèques avec itérations . . . 199
8.6 Les images de mire . . . 203
8.7 Les images de fragments de mire . . . 205

9.1 Erreur relative sur la rotation obtenue en initialisant avec les valeurs exactes . . 212
9.2 Erreur relative sur la direction de translation obtenue en initialisant avec les valeurs exactes 212
9.3 Erreur relative sur la rotation obtenue avec deux initialisations différentes . . . 213
9.4 Erreur relative sur la direction de translation obtenue avec deux initialisations différentes 213
9.5 Sensibilité du calcul du mouvement aux erreurs sur le point principal 215
9.6 Sensibilité du calcul du mouvement aux erreurs sur le facteur d'échelle 216
9.7 Sensibilité du calcul du mouvement aux erreurs sur les facteurs d'échelle, lors que ceux-ci sont ajustés dans la minimisation 219
9.8 Précision d'ajustement des facteurs d'échelle lors du calcul du mouvement . . . 219
9.9 Comparaison des équations de Kruppa et de la méthode avec calcul du mouvement. Gauche 4 paramètres. Droite: 2 paramètres. 222
9.10 Les paramètres intrinsèques d'un système de stéréovision 224
9.11 Déplacement d'un système de stéréovision 225
9.12 Les images de scène d'intérieur, prises à quatre focales différentes 231
9.13 Un exemple de faux appariement détecté 233
9.14 Les images de mire avec droites épipolaires obtenues à partir des matrices de projection obtenues après autocalibration 234
9.15 Reconstructions 3D de la mire 235
9.16 Le triplet de vues ayant servi à reconstruire la scène d'intérieur 235
9.17 Reconstruction de la scène d'intérieur (stéréogramme pour vue croisée) 236
9.18 Reconstruction de la scène d'intérieur (stéréogramme pour vue croisée) 236
9.19 Deux vues reprojetées de la scène d'intérieur 237

Liste des tableaux

1.1 Exemple de paramètres intrinsèques obtenus à partir de différentes positions de la caméra . . . 21
1.2 Exemple de paramètres intrinsèques obtenus à partir de différentes focales d'une caméra munie d'un zoom . . . 22

3.1 Les deux mouvements utilisés pour illustrer la méthode de Sturm . . . 56
3.2 Résultats de l'algorithme de Sturm avec choix exhaustif, pour le mouvement 1 . . . 61
3.3 Résultats de l'algorithme de Sturm avec choix exhaustif, pour le mouvement 2 . . . 62
3.4 Résultats de l'algorithme de Sturm avec consensus aléatoire, pour le mouvement 1 . . . 65
3.5 Résultats de l'algorithme de Sturm avec consensus aléatoire, pour le mouvement 2 . . . 65
3.6 Résultats de l'algorithme algébrique d'intersection, pour le mouvement 1 . . . 67
3.7 Résultats de l'algorithme algébrique d'intersection, pour le mouvement 2 . . . 68
3.8 Résultats de l'algorithme graphique d'intersection, pour le mouvement 1 . . . 73
3.9 Résultats de l'algorithme graphique d'intersection, pour le mouvement 2 . . . 73
3.10 Moyenne des distances relatives obtenus par la méthode des birapports . . . 80
3.11 Minimum exact du critère des birapports, pour le mouvement 1 . . . 81
3.12 Minimum exact du critère des birapports pour le mouvement 2 . . . 81

4.1 Un exemple illustrant le comportement différent d'implémentations du critère linéaire . . . 86
4.2 Un exemple illustrant le comportement du critère linéaire lorsque le mouvement est une translation . . . 90
4.3 Comparaison des quatre coordonnées des épipoles . . . 100
4.4 Comparaison des critères linéaires . . . 101
4.5 Critères utilisées pour le calcul de la matrice fondamentale . . . 101
4.6 Comparaison des résultats de deux algorithmes de minimisation . . . 102
4.7 L'apport de la procédure itérative du M-estimateur . . . 105

5.1 Différences réciproques entre les matrices d'homographies calculées par différentes méthodes . . . 117
5.2 Résultats d'estimation de la matrice fondamentale lorsque les points sont situés dans des plans . . . 120
5.3 Un exemple illustrant les résultats obtenus avec différentes méthodes de calcul de $\mathbf{F}$ à partir de matrices $\mathbf{H}_i$ . . . 125
5.4 Résultats de détermination de l'épipole en utilisant un seul plan et des points extérieurs 127
5.5 Résultats obtenus sur les images de mire avec le critère général et celui fondé sur les plans . . . 128
5.6 Comparaison statistique de la méthode des plans et de la méthode générale . . . 130

