



**HAL**  
open science

# Mesure de formes et de champs de déplacements tridimensionnels par stéréo-corrélation d'images

Dorian Garcia

► **To cite this version:**

Dorian Garcia. Mesure de formes et de champs de déplacements tridimensionnels par stéréo-corrélation d'images. Traitement du signal et de l'image [eess.SP]. Institut National Polytechnique de Toulouse - INPT, 2001. Français. NNT: . tel-00166476

**HAL Id: tel-00166476**

**<https://theses.hal.science/tel-00166476>**

Submitted on 6 Aug 2007

**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.

Thèse présentée pour l'obtention du titre de

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE  
DE TOULOUSE

— spécialité : informatique —

par

**Dorian GARCIA**

ingénieur diplômé de l'INSAT

## Mesure de formes et de champs de déplacements tridimensionnels par stéréo-corrélation d'images

Soutenue le 21 décembre 2001 devant le jury composé de :

rapporteurs	M. Rachid DERICHE	Directeur de Recherche, INRIA Sophia-Antipolis
	M. Michel DHOME	Directeur de Recherche, LASMEA Clermont-Ferrand
examineurs	M. Alain AYACHE	Professeur, INPT-ENSEEIH Toulouse
	M. Christophe LEVAILLANT	Professeur, École des Mines d'Albi
	M. Michael A. SUTTON	Professeur, South Carolina University (USA)
	M. Alain VAUTRIN	Professeur, École des Mines de Saint-Étienne
directeur	M. Michel DEVY	Directeur de Recherche, LAAS-CNRS Toulouse
co-directeur	M. Jean-José ORTEU	Maitre-assistant (HDR), École des Mines d'Albi

Thèse préparée à l'École des Mines d'Albi







# Remerciements

Me voici enfin élevé au rang de Docteur ! C'est un vieux rêve que je caressais depuis ma tendre enfance : être chercheur. S'il est vrai que mes premières amours étaient plutôt orientées vers la synthèse d'images, par une suite d'hasards heureux je me suis dirigé vers la vision artificielle. « Hasards heureux » car le chemin qui m'y a conduit était jalonné de rencontres exceptionnelles, décisives, de celles qui vous dessinent de nouveaux horizons. La première d'entre elles remonte à 1994 avec la connaissance de Jean-José ORTEU, alors chercheur responsable du jeune laboratoire de vision artificielle de la jeune École des Mines d'Albi, puis, deux ans plus tard, avec la rencontre de Michel DEVY, chercheur au Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS de Toulouse. Tous deux m'ont donné le goût de la recherche en vision par ordinateur et devinrent plus tard mes directeurs de thèse. Au delà de cette fonction administrative et scientifique, ils devinrent naturellement des amis. Je tiens à leur rendre un hommage appuyé, pour leurs conseils, leurs remarques toujours pertinentes, la liberté et la confiance qu'ils m'ont accordées, et leur grande sympathie. Merci chefs !

Je dois en partie ma situation actuelle à mes anciens enseignants du lycée Louis RASCOL d'Albi, qui m'ont soutenu et encouragé à poursuivre mes études en école d'ingénieur. En particulier Bruno FENEUIL, Bruno VIGUIÉ, Jean-Claude JOULIÉ et Jean BANQUET, qui font avec beaucoup de passion et de sérieux leur beau métier d'enseignant. Que votre vocation fasse d'autres émules. Merci pour tout Messieurs.

Merci à Rachid DERICHE et Michel DHOME pour m'avoir fait l'immense honneur d'accepter d'être rapporteur de cette thèse. J'adresse également un grand merci à tous les examinateurs. Paradoxalement à l'idée que je m'en faisais, cela aura été un exercice fort agréable de répondre à vos questions lors de la soutenance. Merci beaucoup aussi pour tous vos encouragements.

Évidemment, j'adresse un grand merci à mes collègues, compagnons de route, avec lesquels j'ai eu un plaisir formidable à travailler dans une ambiance très (parfois trop ?) décontractée : Bruno DAVID pour sa gentillesse et ses exceptionnelles qualités en informatique, Vincent VELLAY le mathématicien qui a toujours répondu présent à mes appels au secours, Luc PENAZZI qui m'a sorti des méandres obscurs du logiciel ABAQUS™ et Étienne PERCHAT pour ses discussions passionnantes en algorithmique numérique. Je remercie vivement Gérard BERNHART, et Andrew Franz BENNETT pour leur participation à la relecture de ce manuscrit. Ce travail est aussi la contribution de plusieurs stagiaires ingénieurs avec lesquels j'ai eu beaucoup de plaisir à collaborer : Vincent LEMONDE, Jacques BARNAUD, Thomas VIEIL, Ariel CHOUKROUN et Nicolas CORNILLE. Je remercie également les techniciens et personnels administratifs de l'École des Mines pour leur sympathie et leur efficacité.

Enfin, je dois de merveilleux moments de détente à d'autres amis parmi lesquels, pêle-mêle : Alexis OUDIN, Serge MONTEIX, Christophe DAFFOS, Romain BEYDON, Emmanuel CAILLEUX et tous les autres thésards du CROMeP, Fab, Véro, Gigi, Chouchoutte, Bouddha, Francine, Bob, Sandrine, et tous les autres élèves ingénieurs de l'École des Mines qui ont eu à essayer mes premiers pas dans l'enseignement.



# Table des matières

Remerciements	iii
Introduction générale	1
<b>1 Modélisation et calibrage d'un capteur de stéréovision</b>	<b>7</b>
1.1 Modélisation d'une caméra	8
1.1.1 De la scène au pixel	8
1.1.2 Aberrations géométriques	10
1.1.3 Modèle de caméra avec distorsion	13
1.2 Calibrage d'une caméra	14
1.2.1 Méthode classique	15
1.2.2 Méthode photogrammétrique	17
1.2.3 Correction de la distorsion	18
1.2.4 Évaluation de la qualité du calibrage	19
1.3 Modélisation d'un capteur de stéréovision	20
1.3.1 Pourquoi utiliser deux caméras ?	20
1.3.2 Référentiels et changements de repères	21
1.3.3 Géométrie du capteur de stéréovision	22
1.4 Calibrage d'un capteur de stéréovision	26
1.4.1 Méthodes <b>LIN1</b> et <b>LIN<math>n</math></b>	27
1.4.2 Méthode <b>GLOB</b>	29
1.4.3 Comparaison des différentes approches	29
1.5 Conclusion	33
Bibliographie du chapitre	41
<b>2 Suivi de points par corrélation</b>	<b>43</b>
2.1 État de l'art	43
2.1.1 Appariement de primitives	43
2.1.2 Méthodes différentielles	46
2.1.3 Méthodes par corrélation	46
2.2 Corrélation	48
2.2.1 Approximation de la transformation locale	48
2.2.2 Coefficients de corrélation	49
2.2.3 Techniques d'estimation de la transformation locale	51
2.2.4 Corrélation subpixel : modèle de luminance	52
2.3 Méthode développée	52

2.4	Évaluation des performances . . . . .	55
2.4.1	Génération des images de synthèse . . . . .	57
2.4.2	Évaluation de l'incertitude sur la mesure de déplacements . . . . .	61
2.5	Conclusion . . . . .	74
	Bibliographie du chapitre . . . . .	75
<b>3</b>	<b>Mesure de formes tridimensionnelles par stéréo-corrélation</b>	<b>79</b>
3.1	Appariement d'images stéréoscopiques . . . . .	79
3.1.1	Géométrie épipolaire . . . . .	80
3.1.2	Rectification . . . . .	80
3.1.3	Technique d'appariement classique . . . . .	84
3.1.4	Appariement par corrélation précise . . . . .	88
3.1.5	Méthode développée : amélioration de la stéréo-corrélation précise . . . . .	91
3.2	Triangulation . . . . .	93
3.3	Évaluation des performances . . . . .	94
3.3.1	Pièce emboutie numéro 1 . . . . .	94
3.3.2	Pièce emboutie numéro 2 . . . . .	98
3.3.3	Rapidité . . . . .	102
3.4	Conclusion . . . . .	103
	Bibliographie du chapitre . . . . .	105
<b>4</b>	<b>Mesure de champs de déplacements tridimensionnels par stéréo-corrélation et appariement temporel d'images</b>	<b>107</b>
4.1	Méthode classique . . . . .	107
4.2	Suivi de pixels et stéréo-corrélation simultanés . . . . .	109
4.2.1	Critère de corrélation global . . . . .	113
4.2.2	Paramétrisation minimale des transformations géométriques locales . . . . .	113
4.2.3	Estimation des paramètres . . . . .	115
4.3	Calcul du champ de déformations . . . . .	115
4.4	Extension à la mesure de champ de déplacements en dynamique . . . . .	117
4.5	Évaluation des performances . . . . .	117
4.6	Conclusion . . . . .	119
	Bibliographie du chapitre . . . . .	123
<b>5</b>	<b>Applications</b>	<b>125</b>
5.1	Tôles minces embouties . . . . .	125
5.2	Membranes en élastomère soufflées . . . . .	132
5.3	Bétons réfractaires renforcés de fibres métalliques . . . . .	135
	Bibliographie du chapitre . . . . .	139
	<b>Conclusion générale et perspectives</b>	<b>141</b>
	<b>Annexes</b>	<b>145</b>
<b>A</b>	<b>Du modèle optique d'une caméra au modèle géométrique</b>	<b>145</b>
A.1	Modèle lentille mince . . . . .	145
A.2	Modèle lentille épaisse . . . . .	148

---

A.3	Modèle géométrique . . . . .	150
<b>B</b>	<b>Calibrage d'une caméra : méthode linéaire de RAVN <i>et al.</i></b>	<b>153</b>
<b>C</b>	<b>Calibrage par méthode photogrammétrique : unicité des solutions</b>	<b>157</b>
<b>D</b>	<b>Calcul du gradient des niveaux de gris d'une image rectifiée en fonction du gradient dans l'image initiale</b>	<b>161</b>
<b>E</b>	<b>Paramétrisation minimale des transformations géométriques locales de la méthode de stéréo-corrélation et tracking simultanés</b>	<b>165</b>
<b>F</b>	<b>Algorithme de Levenberg-Marquardt</b>	<b>169</b>
<b>G</b>	<b>Jacobiennes</b>	<b>173</b>
G.1	Jacobiennes du modèle de caméra . . . . .	173
G.2	Jacobiennes du modèle de capteur de stéréovision — méthode <b>GLOB</b> . . . . .	174
G.3	Jacobiennes du critère de corrélation de la méthode de stéréoscopie et suivi de pixels simultanés . . . . .	175
<b>H</b>	<b>Vecteur de rotation instantanée</b>	<b>179</b>
	<b>Bibliographie générale</b>	<b>181</b>



# Introduction générale

La réponse d'un matériau sous sollicitations mécaniques, thermiques ou thermo-mécaniques est définie par sa loi de comportement. Celle-ci, intégrée dans des codes de calcul par éléments finis, permet de prévoir la réponse de la pièce, de l'outillage, de l'ouvrage d'art, *etc.*, conçus avec ce matériau, et soumis à ces sollicitations environnementales. Cette simulation est un outil permettant d'accroître les performances technologiques des pièces par un meilleur dimensionnement, d'assurer une plus grande sécurité de fonctionnement, ou de réduire les coûts de production. Le succès d'une telle approche repose sur la pertinence des modèles de comportement établis à partir des mesures des champs de déformations et de température du matériau lors d'essais mécaniques et/ou thermiques.

Le moyen expérimental classique est l'extensométrie mécanique qui permet la détermination des déformations à la surface de la pièce, le plus souvent à partir d'extensomètres et de jauges [LGA92a, LGA92b, Pol00]. Ces moyens de mesures sont performants car ils offrent une simplicité de mise en œuvre, des mesures en temps réel, une bonne résolution ainsi qu'une grande précision. Cependant ils fournissent le plus souvent des mesures ponctuelles, unidirectionnelles et la qualité de la mesure est fortement tributaire de la liaison du capteur avec l'objet étudié (des efforts parasites peuvent perturber la mesure). En raison de tous ces inconvénients, les techniques optiques ont connu ces dernières années un essor important. Celles-ci se sont rapidement imposées en raison de leur caractère non intrusif, de leur grande résolution spatiale, de leur sensibilité élevée, de l'importance du champ examiné à tout instant et des progrès de l'informatique qui permet le traitement automatique d'un grand volume d'information.

Ainsi sont apparues les méthodes d'extensométrie optique bidimensionnelle utilisant une seule caméra, très utilisées aujourd'hui pour la mécanique expérimentale [Pho95, Pho98, Pho01]. Pour des raisons qui seront exposées plus loin, nous avons décidé de focaliser nos travaux sur les techniques de vision basées sur la corrélation d'images. SUTTON *et al.* furent parmi les premiers à proposer dès 1983 d'utiliser une méthode de corrélation d'images pour mesurer les déplacements dans le plan d'une éprouvette plane [SWP<sup>+</sup>83]. Durant les deux décennies suivantes, le nombre d'applications reposant sur ce moyen de mesure n'a pas cessé de croître [SMHS00]. Parmi celles-ci, et à titre d'exemple, citons les travaux de YANG [Yan92] mettant en œuvre cette technique pour des expériences de cisaillement de sables, de béton en compression, et de fissuration d'une éprouvette CT. TARDIF [Tar98] s'en est servi pour comprendre les mécanismes de dégradation d'un nouveau matériau composite carbone/carbone utilisé dans la conception de disques de freins, de cols de tuyère des propulseurs, des soles de four, *etc.* LARABA-ABBES [LA98] a aussi employé l'une de ces techniques pour l'étude des comportements hyperélastiques et viscohyperélastiques de deux types d'élastomères. DOUMALIN [Dou00] a appliqué ces techniques aux cas d'images acquises avec un microscope électronique à balayage pour l'étude micromécanique du comportement de deux matériaux. Ces quelques exemples donnent un aperçu de la vaste étendue des applications possibles de ce moyen de

mesure optique par caméra.

Malgré cette grande diversité d'applications des méthodes d'extensométrie bidimensionnelle par vision monoculaire, une limitation majeure en restreint son usage : il n'est possible d'accéder qu'à des déplacements dans le plan. Tout déplacement hors-plan introduit un grandissement dans les images qui conduit à des erreurs de mesure significatives des déplacements supposés dans le plan. Pour cette raison, certains auteurs ont proposé d'utiliser une méthode de stéréovision (deux caméras) pour accéder aux trois composantes d'un champ de déplacements tridimensionnels [VL89, KJC90, LCSP93, OGD97, SS99, GH99, LAE00, SMHS00].

La stéréovision, utilisant deux caméras (nous ne nous intéressons pas ici aux méthodes de stéréovision axiale utilisant une seule caméra [RA94]), repose sur le principe de triangulation [Fau93] : lorsqu'un point de l'espace est visible simultanément par les deux caméras, il est possible de déterminer ses coordonnées tridimensionnelles en calculant l'intersection des deux rayons optiques qui passent respectivement par les centres optiques des deux caméras et les points image. Ce calcul nécessite la résolution des deux problèmes fondamentaux de la stéréovision. Le premier est l'appariement stéréoscopique : étant donné un point dans l'image d'une caméra, il faut déterminer son homologue dans l'image de l'autre caméra. Le deuxième problème concerne le calcul de l'intersection dans l'espace des deux rayons optiques associés : le capteur de stéréovision doit avoir été préalablement calibré.

En collaboration avec le Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS-CNRS) de Toulouse, le Laboratoire de Vision Artificielle et Thermographie Infrarouge de l'École des Mines d'Albi a commencé en 1996 [OGD97] à s'intéresser à la stéréovision pour la mesure de champs de déplacements tridimensionnels sur des tôles minces embouties. Ces travaux ont été étendus en 1997 [O AHL98] à la stéréo-corrélation, utilisée couramment dans le domaine de la reconstruction tridimensionnelle en robotique [FHM<sup>+</sup>93], pour la mesure de formes et de champs de déplacements tridimensionnels dans le cadre de la mécanique expérimentale. Ces travaux préliminaires ont montré la faisabilité de l'approche (déjà montrée du reste par SUTTON et son équipe [LCS94]), et ce travail de thèse a consisté à améliorer autant que possible chacune des étapes de la chaîne de mesure.

Cette thèse a **deux objectifs majeurs distincts** : le premier concerne **la mesure précise de la forme tridimensionnelle** d'un objet statique, le deuxième doit permettre **la mesure de champ de déplacements tridimensionnels** d'un objet déformé ou se déformant dynamiquement. La mesure précise de forme présente un intérêt majeur pour des applications de rétro-conception par exemple ; elle est aussi une étape intermédiaire pour la mesure de champs de déplacements tridimensionnels comme nous le verrons dans le chapitre 4. La mesure d'un champ de déplacements tridimensionnels est une étape préliminaire permettant ensuite le calcul d'un champ de déformations en surface de l'objet.

Le **chapitre 1** de cette thèse présente la **modélisation** et le **calibrage d'un capteur de stéréovision** composé de deux caméras. Étant donné la finalité métrologique de notre travail, nous nous sommes intéressés aux techniques de calibrage fort et une attention particulière a été portée sur la qualité du modèle de caméra proposé et sur l'estimation des paramètres de ce modèle. Ainsi, les principales aberrations géométriques liées au système optique de chacune des caméras sont prises en compte dans la modélisation. Nous étudions ensuite les moyens

classiques de calibrer un capteur de stéréovision à partir de deux caméras préalablement calibrées. Enfin, nous proposons une approche photogrammétrique globale considérant le capteur de stéréovision comme un tout indivisible et montrons expérimentalement l'intérêt de cette approche pour l'obtention de mesures précises [GOD00].

Le **chapitre 2** est consacré au **suivi de points dans une séquence d'images acquise par une seule caméra**. Une brève étude des trois techniques majeures existantes (suivi de marqueurs, flot optique et corrélation) nous a conduit à choisir une méthode par corrélation directe d'images pour sa facilité d'utilisation et la grande précision des appariements qu'elle permet. Après une présentation des principaux critères de corrélation utilisés dans la littérature, nous exposons une méthode d'appariement très précise prenant en compte la déformation locale des domaines de corrélation. Les performances de la méthode que nous avons développée sont alors analysées en détail sur des images de synthèse. Les méthodes présentées dans ce chapitre permettent donc de disposer d'un outil d'extensométrie optique bidimensionnelle par vision monoculaire. Cependant le contenu de ce chapitre a pour principal objectif de proposer une solution au problème de l'appariement temporel de points qui nous sera nécessaire pour les techniques de mesure de champs de déplacements tridimensionnels présentées au chapitre 4.

Le **chapitre 3** traite du problème de l'**appariement stéréoscopique** et de la **reconstruction tridimensionnelle**. Après un bref rappel de quelques propriétés géométriques d'un capteur stéréoscopique, l'opération classique de rectification, qui permet notamment de simplifier et d'accélérer les algorithmes d'appariement, est présentée. Ces bases posées, nous étudions une méthode de stéréo-corrélation précise d'images rectifiées, prenant en compte la transformation géométrique locale des domaines de corrélation, comme nous l'avons fait pour nos algorithmes de suivi de points par corrélation du chapitre 2. Cette méthode de stéréo-corrélation précise a déjà été exposée par DEVERNAY et FAUGERAS en 1994 dans [DF94]. Nous en proposons une amélioration qui se base directement sur les images initiales dans le but d'éviter d'exploiter des images dégradées par l'opération de rectification et de correction de la distorsion. Ce chapitre répond alors à la première de nos deux problématiques : la mesure de forme tridimensionnelle par stéréo-corrélation d'images.

Le **chapitre 4** expose en détail les solutions que nous avons retenues pour répondre à notre deuxième problématique : la **mesure de champs de déplacements tridimensionnels** à partir d'une séquence de paires d'images stéréoscopiques d'un objet déformé ou se déformant dynamiquement. La méthode classique est présentée et discutée : elle repose sur les techniques exposées dans les chapitres 2 et 3 et consiste à traiter de façon séquentielle deux paires d'images acquises à deux états de déformation. Nous proposons une nouvelle méthode qui considère de façon globale les quatre images afin de déterminer de façon optimale les appariements stéréoscopiques et temporels. Comme pour la mesure de formes tridimensionnelles, les techniques mises en œuvre reposent sur les images brutes (non dégradées par les opérations de rectification et de correction de la distorsion).

Le **chapitre 5** présente quelques **applications** de nos méthodes de mesure de formes et de champs de déplacements tridimensionnels au travers de plusieurs expériences [GO01]. La première concerne l'emboutissage de tôles minces. Les reconstructions tridimensionnelles avant et après emboutissage sont montrées, ainsi que le champ des déplacements entre ces deux formes et le champ de déformations de surface. Une deuxième application concerne le calcul du champ de déplacements en surface d'une membrane en élastomère soufflée. Enfin, la dernière expérience concerne l'utilisation de notre méthode de mesure de déplacements tridimensionnels pour simuler des extensomètres virtuels sur une éprouvette en béton lors d'un essai de traction uniaxial.

Le lecteur trouvera dans les annexes quelques détails de calculs qui lui permettront d'implanter facilement les méthodes proposées dans cette thèse. Les autres annexes apportent un éclaircissement sur divers points de ce travail.

## Notations

$s$	un scalaire
$\mathbf{v}$	un vecteur de $\mathbb{R}^n$
$\mathbf{v}^\top$	le vecteur $\mathbf{v}$ transposé, <i>i.e.</i> $\mathbf{v}^\top = (x_1 \ x_2 \ \dots \ x_n)$
$\mathbf{A}$	une matrice
$M$	un point de l'espace tridimensionnel $\mathbb{R}^3$ , <i>e.g.</i> de coordonnées $M = (x \ y \ z)^\top$
$\mathbf{m}$	un point de l'espace bidimensionnel $\mathbb{R}^2$ , <i>e.g.</i> de coordonnées $\mathbf{m} = (u \ v)^\top$
$\hat{m}$	un point d'une image dont les coordonnées $\hat{u}$ et $\hat{v}$ sont déterminées expérimentalement
$\tilde{m}, \tilde{M}$	les points $\mathbf{m}$ et $M$ exprimés en coordonnées homogènes, <i>i.e.</i> respectivement dans les espaces projectifs $\mathbb{P}^2$ et $\mathbb{P}^3$
$\cong$	l'égalité projective, <i>i.e.</i> l'égalité définie à un facteur non nul près : $\tilde{m} \cong s\tilde{m}$ , $\forall s \in \mathbb{R}^*$
$\mathcal{R}_c$	un référentiel de l'espace bi- ou tridimensionnel



# Chapitre 1

## Modélisation et calibrage d'un capteur de stéréovision

En vision par ordinateur, dès lors que l'on souhaite obtenir des informations métriques en vue d'applications de mesures dimensionnelles de précision, le calibrage du capteur est d'une importance cruciale. Calibrer un capteur monoculaire (une seule caméra<sup>1</sup>) consiste à déterminer ses paramètres intrinsèques et sa position et orientation par rapport au référentiel du monde. Calibrer un capteur de stéréovision, constitué de deux caméras, consiste à déterminer les paramètres intrinsèques de chacune des caméras, et la position et orientation relative de ces deux capteurs. Ces paramètres de calibrage sont nécessaires pour calculer, par triangulation, les coordonnées tridimensionnelles d'un point correspondant à des pixels appariés dans les deux images.

De nombreux travaux ont été effectués concernant le calibrage de caméra et la reconstruction tridimensionnelle avec un capteur de vision stéréoscopique ou une caméra en mouvement. La plupart des méthodes nécessitent une mire de calibrage de géométrie parfaitement connue pour calibrer hors-ligne une caméra ou un banc de stéréovision [Tos87, FT87]. Plus récemment, des méthodes d'auto-calibrage ont été proposées pour calibrer en ligne une caméra ou un banc de stéréovision sans connaissance préalable sur l'objet d'étalonnage (voir [LD99, LD00] et [Fus00] pour un état de l'art). Ces méthodes conviennent bien aux systèmes de vision active, pour lesquels les paramètres de caméra peuvent varier continuellement, mais ne permettent pas de fournir une information métrique précise (de plus, la plupart de ces méthodes ne prennent pas en compte la distorsion induite par le système optique de la caméra). Pour les deux problématiques posées, la mesure de formes et de champs de déplacements tridimensionnels, nous nous intéressons à l'obtention de mesures dimensionnelles précises, c'est pourquoi nous avons fait le choix de nous intéresser aux méthodes de calibrage fort.