6.1 Résultat de détermination de transformation quadratique: le cas exact . . . 149
6.2 Exemple: influence de l'angle de la rotation sur la stabilité du critère linéaire . . . 155
6.3 Influence de la distance à la surface critique et du bruit sur la stabilité . . . 159
6.4 Influence de la distance à la surface critique sur la stabilité . . . 160
6.5 Facteurs d'instabilité dans une expérience statistique . . . 160

8.1 Résultats de la méthode des continuations, configuration 1 . . . 190
8.2 Résultats de la méthode des continuations, configuration 2 . . . 191
8.3 Résultats statistiques de la méthode de minimisation globale avec 3 mouvements . . . 193
8.4 L'apport de connaissances a priori dans la détermination des paramètres intrinsèques 194
8.5 Comparaison de la minimisation et du filtre de Kalman . . . 201
8.6 Effet de l'application du filtre de Kalman . . . 201
8.7 Paramètres intrinsèques calculés sur les fragments de mire . . . 206

9.1 Exemples numériques: effet de l'imprécision des paramètres intrinsèques sur le calcul du mouvement . . . 214
9.2 Exemples numériques: différentes méthodes de calcul des paramètres intrinsèques . . . 221
9.3 Exemples numériques: évolution des erreurs sur le mouvement . . . 221
9.4 Paramètres intrinsèques obtenus sur la scène d'intérieur . . . 232

Bibliographie

[1] *NAg Fortran library, mark 14*. The Numerical Algorithms Group Limited, 1990.

[2] Y.I. Abdel-Aziz and H.M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. In *Symposium on close-range photogrammetry*, pages 1–18, UI-Urbana, 1971.

[3] G. Adiv. Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:477–489, 1989.

[4] J.K. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images. A review. *Proc. of the IEEE*, 76(8):917–934, Aug 1988.

[5] H. Asada and M. Brady. The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:2–14, 1986.

[6] N. Ayache. *Vision stereoscopique et perception multisensorielle*. InterEditions, 1989.

[7] A. Bani-Hashemi. Finding the aspect-ratio of an imaging system. In *Proc. of the conf. on Computer Vision and Pattern Recognition*, pages 122–126, 1991.

[8] E.B. Barrett, M.H. Brill, N.N. Haag, and P. M. Payton. Invariant linear methods in photogrammetry and model-matching. In J.L. Mundy and A. Zisserman, editors, *Geometric invariance in computer vision*, chapter 14, pages 277–292. MIT Press, 1992.

[9] P. Beardsley, D. Murray, and A. Zissermann. Camera calibration using multiple images. In *Proceedings 2nd ECCV*, pages 312–320, 1992.

[10] P.R. Beaudet. Rotational invariant image operators. In *Proc. International Conference on Pattern Recognition*, pages 579–583, 1978.

[11] H.A. Beyer. Linejitter and geometric calibration of CCD cameras. *ISPRS Journal of photometry and remote sensing*, 45:17–32, 1990.

[12] H.A. Beyer. Accurate calibration of CCD cameras. In *Proc. of the conf. on Computer Vision and Pattern Recognition*, pages 96–101, 1992.

[13] T. Blaszka. Modélisation et caractérisation des variations d'intensité dans une image. Rapport de DEA. Université Nice Sophia-Antipolis, Septembre 1992.

[14] F. L. Bookstein. Fitting conic sections to scattered data. *Computer Graphics and Image Processing*, 9(1):56–71, Jan 1979.

[15] D.C. Brown. Decentering distorsion of lenses. *Photogrammetric Engineering*, 32(3):444–462, 196.

[16] B. Caprile and V. Torre. Using Vanishing Points for Camera Calibration. *International Journal of Computer Vision*, 4:127–140, 1990.

[17] M. Chasles. Question No. 296. *Nouv. Ann. Math.*, 14:50, 1855.

[18] F. Chaumette and P. Rives. Modélisation et Calibration d'une caméra. In *AFCET*, pages 527–536, 1989.