Nous commençons section 1.1 par la modélisation d'une caméra prenant en compte les principales aberrations géométriques liées à son système optique. Puis, section 1.2, nous présentons deux méthodes de calibrage d'une caméra : l'une classique, l'autre utilisant une approche photogrammétrique. Dans la section 1.3 nous modélisons un capteur de stéréovision, composé de deux caméras liées rigidement. Le calibrage d'un tel capteur est étudié dans la section 1.4, avec notamment une nouvelle méthode photogrammétrique basée sur une approche globale.

---

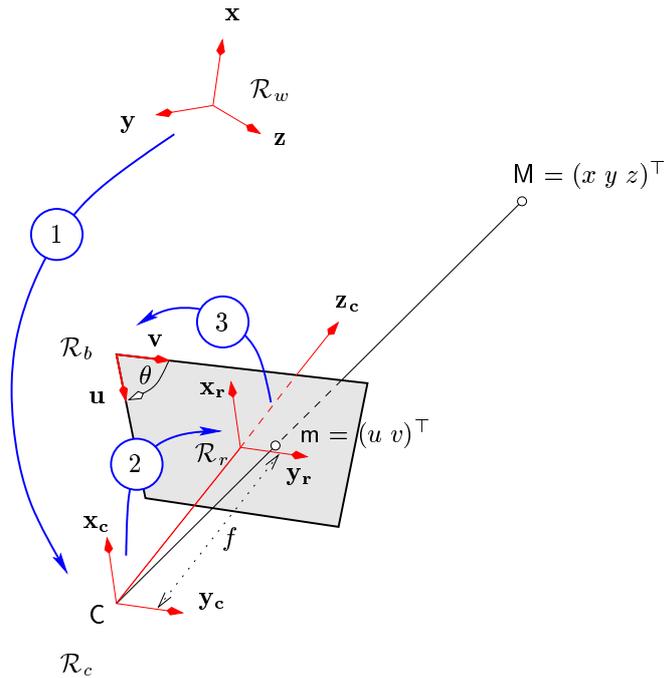
<sup>1</sup>Par abus de langage, nous désignerons par « caméra », le capteur couplé à son système optique : l'objectif.

## 1.1 Modélisation d'une caméra

Une caméra est un système optronique composé de deux éléments : le premier est le détecteur qui est un dispositif à transfert de charges, le deuxième élément est un système optique placé entre la source et le détecteur qui a pour but de focaliser l'image d'un objet sur ce détecteur. Nous étudions dans cette section la modélisation géométrique de ce capteur considéré comme parfait, conduisant au modèle de projection perspective dit **sténopé**. Ensuite sont présentées les principales aberrations géométriques rencontrées avec des objectifs courants ; un modèle de caméra prenant en compte ces distorsions est alors établi. Le lecteur trouvera dans l'annexe A une justification physique du modèle géométrique de caméra exposé ci-après, à partir d'une modélisation optique de ce capteur. L'hypothèse de validité du modèle géométrique y est notamment détaillée (effet de la défocalisation négligeable).

### 1.1.1 De la scène au pixel

Le modèle géométrique classique d'une caméra est une projection perspective, résultat de la composition des trois transformations élémentaires [Fau93, HM95, Gar96] présentées figure 1.1. La première est un simple changement de repère : elle transforme les coordonnées



**FIG. 1.1 :** Les trois transformations élémentaires du modèle sténopé de caméra et ses référentiels. La première transformation exprime les coordonnées d'un point de la scène dans le référentiel de la caméra. La deuxième est la projection perspective de ce point sur le plan rétinien. La troisième exprime ce dernier point dans le référentiel de l'image.

d'un point de l'espace  $M = (x \ y \ z)^\top$  exprimé dans le repère du monde<sup>2</sup>  $\mathcal{R}_w$ , en un point

<sup>2</sup>On appelle aussi ce référentiel « repère de la scène ».

$M_c = (x_c \ y_c \ z_c)^\top$  de l'espace exprimé dans le référentiel  $\mathcal{R}_c$  de la caméra. Les paramètres de cette transformation sont dits **extrinsèques** et sont ceux d'une transformation rigide, i.e. une translation et une rotation. En utilisant les coordonnées homogènes, nous pouvons écrire la relation :

$$\tilde{M}_c \cong \underbrace{\begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\mathbf{T}} \tilde{M} \quad (1.1)$$

avec  $\mathbf{T} \cong \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix}$

La matrice  $\mathbf{R}$  est une matrice de rotation et  $\mathbf{t}$  est un vecteur de translation.

La deuxième transformation décrit l'optique : elle projette le point  $M_c$  sur le plan rétinien de la caméra en un point  $m_r = (x_r \ y_r)^\top$ . C'est une projection perspective, de rapport la distance focale  $f$ , dont l'expression nous est donnée par l'équation (A.3) :

$$\tilde{m}_r \cong \underbrace{\begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\tilde{\mathbf{P}}} \tilde{M}_c \quad (1.2)$$

Enfin, la troisième transformation décrit l'échantillonnage du plan rétinien par la matrice des capteurs CCD. Elle permet donc d'exprimer un point du plan rétinien en « coordonnées pixels » dans l'image  $\mathbf{m} = (u \ v)^\top$  [HM95] :

$$\tilde{\mathbf{m}} \cong \begin{pmatrix} k_u & k_u \cot \theta & u_0 + v_0 \cot \theta \\ 0 & k_v / \sin \theta & v_0 / \sin \theta \\ 0 & 0 & 1 \end{pmatrix} \tilde{m}_r$$

Les paramètres  $k_u$  et  $k_v$  sont les facteurs d'échelle vertical et horizontal (en pixel/unité de longueur),  $u_0$  et  $v_0$  sont les coordonnées dans l'image du centre de projection, et  $\theta$  traduit la non orthogonalité<sup>3</sup> des lignes et colonnes de l'image. Généralement, ce dernier paramètre est très proche de  $\pi/2$ , aussi nous simplifions cette transformation en le considérant égal à  $\pi/2$  :

$$\tilde{\mathbf{m}} \cong \underbrace{\begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{\tilde{\mathbf{A}}} \tilde{m}_r \quad (1.3)$$

$\tilde{\mathbf{A}}$  est donc une transformation affine à quatre paramètres non nuls appelés **intrinsèques**.

La composition de ces trois transformations permet d'exprimer les coordonnées image d'un point en fonction de ces coordonnées dans le référentiel de la scène :

$$\tilde{\mathbf{m}} \cong \underbrace{\tilde{\mathbf{A}} \tilde{\mathbf{P}} \mathbf{T}}_{\mathbf{K}} \tilde{M} \quad (1.4)$$

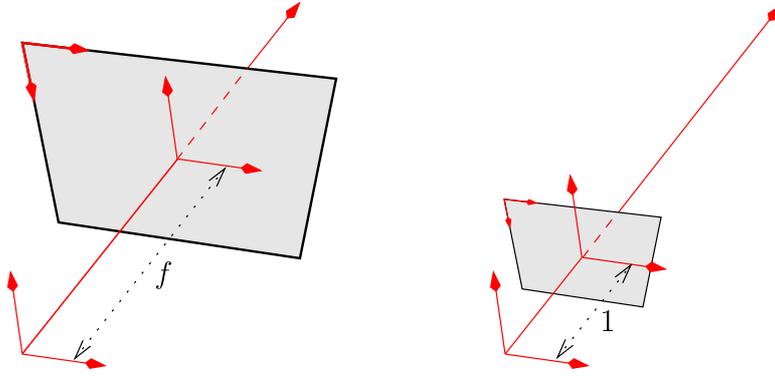
---

<sup>3</sup>Ce paramètre vient de l'observation qu'un point la scène se projette en un point de l'image par une homographie de  $\mathbb{P}^3$  dans  $\mathbb{P}^2$ . Celle-ci est donc une matrice  $3 \times 4$  définie par 11 paramètres. 6 correspondent aux paramètres extrinsèques, il n'en reste donc plus que 5 :  $u_0$ ,  $v_0$ ,  $\alpha_u$ ,  $\alpha_v$  et un autre supplémentaire. Ce dernier est généralement choisi comme l'angle  $\theta$  que forment les deux axes  $u$  et  $v$  du référentiel image.

La longueur focale  $f$  étant toujours en facteur d'autres paramètres ( $k_u$  et  $k_v$ ) du modèle, il est impossible de la dissocier de ces derniers [Gar96]. Cela vient du fait que ce modèle ne peut différencier des images produites par une caméra de focale  $f$  et de pixels de dimensions  $k_u^{-1} \times k_v^{-1}$  par rapport à une caméra de focale 1 et de pixels de dimensions  $(k_u f)^{-1} \times (k_v f)^{-1}$  (cf. figure 1.2) :

$$\mathbf{K} \cong \bar{\mathbf{A}}\bar{\mathbf{P}} \cong \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cong \underbrace{\begin{pmatrix} k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathbf{P}}$$

Il est d'usage de noter  $\alpha_u = k_u f$  et  $\alpha_v = k_v f$ . Avec cette convention, la matrice  $\mathbf{K}$  des



**FIG. 1.2 :** Deux modèles de caméra produisant les mêmes images. À gauche, la focale est égale à  $f$  et les pixels ont pour dimensions  $k_u^{-1} \times k_v^{-1}$ , à droite la focale est unitaire et les pixels ont pour dimensions  $(k_u f)^{-1} \times (k_v f)^{-1}$ .

paramètres intrinsèques du modèle sténopé s'écrit :

$$\mathbf{K} \cong \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (1.5)$$

L'équation (1.4) nous donne les **relations de colinéarité** [KW97] suivantes pour tout point  $\mathbf{M} = (x \ y \ z)^\top$  :

$$\mathbf{m} = (u \ v)^\top \quad \text{avec} \quad \begin{aligned} u &= \alpha_u \frac{r_{11}x + r_{12}y + r_{13}z + t_x}{r_{31}x + r_{32}y + r_{33}z + t_z} + u_0 \\ v &= \alpha_v \frac{r_{21}x + r_{22}y + r_{23}z + t_y}{r_{31}x + r_{32}y + r_{33}z + t_z} + v_0 \end{aligned} \quad (1.6)$$

### 1.1.2 Aberrations géométriques

Dans un système optique centré, un rayon n'intersecte généralement pas le plan image au point déterminé sous l'approximation linéaire de GAUSS. Ces écarts à l'optique paraxiale ne sont plus négligeables si la distance du point objet à l'axe optique est grande et les angles

d'inclinaison des rayons incidents ne sont plus petits [Pér94]. De plus, il est difficile de réaliser techniquement un système optique dont toutes les lentilles sont parfaitement parallèles, leurs centres optiques parfaitement colinéaires, et le plan image parallèle au plan du détecteur. Pour des applications de mesure de précision, il est donc nécessaire de prendre en compte ces défauts dans le modèle de caméra [Bey92].

Si l'on pose  $\mathbf{m}_r = (x_r \ y_r)^\top$  le point idéal déterminé sous l'approximation de GAUSS,  $\mathbf{m}_d = (x_d \ y_d)^\top$  le point réel (distordu), la relation entre  $\mathbf{m}_r$  et  $\mathbf{m}_d$  est

$$\mathbf{m}_d = \mathbf{m}_r + \Delta \quad (1.7)$$

où  $\Delta = (\delta_x \ \delta_y)^\top$  est appelé **vecteur d'aberration** ou **distorsion**. Ce vecteur est lui-même la somme de plusieurs distorsions élémentaires et WENG propose dans [WCH92] d'exprimer  $\Delta$  sous la forme

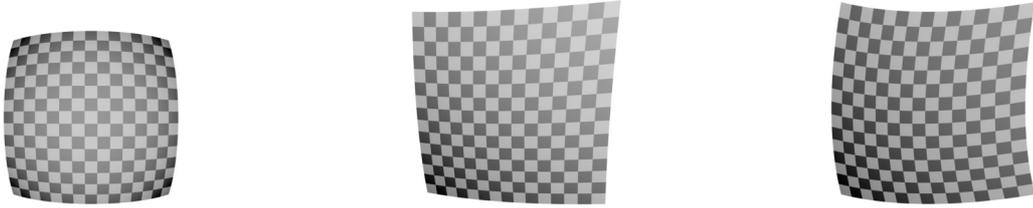
$$\Delta = \Delta_R + \Delta_D + \Delta_P$$

où :

$\Delta_R = (\delta_{R_x} \ \delta_{R_y})^\top$  est la **distorsion radiale** induite par l'approximation paraxiale (appelée « distorsion géométrique » dans la théorie des aberrations des systèmes centrés) — cf. figure 1.3-1 ;

$\Delta_D = (\delta_{D_x} \ \delta_{D_y})^\top$  est la **distorsion de décentrage** ayant pour origine l'imperfection d'alignement des centres optiques des lentilles qui composent le système optique — cf. figure 1.3-2 ;

$\Delta_P = (\delta_{P_x} \ \delta_{P_y})^\top$  est la **distorsion prismatique** causée par un mauvais parallélisme des lentilles entre elles et/ou avec le plan du détecteur — cf. figure 1.3-3.



(1) distorsion radiale      (2) distorsion de décentrage      (3) distorsion prismatique

**FIG. 1.3 :** Images de synthèse montrant les effets des trois types de distorsion sur l'image d'un damier.

Il est plus aisé d'exprimer chacune de ces distorsions élémentaires dans un système de coordonnées polaires dont l'origine varie selon la distorsion considérée. Ces origines sont toutefois très proches du point principal et on peut donc les considérer confondues avec ce dernier. On pose donc :

$$\rho = \sqrt{x_r^2 + y_r^2} \quad \text{et} \quad \cos \phi = x_r / \rho \quad \text{et} \quad \sin \phi = y_r / \rho \quad (1.8)$$

Chaque vecteur de distorsion élémentaire est alors exprimé sous la forme d'un terme de distorsion radiale  $\delta_{.R}(\rho, \phi)$  et d'un terme de distorsion tangentielle  $\delta_{.T}(\rho, \phi)$  tels que [Zel96] :

$$\begin{pmatrix} \delta_{.x}(x_r, y_r) \\ \delta_{.y}(x_r, y_r) \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \delta_{.rad}(\rho, \phi) \\ \delta_{.tan}(\rho, \phi) \end{pmatrix} \quad (1.9)$$

### Distorsion de Seidel (distorsion radiale)

Cette distorsion est liée à l'approximation linéaire de GAUSS, ou à un défaut de courbure de la lentille, qui a pour effet de déplacer radialement le point d'intersection d'un rayon avec le plan image. Ce déplacement est d'autant plus important que le point objet s'éloigne de l'axe optique, *cf.* figure 1.3-1. Son expression est :

$$\begin{aligned}\delta_{R_{rad}}(\rho, \phi) &= \sum_{k=1}^{\infty} r_k \rho^{2k+1} = r_1 \rho^3 + r_2 \rho^5 + r_3 \rho^5 + \dots \\ \delta_{R_{tan}}(\rho, \phi) &= 0\end{aligned}$$

En utilisant les équations (1.8) et (1.9) on obtient l'expression de la distorsion radiale en coordonnées cartésiennes :

$$\begin{aligned}\delta_{R_x}(x_r, y_r) &= x \sum_{k=1}^{\infty} r_k (x_r^2 + y_r^2)^k \\ \delta_{R_y}(x_r, y_r) &= y \sum_{k=1}^{\infty} r_k (x_r^2 + y_r^2)^k\end{aligned}\tag{1.10}$$

### Distorsion de décentrage

La distorsion de décentrage est une aberration géométrique provenant d'un alignement imparfait des centres optiques des lentilles du système optique. Elle a pour expression :

$$\begin{aligned}\delta_{D_{rad}}(\rho, \phi) &= 3 \sin(\phi - \phi_0) \sum_{k=1}^{\infty} s_k \rho^{2k} = 3 \sin(\phi - \phi_0) (s_1 \rho^2 + s_2 \rho^4 + s_3 \rho^5 + \dots) \\ \delta_{D_{tan}}(\rho, \phi) &= \cos(\phi - \phi_0) \sum_{k=1}^{\infty} s_k \rho^{2k} = \cos(\phi - \phi_0) (s_1 \rho^2 + s_2 \rho^4 + s_3 \rho^5 + \dots)\end{aligned}$$

En utilisant les équations (1.8) et (1.9) on obtient l'expression de la distorsion de décentrage en coordonnées cartésiennes :

$$\begin{aligned}\delta_{D_x}(x_r, y_r) &= (2x_r y_r \cos \phi_0 - (3x_r^2 + y_r^2) \sin \phi_0) \sum_{k=1}^{\infty} s_k (x_r^2 + y_r^2)^{k-1} \\ \delta_{D_y}(x_r, y_r) &= (-2x_r y_r \sin \phi_0 + (x_r^2 + 3y_r^2) \cos \phi_0) \sum_{k=1}^{\infty} s_k (x_r^2 + y_r^2)^{k-1}\end{aligned}\tag{1.11}$$

### Distorsion prismatique

Lorsque les lentilles, qui composent le système optique, et le plan du détecteur ne sont pas parfaitement parallèles entre eux, cela provoque une aberration géométrique identique à celle que l'on obtiendrait pour un système parfait dans lequel on insérerait un prisme fin. Cette

aberration, appelée distorsion prismatique, a pour expression :

$$\begin{aligned}\delta_{P_{rad}}(\rho, \phi) &= \sin(\phi - \phi_1) \sum_{k=1}^{\infty} t_k \rho^{2i} = \sin(\phi - \phi_1) (t_1 \rho^2 + t_2 \rho^4 + t_3 \rho^6 + \dots) \\ \delta_{P_{tan}}(\rho, \phi) &= \cos(\phi - \phi_1) \sum_{k=1}^{\infty} t_k \rho^{2i} = \cos(\phi - \phi_1) (t_1 \rho^2 + t_2 \rho^4 + t_3 \rho^6 + \dots)\end{aligned}$$

En utilisant les équations (1.8) et (1.9) on obtient l'expression de la distorsion prismatique en coordonnées cartésiennes :

$$\begin{aligned}\delta_{P_x}(x_r, y_r) &= -\sin \phi_1 \sum_{k=1}^{\infty} t_k (x_r^2 + y_r^2)^k \\ \delta_{P_y}(x_r, y_r) &= \cos \phi_1 \sum_{k=1}^{\infty} t_k (x_r^2 + y_r^2)^k\end{aligned}\tag{1.12}$$

### Approximations

Afin de prendre en compte la distorsion dans notre modèle de caméra, il est nécessaire de limiter l'ordre des développements des expressions (1.10), (1.11) et (1.12). Plusieurs auteurs ont montré qu'un développement à l'ordre 3 de la distorsion radiale est suffisant pour la plupart des objectifs de focale supérieure à 5 mm. La distorsion dans l'image est totalement dominée par sa composante radiale, par conséquent la distorsion de décentrage et la distorsion prismatique sont de faible importance. Nous limitons l'ordre de leur développement à 1. Cela conduit au modèle de distorsion suivant :

$$\begin{aligned}\delta_x &= x_r (r_1 \rho^2 + r_2 \rho^4 + r_3 \rho^6) + s_1 (2x_r y_r \cos \phi_0 - (3x_r^2 + y_r^2) \sin \phi_0) - t_1 \sin \phi_1 \rho \\ \delta_y &= y_r (r_1 \rho^2 + r_2 \rho^4 + r_3 \rho^6) + s_1 (2x_r y_r \sin \phi_0 - (x_r^2 + 3y_r^2) \cos \phi_0) + t_1 \cos \phi_1 \rho\end{aligned}$$

En posant :

$$\begin{aligned}d_1 &= s_1 \cos \phi_0 & p_1 &= -t_1 \sin \phi_1 \\ d_2 &= s_1 \sin \phi_0 & p_2 &= t_1 \cos \phi_1\end{aligned}$$

nous pouvons écrire le modèle de distorsion sous la forme suivante :

$$\begin{aligned}\delta_x &= x_r (r_1 \rho^2 + r_2 \rho^4 + r_3 \rho^6) + 2d_1 x_r y_r + d_2 (3x_r^2 + y_r^2) + p_1 \rho \\ \delta_y &= y_r (r_1 \rho^2 + r_2 \rho^4 + r_3 \rho^6) + 2d_2 x_r y_r + d_1 (x_r^2 + 3y_r^2) + p_2 \rho\end{aligned}\tag{1.13}$$

### 1.1.3 Modèle de caméra avec distorsion

La prise en compte de la distorsion radiale au troisième ordre, et des distorsions de décentrage et prismatique au premier ordre, a pour conséquence d'ajouter une transformation non linéaire dans notre modèle de caméra. Ainsi, un point image distordu  $\mathbf{m}$  n'est plus immédiatement relié à ses coordonnées rétinienne idéales  $\mathbf{m}_r$ , mais au point  $\mathbf{m}_d$ , coordonnées rétinienne distordues, par la relation :

$$\tilde{\mathbf{m}} \cong \mathbf{A} \tilde{\mathbf{m}}_d \quad \Longrightarrow \quad \begin{aligned}u &= \alpha_u x_d + u_0 \\ v &= \alpha_v y_d + v_0\end{aligned}$$

Des équations (1.7) et (1.13) nous déterminons l'expression du point rétinien distordu  $\mathbf{m}_d$  en fonction de ses coordonnées idéales  $\mathbf{m}_r$  :

$$\begin{aligned} x_d &= x_r + x_r \left( r_1 (x_r^2 + y_r^2) + r_2 (x_r^2 + y_r^2)^2 + r_3 (x_r^2 + y_r^2)^3 \right) \\ &\quad + 2d_1 x_r y_r + d_2 (3x_r^2 + y_r^2) + p_1 (x_r^2 + y_r^2) \\ \mathbf{m}_d = \mathbf{m}_r + \Delta &\quad \implies \\ y_d &= y_r + y_r \left( r_1 (x_r^2 + y_r^2) + r_2 (x_r^2 + y_r^2)^2 + r_3 (x_r^2 + y_r^2)^3 \right) \\ &\quad + 2d_2 x_r y_r + d_1 (x_r^2 + 3y_r^2) + p_2 (x_r^2 + y_r^2) \end{aligned}$$

Les relations entre les coordonnées rétinienne idéales  $\mathbf{m}_r$  et le point de la scène  $\mathbf{M}_c$  exprimé dans le repère caméra, puis entre ce dernier et ses coordonnées  $\mathbf{M}$  exprimées dans le repère de la scène sont inchangées et données respectivement par l'équation  $\tilde{\mathbf{m}}_r \cong \mathbf{P}\tilde{\mathbf{M}}_c$ , et l'équation (1.1).

Appelons  $\mathbf{k}$  le vecteur des paramètres intrinsèques définis par la matrice  $\mathbf{K}$ ,  $\mathbf{d}$  le vecteur des coefficients de distorsion (qui sont également intrinsèques à la caméra) :

$$\begin{aligned} \mathbf{k} &= (u_0 \ v_0 \ \alpha_u \ \alpha_v)^\top \\ \mathbf{d} &= (r_1 \ r_2 \ r_3 \ d_1 \ d_2 \ p_1 \ p_2)^\top \end{aligned}$$

Le modèle de caméra est non linéaire et peut s'écrire sous la forme d'une fonction vectorielle  $f$  :

$$\mathbf{m} = f(\mathbf{k}, \mathbf{d}, \mathbf{R}, \mathbf{t}, \mathbf{M}) \quad (1.14)$$

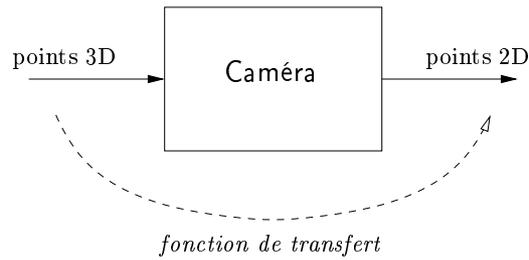
Nous venons d'établir les modèles linéaire (1.6) et non linéaire (1.14) d'une caméra, nous allons voir maintenant les méthodes dites de calibrage permettant d'estimer les paramètres de ces modèles.