[19] H.H. Chen. A screw motion approach to uniqueness analysis of head-eye geometry. In *Proc. of the conf. on Computer Vision and Pattern Recognition*, pages 145–151, 1991.

[20] J.C.K. Chou and M. Kamel. Quaternions approach to solve the kinematic equation of rotation, $A_a A_x = A_x A_a$, of a sensor-mounted robotic manipulator. In *Proc International Conference on Robotics and Automation*, pages 656–662, 1988.

[21] C. Coelho, A. Heller, J. Mundy, D. Forsyth, and Zisserman A. An experimental evaluation of projective invariants. In J. Mundy and Zisserman A., editors, *Geometric invariance in computer vision*, chapter 4, pages 87–104. MIT Press, 1992.

[22] R.R. Cohen and E.A. Feigenbaum. *The Handbook of Artificial Intelligence*, volume III. William Kaufmann, 1982.

[23] H.S.M. Coxeter. *Projective Geometry*. Springer Verlag, second edition, 1987.

[24] J. Dahler. Problems with Digital Image Acquistion with CCD Cameras. In *Intercommission Conference on Fast Processing of Photogrammetric Data*, 1987.

[25] K. Daniilidis and H.-H. Nagel. Analytical results on error sensitivity of motion estimation from two views. In *Proc. European Conference on Computer Vision*, pages 199–208, Antibes, April 1990.

[26] N. Daucher, M. Dhome, and J.T. Lapreste. Etalonage de caméra à partir d'images de sphères. In *Reunion Orasis*, ENST Paris, Oct 1992. aussi soumis a *Traitement du signal*.

[27] R. Deriche. Fast algorithms for low-level vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1), 1990.

[28] R. Deriche and T. Blaszka. Recovering and characterizing image features using an efficient model based approach. In *Proc. International Conference on Computer Vision and Pattern Recognition*, 1993. Submitted.

[29] R. Deriche and O.D. Faugeras. Tracking line segments *Image and Vision Computing*, 8(4):261–270, November 1990. A shorter version appeared in the Proceedings of the 1st ECCV.

[30] R. Deriche and G. Giraudon. Accurate corner detection : An analytical study. In *Proc. International Conference on Computer Vision*, pages 66–70, Osaka, Japan, December 1990.

[31] R. Deriche and G. Giraudon. A computational approach for corner and vertex detection. *The International Journal of Computer Vision*, 1993 Accepted for Publication.

[32] L. Dreschler and H.H. Nagel On the selection of critical points and local curvature extrema of region boundaries for interframe matching. In *Proc. International Conference on Pattern Recognition*, pages 542–544, 1982.

[33] R. Dutta, R. Manmatha, E. Riseman, and M. Snyder. Issues in extracting motion parameters and depth from approximate translational motion. In *Proc. DARPA Image Understanding workshop*, pages 945–960, April 1988.

[34] T. Echigo. A Camera Calibration Technique using Three Sets of Parallel Lines. *Machine Vision and Applications*, 3:159–167, 1990.

[35] W. Faig Calibration of close range photogrammetry systems: mathematical formulation. *Photogrammetric engineering and remote sensing*, 41(12):1479–1486, 1975.

[36] J.-Q. Fang and T.S. Huang. Solving three-dimensionnal small rotation motion equations: uniqueness, algorithms and numerical results. *Computer Vision, Graphics, and Image Processing*, 16:183–206, 1984.

[37] J.Q. Fang and T.S. Huang. Some experiments on estimating the 3D motion parameters of a rigid body from two consecutive image frames. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:545–554, 1984.

[38] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig In *Proceedings of the 2nd European Conference on Computer Vision*, pages 563–578, 1992.

[39] O.D. Faugeras. *Three-dimensional computer vision: a geometric viewpoint.* MIT Press, 1993. To appear.

[40] O.D. Faugeras, Q.-T. Luong, and S.J. Maybank. Camera self-calibration: theory and experiments In *Proc. 2nd ECCV*, pages 321–334, 1992.

[41] O.D. Faugeras and F. Lustman. Motion and Structure from Motion in a piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(3):485–508, 1988.

[42] O.D. Faugeras and S.J. Maybank. Motion from point matches: multiplicity of solutions *The International Journal of Computer Vision*, 4(3):225–246, 1990. also INRIA Tech. Report 1157.

[43] O.D. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings of CVPR'86*, pages 15–20, 1986.

[44] Olivier D. Faugeras, Francis Lustman, and Giorgio Toscani. Motion and Structure from point and line matches. In *Proc International Conference on Computer Vision*, pages 25–34, June 1987.

[45] M A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. Also in Readings in Computer vision, ed. Fischler.

[46] W. Forstner and A. Pertl. Photogrammetric standard methods and digital image matching techniques for high precision surface measurements. In Gelsema, E.S. and Kanal, L.N., editor, *Pattern Recognition in Practice II*, pages 57–72. Elsevier Science Publishers, 1986.

[47] C.S. Fraser. Photogrammetric camera component calibration: a review of analutical techniques. Workshop on calibration and orientation of cameras in computer vision. XVII Congress ISPRS, Aug 1992.

[48] S. Ganapathy Decomposition of transformation matrices for robot vision In *Proc. of IEEE conference on robotics and automation*, pages 130–139, 1984.

[49] L.E. Garner. *An outline of projective geometry*. Elsevier North Holland, 1981

[50] D B. Gennery. Stereo-camera calibration. In *Proc. Image understanding workshop*, pages 101–108, 1979.

[51] D.B. Gennery. *Modelling the Environment of an Exploring Vehicle by means of Stereo Vision*. PhD thesis, Stanford University, June 1980.

[52] G. Giraudon and R. Deriche. On corner and vertex detection. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 650–655, Lahaina,Maui,Hawai, June 1991.

[53] G.H. Golub and C.F. Van Loan. *Matrix computations*. The John Hopkins University Press, 1989.

[54] K. D. Gremban, C. E. Thorpe, and T. Kanade. Geometric Camera Calibration using Systems of Linear Equations. In *Image Understanding Workshop*, pages 820–825, 1988.

[55] W.I. Grotsky and L.A. Tamburino. A unified approach to the linear camera calibration problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:663–671, 1990.

[56] A. Guiducci. Corner characterization by differential geometry techniques. *Pattern Recognition Letters*, 8:311–318, 1988.

[57] C. Harris. Camera calibration. Workshop VOILA on calibration, 1990.

[58] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conference*, pages 189–192, 1988.

[59] C.G. Harris:. Determination of ego-motion from matched points. In *Proc. Alvey Vision Conference*, 1987.

[60] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. of the conf. on Computer Vision and Pattern Recognition*, pages 761–764, Urbana, 1992.

[61] R.I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *Proceedings 2nd ECCV*, pages 579–587, 1992.

[62] O. Hesse. Die cubische Gleichung, von welcher die Lösung des Pro blems der Homogıaphie von M. Chasles abhängt. *J. reine angew. Math.*, 62:188–192, 1863.

[63] B.K.P. Horn. Motion fields are hardly ever ambiguous. *The International Journal of Computer Vision*, 1(3):263–278, 1987.

[64] B.K.P. Horn. Relative orientation. *The International Journal of Computer Vision*, 4(1):59–78, Jan. 1990.

[65] B.K.P. Horn and E.J. Weldon. Direct methods for recovering motion. *The International Journal of Computer Vision*, 2(1):51–76, 1988.

[66] B.P.K. Horn. *Robot vision*. MIT Press, 1986.

[67] T.S. Huang and A.N. Netravali. Motion and structure from features correspondances: a review. *Proc. of the IEEE*, 1992. To appear.

[68] P.J. Huber. *Robust Statistics*. John Wiley, 1981.

[69] Y.P. Hung and S.W. Shieh. When should we consider lens distortion in camera calibration ? In *Proc. IAPR worshop on machine vision applications*, pages 367–370, Tokyo, November 1990.

[70] Koenderink J. J. and A. J. van Doorn. Affine Structure from Motion. *Journal of the Optical Society of America A*, 8(2):377–385, 1992.

[71] A.M. Jazwinsky. *Stochastic processes and filtering theory*. Academic Press, London, 1970.

[72] C.P. Jerian and R. Jain. Structure from motion. a critical analysis of methods. *IEEE Transactions on Systems, Man and Cybernetics*, 21(3):572–587, 1991.