## 1.2 Calibrage d'une caméra

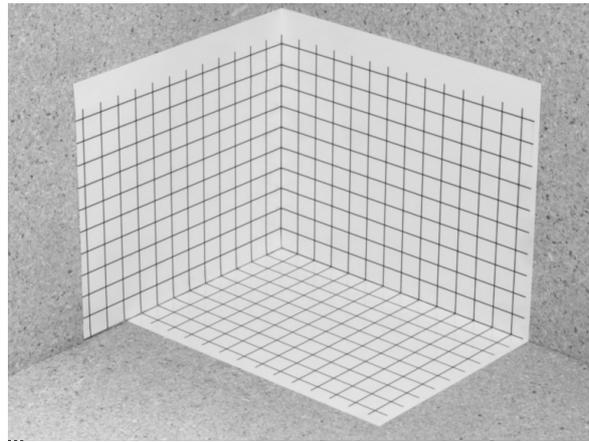
Calibrer une caméra consiste à estimer ses paramètres intrinsèques et sa position et orientation par rapport à un référentiel que l'on choisit : on le nomme **référentiel du monde ou de la scène**. Le calibrage d'un tel capteur revient à estimer les paramètres de sa fonction de transfert (ou modèle), équation (1.14), qui transforme un point de l'espace du monde en un point de l'image (*cf.* figure 1.4). Ce problème d'estimation peut alors être résolu si l'on connaît des points tridimensionnels exprimés dans le repère du monde et les coordonnées de leur projection dans les images : on parle alors de **calibrage fort**.

Les points tridimensionnels utilisés sont habituellement répartis sur un objet que l'on nomme **mire de calibrage** ou **d'étalonnage**. Les premières méthodes qui ont été proposées consistent à utiliser une mire de calibrage précise dite « volumique » parce que ses points d'étalonnage sont répartis dans un volume. Elles ont alors la géométrie d'un dièdre ou d'un trièdre (*cf.* figure 1.5). Ce type de mire permet d'estimer tous les paramètres du modèle de caméra avec une seule image, en revanche sa précision de construction est difficile à garantir et à maintenir, et sa manipulation n'est pas aisée. Ce sont les raisons pour lesquelles d'autres méthodes sont apparues, basées sur le déplacement précis (par une table de translation micrométrique) d'une mire plane précise.

Le calibrage utilisant une mire plane nécessite l'acquisition et le traitement d'au moins deux images de la mire pour pouvoir être résolu. Toutefois, il est plus facile de garantir une



**FIG. 1.4 :** *Le calibrage d'une caméra est un problème d'estimation des paramètres de sa fonction de transfert.*



**FIG. 1.5 :** *Exemple d'une mire volumique utilisée pour le calibrage d'une caméra.*

bonne précision sur la connaissance des points qui la compose, et sa manipulation devient simple du fait de son faible encombrement. La figure 1.8, page 19, montre une mire plane en verre de 20cm de côté dont le marquage a été réalisé par un procédé de microlithographie garantissant une précision de l'ordre du micromètre de la position des nœuds de sa grille.

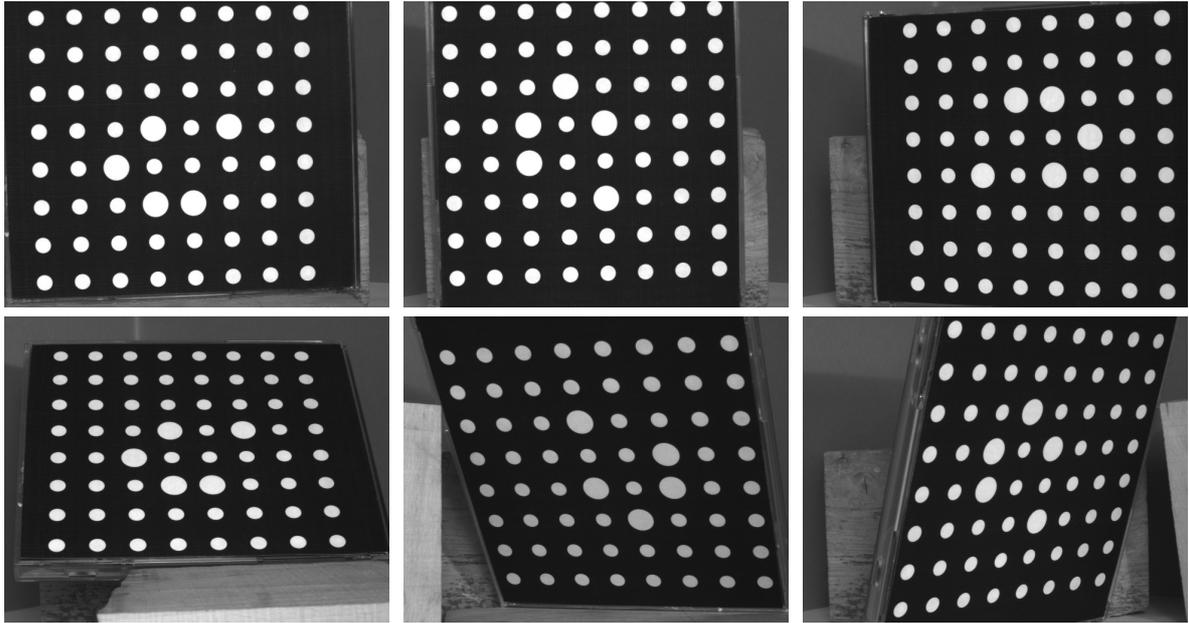
Dans la section suivante, nous allons d'abord présenter la méthode de calibrage classique, supposant les points de la mire connus avec exactitude. Nous étudierons ensuite une autre méthode de calibrage dite **photogrammétrique** qui nous permettra de nous affranchir d'une connaissance précise des points de la mire, et donc qui autorisera l'utilisation de mires de calibrage de faible coût de réalisation.

### 1.2.1 Méthode classique

L'approche classique du calibrage d'une caméra consiste à supposer le modèle de la mire connu avec exactitude. Les seuls paramètres à estimer sont donc les paramètres intrinsèques à la caméra (incluant les coefficients de distorsion), et chaque ensemble de paramètres extrinsèques correspondant à chaque position de la mire.

Nous considérons  $n$  images d'une mire (plane<sup>4</sup>) composée de  $p$  points acquises par une caméra. Les points de la mire peuvent être les nœuds d'intersections de droites horizontales

<sup>4</sup>La mire n'a pas besoin d'être plane mais, en pratique, il est aisé de tracer celle-ci sur une feuille de papier et de la coller sur une surface rigide plane ou quasi plane.



**FIG. 1.6 :** Exemple d'une séquence de six images d'une mire de calibrage composée de soixante quatre taches circulaires, utilisée pour le calibrage d'une caméra.

et verticales, estimés par l'extracteur de PEUCHOT décrit dans [Peu94, OGD97] ou par la méthode des splines de BRAND [Bra95], ou bien le centre d'une tache elliptique déterminé via l'estimateur de LAVEST [LVD99] ou la méthode de BRAND [Bra95]. D'autres méthodes efficaces d'extraction basées sur des modèles peuvent aussi être utilisées [Bla97]. La figure 1.6 illustre une séquence de six images d'une mire de calibrage composée de soixante quatre taches circulaires.

Il est intéressant de noter que des taches circulaires de trop grande dimension causent des erreurs de mesure. En effet, la projection d'une tache circulaire est elliptique si la surface de la mire n'est pas parallèle au plan image. Le centre de la tache circulaire de la mire n'est alors pas projeté au centre de la tache elliptique mesuré dans l'image. L'écart entre le centre réel projeté et celui mesuré peut alors être supérieur à la précision de la méthode d'extraction. Toutefois, l'approche photogrammétrique présentée dans la section suivante permet de compenser l'influence de cet écart [Bra95, Dol96].

Une fois les points de la mire extraits dans la séquence d'images, l'estimation des paramètres du modèle de caméra est posée comme un processus de minimisation. La fonctionnelle à minimiser est généralement la somme des erreurs quadratiques entre les coordonnées des points extraits dans les images, et les coordonnées calculées par le modèle de caméra pour chaque point correspondant de la mire. La minimisation de cette fonctionnelle passe par l'utilisation de méthodes numériques telles que celle de LEVENBERG-MARQUARDT [PTVF92] et nécessite alors une estimée initiale des paramètres à optimiser. Pour cela, de nombreuses méthodes analytiques directes existent, dont celles de RAVN *et al.* (décrite dans l'annexe B), de TSAI [Tsa87], ou de ZHANG [Zha98a].

Cette approche classique du calibrage a l'inconvénient majeur de considérer la géométrie de la mire connue avec exactitude, et donc réalisée sur un support très peu déformable (tel que

le verre) par un procédé de grande précision. Malgré ces précautions, la qualité du calibrage de la caméra est principalement tributaire de la précision de la mire. Pour s'affranchir de cette contrainte, certains auteurs ont proposé d'estimer simultanément les paramètres du modèle de caméra et les coordonnées tridimensionnelles des points de la mire. Nous allons étudier ce type de méthode dans la section suivante.

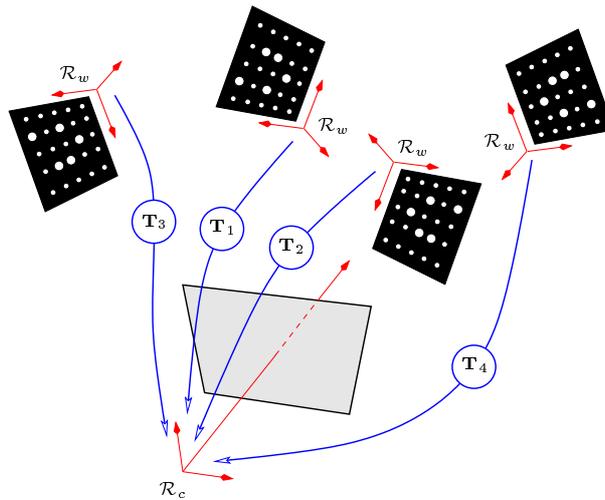
## 1.2.2 Méthode photogrammétrique

Afin d'être en mesure d'utiliser une mire de calibrage imprécise, il est possible d'utiliser une approche photogrammétrique comme dans [LVD99] pour estimer en même temps tous les paramètres du modèle de caméra ainsi que les points tridimensionnels de la mire.

Appelons  $m_i^j$  les coordonnées de la projection du  $j$ -ème point ( $j = 1 \dots p$ ) de la  $i$ -ème vue ( $i = 1 \dots n$ ) sur le plan image de la caméra. Nous prenons en compte la distorsion et pouvons écrire à partir de (1.14) :

$$m_i^j = f(\mathbf{k}, \mathbf{d}, \mathbf{R}_i, \mathbf{t}_i, M_j) \quad (1.15)$$

Dans ces équations, le référentiel de la mire est utilisé pour chaque vue comme référentiel du monde, *cf.* figure 1.7.



**FIG. 1.7 :** *Calibrage d'une caméra en utilisant quatre positions de mire : il existe quatre transformations rigides  $\mathbf{T}_i$  reliant les quatre positions de mire au référentiel de la caméra.*

Énumérons les paramètres inconnus à estimer : quatre paramètres intrinsèques, sept paramètres de distorsion,  $6n$  paramètres extrinsèques (trois pour la rotation et trois autres pour la translation de chaque transformation rigide  $\mathbf{T}_i$ ) et  $3p$  coordonnées tridimensionnelles ; cela nous donne un total de  $11 + 6n + 3p$  inconnues. En utilisant (1.15), chaque projection d'un point tridimensionnel fournit deux équations. Nous avons donc  $2np$  équations et  $11 + 6n + 3p$  inconnues. Si  $n$  et  $p$  sont suffisamment grands (par exemple si  $n = 6$  et  $p = 64$ , il y a 768 équations pour 239 inconnues) nous pouvons estimer tous les paramètres en minimisant la somme des distances entre la projection du  $j$ -ème point de la  $i$ -ème vue dans l'image et le

point  $\hat{m}_i^j$  correspondant extrait dans l'image :

$$\theta = \arg \min_{\theta} \sum_{i=1}^n \sum_{j=1}^p \left\| \hat{m}_i^j - f(\mathbf{k}, \mathbf{d}, \mathbf{R}_i, \mathbf{t}_i, M_j) \right\|^2 \quad (1.16)$$

avec  $\theta = (\mathbf{k}, \mathbf{d}, \mathbf{R}_{1\dots n}, \mathbf{t}_{1\dots n}, M_{1\dots p})$ .