[73] J.-L. Jezouin. *Vision dynamique monoculaire: reconstruction 3D et interpretation*. PhD thesis, Universite Paris-Sud. Centre d'Orsay, 1991.

[74] K. Kanatani. Computational projective geometry. *Computer Vision, Graphics, and Image Processing. Image Understanding*, 54(3), 1991.

[75] K. Kanatani. *Geometric computation for machine vision*. Oxford university press, 1992. to appear.

[76] K. Kanatani and Y. Onodera. Noise robust camera calibration using vanishing points. *IEICE transactions on informations and systems*, e74(10), October 1991.

[77] R.E. Kelly, P.R.H. McConnell, and S.J. Mildenberger. The gestalt photomapper. *Photogrammetric Engineering and Remote Sensing*, 43:1407–1417, 1977.

[78] L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, pages 95–102, 1982.

[79] E.P. Krotkov. *Active computer vision by cooperative focus and stereo*. Springer-Verlag, 1989.

[80] E. Kruppa. Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orientierung. *Sitz.-Ber. Akad. Wiss., Wien, math. naturw. Kl., Abt. IIa.*, 122:1939–1948, 1913.

[81] R. Kumar and A. Hanson. Sensibility of the pose refinement problem to accurate estimation of camera parameters. In *Proceedings of the International Conference on Computer Vision*, pages 365–369, Osaka, Japan, 1990.

[82] R.V.R. Kumar, A. Tirumalai, and R.C. Jain. A non-linear optimization algorithm for the estimation of structure and motion parameters. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 136–143, 1989.

[83] J.M. Lavest. *Stereovision axiale par zoom pour la robotique*. PhD thesis, Universite Blaise Pascal de Clermont-Ferrand, 1992.

[84] J.M. Lavest, G. Rives, and M. Dhome. Using a zoom in monocular vision for 3D reconstruction. In *Proc. IARP-91*, Oxford, Sept. 1991.

[85] J.M. Lavest, G. Rives, and M. Dhome. Utilisation d'un objectif a focale variable en vision monoculaire en vue de la reconstruction 3D. In *Proc. AFCET*, pages 293–301, 1991.

[86] C.H. Lee. Time-varying images: the effect of finite resolution on uniqueness. *Computer Vision, Graphics, and Image Processing. Image Understanding*, 54(3):325–332, 1991.

[87] R.K Lenz. Lens distorsion corrected CCD camera calibration with coplanar calibration points for real-time 3D measurements. In *Proc. ISPRS Intercomission conference on fast processing of photogrammetric data*, pages 60–67, Interlaken, June 87.

[88] R.K. Lenz and R.Y. Tsai. Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3D Machine Vision Metrology. In *IEEE International Conference on Robotics and Automation*, pages 68–75, Raleigh, 1987.

[89] C. Longuet-Higgins. The reconstruction of a scene from two projections: configurations that defeat the 8-point algorithm. In *Proc. 1st Conf. on Artificial intelligence applications*, pages 395–397, Denver, 1984.

[90] H.C. Longuet-Higgins. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, 293:133–135, 1981.

[91] H.C. Longuet-Higgins. Multiple interpretations of a pair of images of a surface. *Proc. of the Royal Society London A*, 418:1–15, 1988.

[92] Q.-T. Luong, R. Deriche, O.D. Faugeras, and T. Papadopoulo. On determining the Fundamental matrix: analysis of different methods and experimental results. In *ICCV*, 1993. Submitted.

[93] Q.-T. Luong and O.D. Faugeras. Active head movements help solve stereo correspondance. In *Proc. ECAI 92*, pages 800–802, 1992.

[94] Q.-T. Luong and O.D. Faugeras. Active stereo with head movement. In *Proc. 2nd ICIP*, pages 507–511, 1992.

[95] Q.-T. Luong and O.D. Faugeras. Self-calibration of a camera using multiples images. In *Proc. 11th ICPR*, pages 9–12, 1992.

[96] Q.-T. Luong and O.D. Faugeras. Self-calibration of a stereo rig. Poster presented at XVII ISPRS, 1992.

[97] Q.-T. Luong and O.D. Faugeras. Automatic calibration of a stereo rig. In A. Grun and T.S. Huang, editors, *Calibration and orientation of cameras in computer vision*. Springer-Verlag, 1993. To appear.

[98] Q.-T. Luong and O.D. Faugeras. Determining the Fundamental matrix with planes: unstability and new algorithms. In *CVPR*, 1993. Submitted.

[99] F. Lustman. *Vision stereoscopique et perception du mouvement en vision artificielle*. PhD thesis, Universite de Paris-Sud, Dec. 1987.

[100] F. Mangili. Calibration of a CCD camera on a hybrid coordinate measuring machine for industrial metrology. Workshop VOILA on calibration, 1990.

[101] S.J. Maybank. The angular velocity associated with the optical flow field arising from motion through a rigid environment. *Proc. of the Royal Society London A*, 401:317–326, 1985.

[102] S.J. Maybank. *A theoretical study of optical flow*. PhD thesis, Birkbeck College, University of London, 1987.

[103] S.J. Maybank. Filter based estimates of depth. In *Proc. British Machine Vision Conf.*, pages 349–354, London, 1990.

[104] S.J. Maybank. The projective geometry of ambiguous surfaces. *Proc. of the Royal Society London A*, 332:1–47, 1990.

[105] S.J. Maybank. Properties of essential matrices. *International journal of imaging systems and technology*, 2:380–384, 1990.

[106] S.J. Maybank. *Theory of reconstruction from inage motion*. Springer-Verlag, 1992.

[107] S.J. Maybank and O.D. Faugeras. A Theory of Self-Calibration of a Moving Camera. *The International Journal of Computer Vision*, 8(2):123–151, 1992.

[108] P.S. Maybeck. *Stochastic models, estimation and control*. Academic Press, London, 1979.

[109] G. Medioni and Y. Yasumuto. Corner detection and curve representation using cubic b-spline. In *Proc. International Conference on Robotics and Automation*, pages 764–769, 1986.

[110] P. Meer, D. Mintz, and A. Rosenfeld. Robust regression methods for computer vision: a review. *The International Journal of Computer Vision*, 6(1):59–70, 1991.

[111] A. Mitiche, X. Zhuang, and R. Haralick. Interpretation of optical flow by rotation decoupling. In *Proc. IEEE workshop on computer vision*, pages 195–200, Miami Beach, FL, 1987.

[112] R. Mohr and E. Arbogast. It can be done without camera calibration. *Pattern Recognition Letters*, 12:39–43, 1990.

[113] R. Mohr, L. Morin, and E. Grosso. Relative positionning with poorly calibrated cameras. In J.L. Mundy and A. Zisserman, editors, *Geometric invariance in computer vision*, chapter 22, pages 440–462. MIT Press, 1992.

[114] R. Mohr, L. Morin, C. Inglebert, and L. Quan. Geometric solutions to some 3D vision problems. In R. Storer J.L. Crowley, E. Granum, editor, *Integration and Control in Real Time Active Vision*, ESPRIT BRA Series. Springer Verlag, 1991.

[115] R. Mohr, L. Quan, F. Veillon, and B. Boufama. Relative 3d reconstruction using multiple uncalibrated images. Technical Report RT84-IMAG12, LIFIA, June 1992.

[116] F. Mokhtarian and A. Mackworth. Scale-based description and recognition of planar curves and 2d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):34–43, 1986.

[117] H. Moravec. *Robot rover visual navigation*. UMI Research Press, 1981.

[118] A. Morgan. *Solving polynomial systems using continuation for engineering and science problems*. Prentice-Hall, 1987.

[119] J. L. Mundy and A. Zisserman, editors. *Geometric invariance in computer vision*. MIT Press, 1992.

[120] H.-H. Nagel. Image sequences. Ten (octal) years. From phenomenology towards a theoretical foundation. In *Proc. International Conference on Pattern Recognition*, pages 1174–1185, Paris, France, 1986.

[121] H.H. Nagel. Constraints for the estimation of displacement vector fields from image sequences. In *Proc. International Joint Conference on Artificial Intelligence*, pages 156–160, 1983.

[122] V.S. Nalwa and E. Pauchon. Edgel aggregation and edge description. *Computer Vision, Graphics, and Image Processing*, 40(1):79–94, Oct. 1987.

[123] S. Negahdaripour. Critical surface pairs and triplets. *The International Journal of Computer Vision*, 3(4):293–312, Nov. 1989.