Minimiser (1.16) est un problème d'optimisation non linéaire que nous résolvons en utilisant l'algorithme de LEVENBERG-MARQUARDT [PTVF92] (*cf.* annexe F), avec les rotations  $\mathbf{R}_i$  exprimées sous la forme minimale donnée en annexe H. Pour pouvoir converger, l'algorithme de minimisation nécessite des estimées initiales des paramètres : nous utilisons la méthode analytique de RAVN *et al.*, décrite dans l'annexe B, pour obtenir les estimées des quatre paramètres intrinsèques du modèle sténopé et des paramètres extrinsèques. Il est aussi possible d'utiliser la méthode décrite dans [Zha98a] pour obtenir ces estimées initiales. Nous donnons des valeurs nulles aux paramètres de distorsions initiaux. Les coordonnées tridimensionnelles initiales des points de la mire sont celles de son modèle ayant servi à sa réalisation. Étant donné qu'elles seront réestimées, ces coordonnées n'ont pas besoin d'être connues avec précision, ce qui représente un avantage par rapport à la méthode décrite dans [Zha98a].

La minimisation de (1.16) conduit à une solution définie à une transformation de similarité près, *i.e.* un déplacement de corps rigide et un facteur d'échelle liés au positionnement non contraint du référentiel de l'objet par rapport aux points qui le composent. Afin de déterminer le facteur d'échelle, la distance mesurée dans l'espace entre deux points peut être utilisée. Nous préférons fixer deux points du modèle dans le processus de minimisation ce qui conduit à estimer six paramètres de moins et à obtenir une estimation du modèle sans facteur d'échelle. Fixer deux points introduit seulement six contraintes, la solution est alors définie à une rotation près autour de la droite joignant ces deux points. Nous introduisons une septième contrainte en fixant à  $z = 0$  la coordonnée d'un troisième point non colinéaire aux deux premiers (en supposant que le modèle de l'objet plan de calibrage est composé de points définis sur le plan  $z = 0$ ) — voir l'annexe C pour les détails de cet exposé.

### 1.2.3 Correction de la distorsion

Il est parfois nécessaire de connaître les coordonnées pixel idéales, *i.e.* non distordues, correspondantes à celles distordues. C'est par exemple le cas pour le critère d'évaluation de la qualité de calibrage présenté section suivante. Le modèle de distorsion n'est pas inversible, aussi il est nécessaire d'utiliser une méthode numérique pour estimer les coordonnées idéales d'un point distordu.

Posons  $\bar{\mathbf{m}} = (\bar{u} \ \bar{v})^\top$  le point image non distordu que l'on cherche à déterminer à partir de son homologue distordu  $\mathbf{m} = (u \ v)^\top$ . Nous pouvons utiliser par exemple la méthode de NEWTON comme suit. On exprime le point  $(u \ v)^\top$  dans son plan rétinien :

$$\begin{aligned} x_d &= \frac{u - u_0}{\alpha_u} \\ y_d &= \frac{v - v_0}{\alpha_v} \end{aligned}$$

Ensuite, on utilise comme estimée initiale de  $(x_r \ y_r)^\top$  les coordonnées rétinienne distordues  $(x_d \ y_d)^\top$  :

$$\mathbf{m}_r^{(0)} = \mathbf{m}_d$$

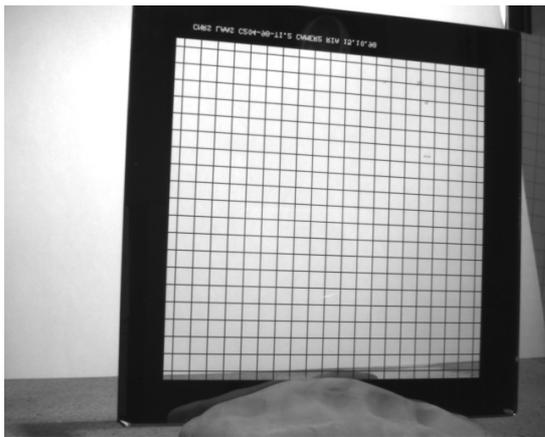
Puis, on itère la méthode de NEWTON :

$$\mathbf{m}_r^{(i)} = \mathbf{m}_r^{(i-1)} - \mathbf{J}^{-1} \underbrace{\left( \begin{matrix} (i-1) \\ \mathbf{m}_d - \mathbf{m}_d \end{matrix} \right)}_{\epsilon} \quad \text{avec} \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_d}{\partial x_r} & \frac{\partial x_d}{\partial y_r} \\ \frac{\partial y_d}{\partial x_r} & \frac{\partial y_d}{\partial y_r} \end{pmatrix} \Big|_{(x_r, y_r) = \mathbf{m}_r^{(i-1)}}$$

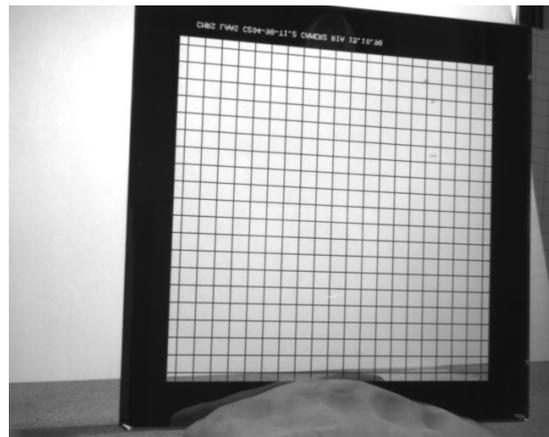
Le critère d'arrêt peut être basé sur la valeur du module de l'erreur  $\epsilon$  après chaque itération. En pratique, la convergence est atteinte après seulement quelques itérations ( $\approx 4$ ).

### 1.2.4 Évaluation de la qualité du calibrage

Évaluer la qualité d'un calibrage, c'est déterminer la qualité des paramètres estimés compte tenu d'un choix de modèle de caméra fait a priori et supposé conforme à la réalité physique. Différents critères ont été proposés pour évaluer la qualité du calibrage d'une caméra [WCH92, Gar96]. Nous avons choisi d'utiliser un critère issu du constat qu'avec une caméra obéissant au modèle sténopé (donc exempt de distorsion) une droite de l'espace de la scène se projette en une droite dans l'image [Bra95]. La caméra est d'abord calibrée en utilisant la procédure



(1) image originale

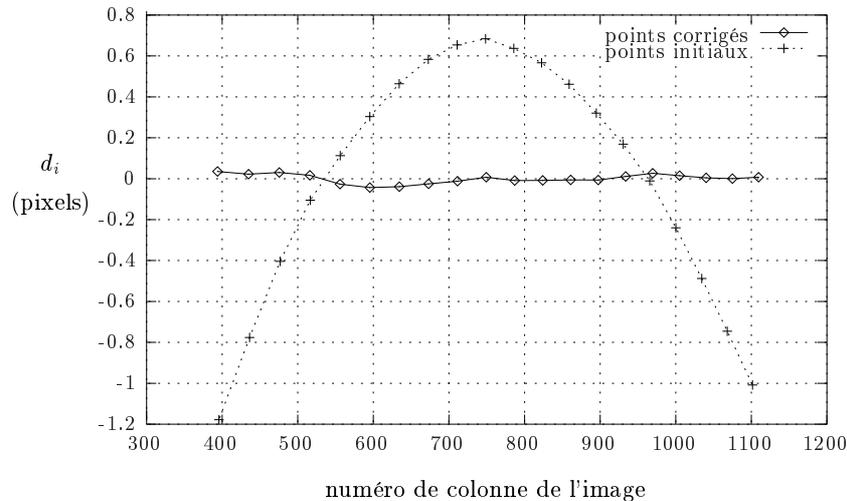


(2) image corrigée

**FIG. 1.8 :** (1) Image 1280 × 1024 originale acquise avec un objectif de focale 12,5 mm (la distorsion est peu visible à l'œil). (2) Image corrigée : les lignes de la mire sont des droites.

expérimentale décrite section 1.2. Nous faisons ensuite l'acquisition d'une image d'une mire plane constituée de lignes horizontales et verticales qui s'intersectent — cf. figure 1.8-1 : la distorsion est peu visible à l'œil mais atteint une amplitude maximale d'environ 20 pixels dans les bords d'image. Cette mire a été réalisée sur un support en verre par un procédé de microlithographie qui garantit sa géométrie avec une précision de l'ordre du micromètre. Les nœuds d'intersection de la mire sont extraits avec une précision de l'ordre de  $\frac{1}{50}$  pixel en utilisant l'extracteur de croix décrit dans [OGD97].

Les points-image correspondant aux nœuds extraits sont ensuite corrigés de la distorsion en utilisant les paramètres intrinsèques fournis par le calibrage. La qualité d'alignement des points corrigés est évaluée par régression linéaire : leurs distances algébriques orthogonales  $d_i$



**FIG. 1.9 :** Distances orthogonales à la droite de régression avant et après correction de la distorsion de points réellement alignés.

à cette droite de régression sont calculées, puis le critère suivant est évalué :

$$\text{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2}$$

Pour la dernière ligne de la mire (celle présentant le plus de distorsion) nous obtenons les résultats montrés figure 1.9. L'évaluation du critère avant correction est  $\text{rms} = 0,58$  ( $\approx \frac{1}{2}$  pixel), et devient après correction  $\text{rms} = 0,021$  ( $\approx \frac{1}{50}$  pixel). Cette amélioration est importante et on constate que la précision de calibrage ainsi évaluée se confond avec celle de notre « vérité terrain » :  $\approx \frac{1}{50}$  pixel chacune.

À titre d'exemple, nous avons aussi corrigé la distorsion sur l'image originale entière et obtenu celle montrée figure 1.8-2 en utilisant l'algorithme décrit dans [Ort91].

## 1.3 Modélisation d'un capteur de stéréovision

Nous nous intéressons dans cette section à la modélisation d'un capteur composé de deux caméras liées rigidement : un **capteur de vision stéréoscopique**. Nous commençons brièvement par justifier l'emploi d'une deuxième caméra dans le but de percevoir l'environnement en trois dimensions, puis nous énumérons les différents référentiels en présence et les transformations de passage de l'un à l'autre dans la section suivante. Nous terminons l'étude de la modélisation de ce capteur par un rappel de ses propriétés géométriques fondamentales qui nous seront utiles dans la suite de ce chapitre.

### 1.3.1 Pourquoi utiliser deux caméras ?

Si l'on se place d'un point de vue géométrique, une caméra est un capteur qui transforme tout « point visible » de l'espace tridimensionnel en point dans l'espace bidimensionnel de

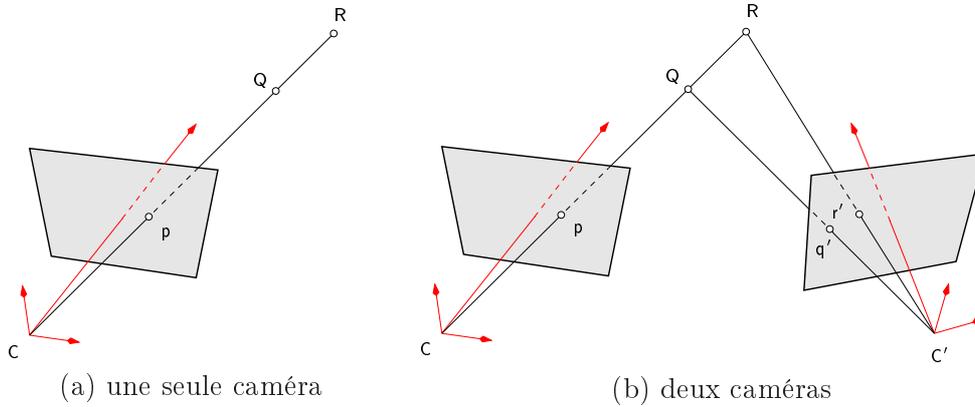


FIG. 1.10 : Retrouver la troisième dimension par l'emploi de deux caméras.

l'image. Cette transformation supprime donc la troisième dimension et est, par conséquent, irréversible. Cela se traduit graphiquement par la figure 1.10.a : les points Q et R de l'espace se projettent tous deux sur le plan image en un seul et même point  $p$  car ils sont sur la même droite projective  $(C, P)$ ,  $C$  est appelé le **centre de projection** ou **centre optique**. Ceci signifie qu'étant donné un point image  $p$ , il existe une infinité de points tridimensionnels pouvant en être la projection. En utilisant deux caméras comme montré figure 1.10.b, il est possible de déterminer la position tridimensionnelle du point par **triangulation**. Il existe en effet un seul point de l'espace correspondant à la paire de projetés  $\{p, q'\}$  et un seul correspondant à  $\{p, r'\}$ . La triangulation consiste donc à déterminer l'intersection dans l'espace des deux droites projectives. Par conséquent, il est nécessaire d'exprimer ces deux droites par rapport à un référentiel commun, par exemple celui de la caméra gauche. Pour y parvenir, nous allons chercher à exprimer une relation géométrique entre les deux caméras.