[124] T.C. Nguyen and T.S. Huang. Quantization errors in avial motion stereo on rectangular-tesselated image sensors. In *Proc. International Conference on Pattern Recognition*, pages 13–16, 1992.

[125] J.A. Noble. Finding corners. *Image and Vision Computing*, 6:121–128, May 1988.

[126] S.I. Olsen. Epipolar line estimation. In *Proc. European Conference on Computer Vision*, pages 307–311, 1992.

[127] B. Peuchot. Camera virtual equivalent model: 0.01 pixel detectors. In *14th Inter. Conf. of the IEEE Engineering in Medecine and Biology Society*, Rennes, Nov. 1992.

[128] J. Philip. Estimation of three-dimensional motion of rigid objects from noisy observations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(1):61–66, 1991.

[129] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical recipes in C*. Cambridge University Press, 1988.

[130] P. Puget and T. Skordas. An Optimal Solution for Mobile Camera Calibration. In *European Conference on Computer Vision*, April 1990.

[131] F. Veillon R. Horaud and T. Skordas. Finding geometric and relational structures in an image. In *Proc. European Conference on Computer Vision*, pages 374–384, 1990.

[132] L. Robert. *Reconstruction de courbes et de surfaces par vision stéréoscopique. Applications a la robotique mobile*. PhD thesis, Ecole Polytechnique, 1993.

[133] K. Rohr. Modelling and identification of characteristic intensity variations. *Image and Vision Computing*, 10(2):66–76, 1992.

[134] S. Rubino. Reconstruction interactive de points. Rapport de DEA. Université Nice Sophia-Antipolis, Septembre 1992.

[135] P.D. Sampson. Fitting conic sections to very scattered data: an iterative refinement of the Bookstein algorithm. *Computer Graphics and Image Processing*, 18(1):97–108, Jan. 1982.

[136] G.L. Scott. *Local and global interpretation of moving images*. Pitman, 1987.

[137] J.G. Semple and G.T. Kneebone. *Algebraic projective geometry*. Oxford science publication, 1952.

[138] M.A. Shah and R. Jain. Detecting time-varying corners. *Computer Vision, Graphics, and Image Processing*, 28:345–355, 1984.

[139] A. Shashua. Projective structure from two uncalibrated images: structure from motion and recognition. Technical Report A.I. Memo No. 1363, MIT, Sept 1992.

[140] Y.S. Shiu and S. Ahmad. Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX = XB$. *IEEE Transactions on robotics and automation*, 5(1):16–29, 1989.

[141] M.R. Shortis, A.W. Burner, W.L. Snow, and W.K. Goad. Calibration tests of industrial and scientific CCD cameras. In *Proc. 1st Australian Photogrammetry Conf.*, Sydney, Nov. 1991.

[142] D. Sinclair, A. Blake, S. Smith, and C. Rothwell. Planar region detection and motion recovery. In *Proc. British Machine Vision Conf.*, pages 59–68, 1992.

[143] I. Sobel. On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes. *Artificial Intelligence*, 5:184–198, 1974.

[144] M.E. Spetsakis. A linear algorithm for point and line-based structure from motion. *Computer Vision, Graphics, and Image Processing. Image Understanding*, 56(2):230–241, 1992.

[145] M.E. Spetsakis and J. Aloimonos. Optimal computing of structure from motion using point correspondances in two frames. In *Proc. International Conference on Computer Vision*, pages 449–453, 1988.

[146] A. Stein and M. Werman. Robust statistics in shape fitting. In *Proc. 8th Israeli symposium on Artif. Intell. and Comp. Vision*, Dec. 1991.

[147] Rudolf Sturm. Das Problem der Projektivität und seine Anwendung auf die Flächen zweiten Grades. *Math. Ann.*, 1:533–574, 1869.

[148] I. Sutherland. Three-dimensional data input by tablet. *Proc. of IEEE*, 62(4):453–461, 1974.

[149] K. Tarabanis and R.Y. Tsai. Computing viewpoints that satisfy optical constraints. In *Proc. of the conf. on Computer Vision and Pattern Recognition*, pages 152–158, 1991.

[150] K. Tarabanis, R.Y. Tsai, and D.S. Goodman. Modeling of a computer-controlled zoom lens. In *Proc. International Conference on Robotics and Automation*, pages 1545–1551, Nice, May 1992.