### 1.3.2 Référentiels et changements de repères

L'espace tridimensionnel de la scène est muni de son référentiel orthonormé  $\mathcal{R}_w$ . Chacune des deux caméras possède son propre référentiel orthonormé : nous les appellerons repère caméra gauche  $\mathcal{R}_c$  et repère caméra droite  $\mathcal{R}'_c$ . La figure 1.11 illustre ces trois référentiels ainsi que les transformations rigides permettant d'exprimer un point dans un autre référentiel. Avec ces conventions, nous pouvons écrire les relations suivantes :

$$\begin{aligned}\tilde{\mathbf{M}}_c &\cong \mathbf{T} \tilde{\mathbf{M}} \\ \tilde{\mathbf{M}}'_c &\cong \mathbf{T}' \tilde{\mathbf{M}} \\ \tilde{\mathbf{M}}'_c &\cong \mathbf{T}_s \tilde{\mathbf{M}}_c\end{aligned}\tag{1.17}$$

Ces équations nous montrent que les trois transformations ne sont pas indépendantes puisque nous pouvons déterminer l'une d'entre elles en fonction des deux autres :

$$\mathbf{T} \cong \mathbf{T}_s^{-1} \mathbf{T}'\tag{1.18}$$

$$\mathbf{T}' \cong \mathbf{T}_s \mathbf{T}\tag{1.19}$$

$$\mathbf{T}_s \cong \mathbf{T}' \mathbf{T}^{-1}\tag{1.20}$$

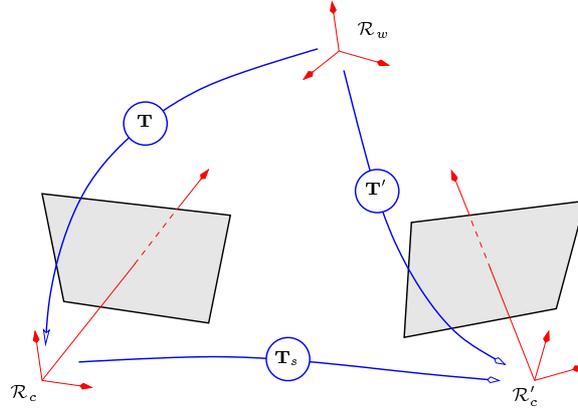


FIG. 1.11 : Les trois référentiels tridimensionnels du capteur de stéréovision.

Lorsqu'un point  $M$  de la scène est visible simultanément par les deux caméras, ces dernières nous donnent deux points :  $m$  pour la caméra gauche, et  $m'$  pour celle de droite. En utilisant le modèle géométrique de caméra et la relation de dépendance entre les trois repères  $\mathcal{R}_w$ ,  $\mathcal{R}_c$  et  $\mathcal{R}'_c$ , nous pouvons écrire les équations de  $m$  et  $m'$  en fonction de  $M$  :

$$\tilde{m} \cong \mathbf{K}\mathbf{T}\tilde{M} \quad (1.21)$$

$$\tilde{m}' \cong \mathbf{K}'\mathbf{T}'\tilde{M} \cong \mathbf{K}'\mathbf{T}_s\mathbf{T}\tilde{M} \quad (1.22)$$

### 1.3.3 Géométrie du capteur de stéréovision

Nous rappelons dans cette section certaines propriétés importantes liées à la géométrie d'un capteur de stéréovision. Elles nous seront utiles pour établir un critère de comparaison des différentes méthodes de calibrage que nous proposerons. Nous invitons le lecteur à se référer à [Fau93] pour une étude détaillée de ces résultats remarquables. Sans perte de généralité, nous exposons ces derniers pour le cas de caméras sans distorsion. Il est en effet toujours possible de se ramener au cas linéaire en corrigeant la distorsion par une méthode telle que celle décrite section 1.2.3.

Cherchons à établir une relation simple entre un point  $m_r$  du plan rétinien normalisé gauche et son homologue  $m'_r$  sur le plan rétinien normalisé droit. Nous savons que le projeté  $m'_r = (x'_r \ y'_r)^\top$  sur le plan rétinien droit du point  $M'_c = (x'_c \ y'_c \ z'_c)^\top$  a pour expression :

$$m'_r = \begin{pmatrix} x'_r \\ y'_r \end{pmatrix} = \frac{1}{z'_c} \begin{pmatrix} x'_c \\ y'_c \end{pmatrix} \quad (1.23)$$

Si l'on utilise la notation  $\mathbf{r}_{s_i}^\top$  pour désigner le vecteur ligne correspondant à la  $i$ -ième ligne de la rotation  $\mathbf{R}_s$  de la transformation  $\mathbf{T}_s$ , nous pouvons utiliser les équations (1.17) et (1.23) pour exprimer les coordonnées dans le plan rétinien droit d'un point de l'espace défini dans le

référentiel  $\mathcal{R}_c$  de la caméra gauche :

$$\begin{aligned} x'_r &= \frac{x'_c}{z'_c} = \frac{\mathbf{r}_{s_1}^\top \mathbf{M}_c + t_{s_x}}{\mathbf{r}_{s_3}^\top \mathbf{M}_c + t_{s_z}} \\ y'_r &= \frac{y'_c}{z'_c} = \frac{\mathbf{r}_{s_2}^\top \mathbf{M}_c + t_{s_y}}{\mathbf{r}_{s_3}^\top \mathbf{M}_c + t_{s_z}} \end{aligned}$$

En utilisant l'équation  $\tilde{\mathbf{m}}_r \cong \mathbf{P}\tilde{\mathbf{M}}_C$  nous pouvons réécrire les relations précédentes sous la forme :

$$x'_r = \frac{z_c \mathbf{r}_{s_1}^\top \tilde{\mathbf{m}}_r + t_{s_x}}{z_c \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r + t_{s_z}} \quad (1.24)$$

$$y'_r = \frac{z_c \mathbf{r}_{s_2}^\top \tilde{\mathbf{m}}_r + t_{s_y}}{z_c \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r + t_{s_z}} \quad (1.25)$$

Ces deux équations montrent qu'il est possible d'exprimer les coordonnées d'un point de l'image de droite en fonction de son **stéréo-correspondant** dans l'image gauche, des paramètres intrinsèques et extrinsèques du capteur stéréoscopique et de la « profondeur » du point  $\mathbf{M}$  dans le référentiel de la caméra gauche.

Nous allons chercher l'équation du lieu de l'homologue droit d'un point du plan rétinien gauche. Pour cela, nous allons commencer par éliminer  $z_c$  des équations (1.24) et (1.25) :

$$\begin{aligned} x'_r = \frac{z_c \mathbf{r}_{s_1}^\top \tilde{\mathbf{m}}_r + t_{s_x}}{z_c \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r + t_{s_z}} &\Rightarrow z_c = \frac{x'_r t_{s_z} - t_{s_x}}{\mathbf{r}_{s_1}^\top \tilde{\mathbf{m}}_r - \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r x'_r} \\ y'_r = \frac{z_c \mathbf{r}_{s_2}^\top \tilde{\mathbf{m}}_r + t_{s_y}}{z_c \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r + t_{s_z}} &\Rightarrow z_c = \frac{y'_r t_{s_z} - t_{s_y}}{\mathbf{r}_{s_2}^\top \tilde{\mathbf{m}}_r - \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r y'_r} \end{aligned}$$

De ces deux équations nous déduisons l'expression suivante reliant  $\tilde{\mathbf{m}}'_r$  à  $\tilde{\mathbf{m}}_r$  :

$$\left( t_{s_z} \mathbf{r}_{s_2}^\top \tilde{\mathbf{m}}_r - t_{s_y} \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r \right) x'_r + \left( t_{s_x} \mathbf{r}_{s_3}^\top \tilde{\mathbf{m}}_r - t_{s_z} \mathbf{r}_{s_1}^\top \tilde{\mathbf{m}}_r \right) y'_r = t_{s_x} \mathbf{r}_{s_2}^\top \tilde{\mathbf{m}}_r - t_{s_y} \mathbf{r}_{s_1}^\top \tilde{\mathbf{m}}_r \quad (1.26)$$

Cette équation, de la forme  $ax'_r + by'_r + c = 0$ , est le lieu des projetés dans le plan rétinien droit pouvant correspondre à un projeté  $\mathbf{m}_r$  dans le plan rétinien gauche. C'est l'équation d'une droite appelée **ligne épipolaire droite**, cf. figure 1.12. Réciproquement nous pourrions déterminer l'équation de la ligne épipolaire gauche pour un point  $\mathbf{m}'_r$  donné.

Si le point  $\mathbf{m}_r$  décrit le plan rétinien gauche, alors les lignes épipolaires droites forment un faisceau dont le point d'intersection de toutes ces droites s'appelle l'**épipole droit**. Ce point est aussi le projeté du centre optique de la caméra gauche sur le plan rétinien de la caméra droite. Nous pouvons déterminer facilement ses coordonnées en prenant  $z_c = 0$  dans les équations (1.24) et (1.25) :

$$\begin{aligned} x'_e &= t_{s_x}/t_{s_z} \\ y'_e &= t_{s_y}/t_{s_z} \end{aligned}$$

D'une façon analogue, nous pourrions définir un **épipole gauche**, projeté du centre optique de la caméra droite sur le plan rétinien de la caméra gauche.

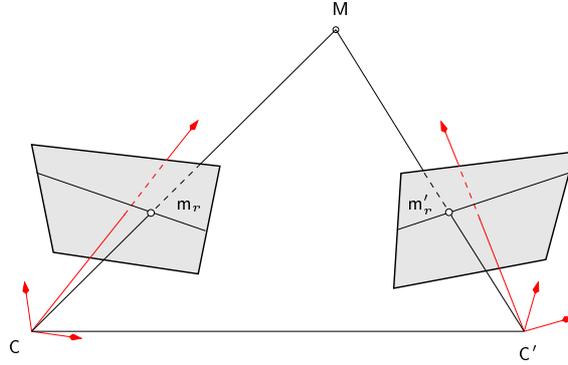


FIG. 1.12 : La géométrie épipolaire.

### Matrice essentielle

L'équation (1.26) peut s'écrire sous la forme :

$$(x'_r \quad y'_r \quad 1) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \quad (1.27)$$

avec  $a$ ,  $b$  et  $c$  fonctions de la transformation  $\mathbf{T}_s$  et du point  $m_r$ . Nous pouvons exprimer ces trois coefficients comme suit :

$$\begin{aligned} a &= (t_{s_y} r_{s31} - t_{s_z} r_{s21})x_r + (t_{s_y} r_{s32} - t_{s_z} r_{s22})y_r + (t_{s_y} r_{s33} - t_{s_z} r_{s23}) \\ b &= (t_{s_z} r_{s11} - t_{s_x} r_{s31})x_r + (t_{s_z} r_{s12} - t_{s_x} r_{s32})y_r + (t_{s_z} r_{s13} - t_{s_x} r_{s33}) \\ c &= (t_{s_x} r_{s21} - t_{s_y} r_{s11})x_r + (t_{s_x} r_{s22} - t_{s_y} r_{s12})y_r + (t_{s_x} r_{s23} - t_{s_y} r_{s13}) \end{aligned}$$

Ce qui peut se mettre sous la forme matricielle :

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -t_{s_z} & t_{s_y} \\ t_{s_z} & 0 & -t_{s_x} \\ -t_{s_y} & t_{s_x} & 0 \end{pmatrix}}_{(\mathbf{t}_s)_\times} \underbrace{\begin{pmatrix} r_{s11} & r_{s12} & r_{s13} \\ r_{s21} & r_{s22} & r_{s23} \\ r_{s31} & r_{s32} & r_{s33} \end{pmatrix}}_{\mathbf{R}_s} \begin{pmatrix} x_r \\ y_r \\ 1 \end{pmatrix} \quad (1.28)$$

La matrice  $(\mathbf{t}_s)_\times$  est une matrice antisymétrique de rang deux, et  $\mathbf{E} = (\mathbf{t}_s)_\times \mathbf{R}_s$ . En utilisant l'équation précédente, la droite épipolaire (1.27) peut se mettre sous la forme suivante :

$$\tilde{m}'_r{}^\top \mathbf{E} \tilde{m}_r = 0 \quad (1.29)$$

La matrice  $\mathbf{E}$  est appelée **matrice essentielle**, elle décrit la transformation épipolaire gauche-droite. On peut montrer facilement que la transformation épipolaire droite-gauche s'exprime par la transposée de la matrice essentielle précédente [LH81, Har93, HM95] :

$$\tilde{m}_r{}^\top \mathbf{E}^\top \tilde{m}'_r = 0$$

### Matrice fondamentale

Nous avons établi, avec la matrice essentielle, l'expression de la ligne épipolaire droite correspondante à un point du plan rétinien gauche. Nous pouvons faire de même en nous plaçant dans le repère image en utilisant l'équation (1.29) et les relations :  $\tilde{\mathbf{m}} = \mathbf{A}\tilde{\mathbf{m}}_r$  et  $\tilde{\mathbf{m}}' = \mathbf{A}'\tilde{\mathbf{m}}'_r$ . En utilisant la notation  $(\mathbf{A}^{-1})^\top \equiv \mathbf{A}^{-\top}$ , on obtient alors :

$$\tilde{\mathbf{m}}'^\top \underbrace{\mathbf{A}'^{-\top} \mathbf{E} \mathbf{A}^{-1}}_{\mathbf{F}} \tilde{\mathbf{m}} = 0 \quad (1.30)$$

La matrice  $\mathbf{F} = \mathbf{A}'^{-\top} \mathbf{E} \mathbf{A}^{-1}$  est appelée **matrice fondamentale** [Luo92, LDFP93, Har95, HM95].