[151] N.A. Thacker. On-line calibration of a 4-dof robot head for stereo vision. In *British Machine Vision Association meeting on active vision*, London, 1992.

[152] G. Toscani. *Systèmes de Calibration et Perception du Mouvement en Vision Articielle*. PhD thesis, Université de Paris-Sud Centre d'Orsay, December 1987.

[153] G. Toscani and O.D. Faugeras. Structure and motion from two perspective views. In *Proc. International Conference on Robotics and Automation*, pages 221–227, Raleigh, 1987.

[154] G. Toscani, R. Vaillant, R. Deriche, and O.D. Faugeras. Stereo camera calibration using the environment. In *Proceedings of the 6th Scandinavian conference on image analysis*, pages 953–960, 1989.

[155] H.P. Trivedi. On the reconstruction of a scene from two unregistered images. In *Proc. AAAI*, pages 652–656, 1986.

[156] H.P. Trivedi. Estimation of stereo and motion parameters using a variational principle. *Image and Vision Computing*, 5:181–183, 1987.

[157] H.P. Trivedi. Can multiple views make up for lack of camera registration ? *Image and Vision Computing*, 6(1):29–32, 1988.

[158] H.P. Trivedi. Semi-analytic method for estimating stereo camera geometry from matched points. *Image and Vision Computing*, 9, 1991.

[159] R.Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In *Proceedings CVPR '86, Miami Beach, Florida*, pages 364–374. IEEE, June 1986.

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

[161] R.Y. Tsai. Synopsis of Recent Progress on Camera Calibration for 3D Machine Vision. In Oussama Khatib, John J. Craig, and Tomás Lozano-Pérez, editors, *The Robotics Review*, pages 147–159. MIT Press, 1989.

[162] R.Y. Tsai and T.S. Huang. Estimating Three-dimensional motion parameters of a rigid planar patch, II: singular value decomposition. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 30, 1982.

[163] R.Y. Tsai and T.S. Huang. Uniqueness and estimation of three-dimensional motion parameters of rigid objects wirth curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:13–27, 1984.

[164] R.Y. Tsai and R.K. Lenz. Real time versatile robotics hamd/eye calibration using 3D machine vision. In *Proc. International Conference on Robotics and Automation*, pages 554–561, 1988.

[165] R. Vaillant. *Geometrie différentielle et vision par ordinateur*. PhD thesis, Université de Paris-Sud, 1990.

[166] T. Viéville. Auto-calibration of visual sensor parameters on a robotic head. *Image and Vision Computing*, 1993. To appear.

[167] C.W. Wampler, A.P. Morgan, and A.J. Sommese. Numerical continuation methods for solving polynomial systems arising in kinematics. Technical Report GMR-6372, General Motors Research Labs, August 1988.

[168] L. Wang and W. Tsai. Computing cameras parameters using vanishing line information from a rectangular parallepiped. *Machine Vision and Applications*, 3:129–141, 1990.

[169] G.-Q. Wei and S.D. Ma. Two plane camera calibration: a unified model. In *Proc of the conf. on Computer Vision and Pattern Recognition*, pages 133–138, 1991.

[170] J. Weng, N. Ahuja, and T.S. Huang. Optimal motion and structure estimation. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 144–152, 1989.

[171] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distorton models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, 1992.

[172] J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from two perspective views: algorithms, error analysis and error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5):451–476, 1989.

[173] Y. Yakimovsky and R. Cunningham. A system for extracting three-dimensional measurements from a stereo pair of tv cameras. *Computer Graphics and Image Processing*, 7:195–210, 1978.

[174] B.L. Yen and T.S. Huang. Determining 3-D motion and structure of a rigid body using the spherical projection. *Computer Vision, Graphics, and Image Processing*, 21:21–32, 1983.

[175] A. Yuille and D. Geiger. Stereo and controlled movement. *The International Journal of Computer Vision*, 4:141–152, 1990. A similar paper was published by the same authors in Biological Cybernetics.

[176] Z. Zhang and O.D. Faugeras. *3D dynamic scene analysis*. Springer-Verlag, 1992.

[177] O.A. Zuniga and R.M. Haralick. Corner detection using the facet model. In *Proc. International Conference on Computer Vision and Pattern Recognition*, pages 30–37, 1983.