Étant donné un couple de points  $\{\mathbf{m}, \mathbf{m}'\}$ , il est possible de calculer la distance algébrique  $d$  entre le point  $\mathbf{m}'$  et la droite épipolaire associée à  $\mathbf{m}$  [Zha98b] :

$$d = \frac{\tilde{\mathbf{m}}'^\top \mathbf{F} \tilde{\mathbf{m}}}{\sqrt{\tilde{\mathbf{m}}'^\top \mathbf{F}^\top \mathbf{Z} \mathbf{Z}^\top \mathbf{F} \tilde{\mathbf{m}}}} \quad \text{avec} \quad \mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (1.31)$$

De la même façon, nous pouvons calculer la distance  $d'$  entre le point  $\mathbf{m}$  et la droite épipolaire associée à  $\mathbf{m}'$  :

$$d' = \frac{\tilde{\mathbf{m}}'^\top \mathbf{F} \tilde{\mathbf{m}}}{\sqrt{\tilde{\mathbf{m}}'^\top \mathbf{F}^\top \mathbf{Z} \mathbf{Z}^\top \mathbf{F} \tilde{\mathbf{m}}'}} \quad (1.32)$$

### Triangulation

Les équations (1.21) et (1.22) expriment les coordonnées des deux points  $\mathbf{m} = (u \ v)^\top$  et  $\mathbf{m}' = (u' \ v')^\top$ , respectivement les projections gauche et droite d'un point  $\mathbf{M}$  de la scène. En posant :

$$\mathbf{T} \cong \begin{pmatrix} \mathbf{r}_1^\top & t_x \\ \mathbf{r}_2^\top & t_y \\ \mathbf{r}_3^\top & t_z \\ \mathbf{0}^\top & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{T}' \cong \begin{pmatrix} \mathbf{r}'_1{}^\top & t'_x \\ \mathbf{r}'_2{}^\top & t'_y \\ \mathbf{r}'_3{}^\top & t'_z \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

nous pouvons écrire le système de quatre équations suivant :

$$\begin{aligned} u &= \alpha_u \frac{\mathbf{r}_1^\top \mathbf{M} + t_x}{\mathbf{r}_3^\top \mathbf{M} + t_z} + u_0 & v &= \alpha_v \frac{\mathbf{r}_2^\top \mathbf{M} + t_y}{\mathbf{r}_3^\top \mathbf{M} + t_z} + v_0 \\ u' &= \alpha'_u \frac{\mathbf{r}'_1{}^\top \mathbf{M} + t'_x}{\mathbf{r}'_3{}^\top \mathbf{M} + t'_z} + u'_0 & v' &= \alpha'_v \frac{\mathbf{r}'_2{}^\top \mathbf{M} + t'_y}{\mathbf{r}'_3{}^\top \mathbf{M} + t'_z} + v'_0 \end{aligned}$$

Si l'on connaît les deux points image  $\mathbf{m}$  et  $\mathbf{m}'$  et que le capteur stéréoscopique est calibré<sup>5</sup>, alors nous pouvons déterminer les trois inconnues qui sont les coordonnées tridimensionnelles du point  $\mathbf{M}$  en résolvant ce système surdéterminé de quatre équations [HS97].

<sup>5</sup>C'est-à-dire que l'on connaît les valeurs de  $\mathbf{R}$ ,  $\mathbf{R}'$ ,  $\mathbf{t}$ ,  $\mathbf{t}'$ ,  $\alpha_u$ ,  $\alpha'_u$ ,  $\alpha_v$ ,  $\alpha'_v$ ,  $u_0$ ,  $u'_0$ ,  $v_0$  et  $v'_0$ .

## 1.4 Calibrage d'un capteur de stéréovision

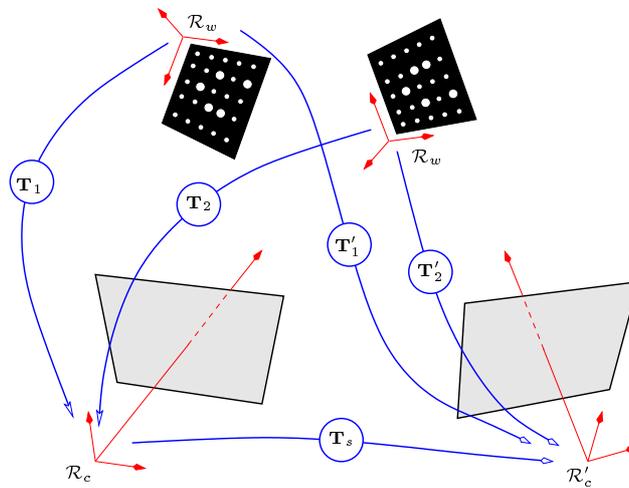
Lorsqu'on calibre une seule caméra, on s'intéresse aux paramètres intrinsèques définis par la matrice  $\mathbf{K}$  et aux paramètres extrinsèques définis par la transformation rigide  $\mathbf{T}$ . Lorsqu'on calibre un capteur de vision stéréoscopique, nous nous intéressons aux deux ensembles de paramètres intrinsèques définis par les matrices  $\mathbf{K}$  et  $\mathbf{K}'$  et à la position et l'orientation relative des deux caméras définies par la transformation rigide  $\mathbf{T}_s$ . Le calibrage de ce capteur a pour but de pouvoir reconstruire la position tridimensionnelle de points observés par les deux caméras et est donc très importante pour ceux qui, comme nous, s'inscrivent dans un contexte applicatif de mesures dimensionnelles de précision.

Nous noterons, respectivement pour la caméra gauche et la caméra droite, les transformations rigides  $\mathbf{T}_i$  et  $\mathbf{T}'_i$  de la façon suivante :

$$\mathbf{T}_i = \begin{pmatrix} \mathbf{R}_i & \mathbf{t}_i \\ \mathbf{0}^\top & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{T}'_i = \begin{pmatrix} \mathbf{R}'_i & \mathbf{t}'_i \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

Elles relient la  $i$ -ième vue de la mire respectivement au référentiel de la caméra gauche et à celui de la caméra droite. Pour chaque position de mire, nous avons d'après (1.20) la relation (cf. figure 1.13) :

$$\mathbf{T}_s \mathbf{T}_i = \mathbf{T}'_i$$



**FIG. 1.13 :** Calibrage d'un capteur de stéréovision avec deux positions de mire : il existe deux transformations rigides pour chaque caméra reliant les deux positions de mire au référentiel des caméras, et une seule transformation reliant les référentiels des deux caméras.

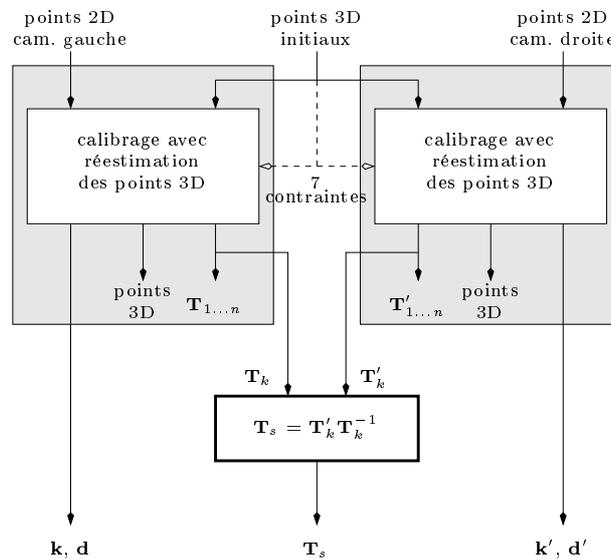
Nous allons voir différentes méthodes permettant d'estimer la transformation  $\mathbf{T}_s$  : deux méthodes linéaires puis une nouvelle méthode globale estimant simultanément cette transformation, les paramètres intrinsèques des deux caméras et les coordonnées tridimensionnelles des points de la mire de calibrage.

### 1.4.1 Méthodes LIN1 et LINn

La première approche consiste à calibrer chaque caméra indépendamment, en utilisant la méthode décrite section 1.2, pour déterminer les paramètres intrinsèques et les coefficients de distorsion des deux caméras, et les deux ensembles  $\{\mathbf{T}_i\}$  et  $\{\mathbf{T}'_i\}$  de matrices des paramètres extrinsèques.  $\mathbf{R}_s$  et  $\mathbf{t}_s$  peuvent alors être calculés en utilisant n'importe quelle paire  $k \in \{1 \dots n\}$  de matrices des paramètres extrinsèques en utilisant l'équation (1.20) :

$$\mathbf{T}_s = \mathbf{T}'_k \mathbf{T}_k^{-1} \quad (1.33)$$

Cette méthode est celle classiquement utilisée par de nombreux chercheurs, nous l'appellerons **LIN1**, cf. figure 1.14.



**FIG. 1.14 :** Méthode **LIN1** : la transformation gauche-droite  $\mathbf{T}_s$  est calculée en utilisant une paire de matrices des extrinsèques (ici la paire numéro  $k$ ) choisie parmi  $n$ .

Le choix de la paire de matrices des paramètres extrinsèques  $\mathbf{T}_k$  et  $\mathbf{T}'_k$  dans la méthode **LIN1** est délicat et au moins trois heuristiques sont possibles :

- choisir arbitrairement toujours la  $k$ -ième paire de matrices de l'expérimentation, par exemple la première paire  $\mathbf{T}_1$  et  $\mathbf{T}'_1$  ;
- prendre la paire de matrices qui correspond à la plus faible erreur globale de reprojection des points de la mire dans les deux images ;
- les choisir de telle sorte que la somme des distances des points à leur droite épipolaire soit minimale en utilisant les équations (1.31) et (1.32).

Ces heuristiques ont toutes l'inconvénient de ne pas utiliser la redondance fournie par l'utilisation simultanée de toutes les paires de matrices des paramètres extrinsèques pour estimer la transformation  $\mathbf{T}_s$ . On peut penser que cela permettrait de réduire l'erreur commise sur cette estimation. Nous avons donc cherché à estimer  $\mathbf{R}_s$  et  $\mathbf{t}_s$  en utilisant toutes les matrices des extrinsèques  $\{\mathbf{T}_i\}$  et  $\{\mathbf{T}'_i\}$ . Si nous considérons un point  $M$  de la mire, nous pouvons calculer dans la  $i$ -ième vue ses coordonnées  $M_{c_i}$  et  $M'_{c_i}$  dans les deux référentiels caméra en utilisant

les relations :

$$\begin{aligned}\tilde{M}_{c_i} &\cong \mathbf{T}_i \tilde{M} \\ \tilde{M}'_{c_i} &\cong \mathbf{T}'_i \tilde{M}\end{aligned}$$

Pour tous les  $p$  points de la mire, et en utilisant l'équation (1.17), nous pouvons calculer la transformation  $\mathbf{T}_s$  optimale qui transforme l'ensemble de points  $\{M_{c_i}^j\}$  en l'ensemble  $\{M'_{c_i}^j\}$  en minimisant la fonctionnelle (cf. figure 1.15) :

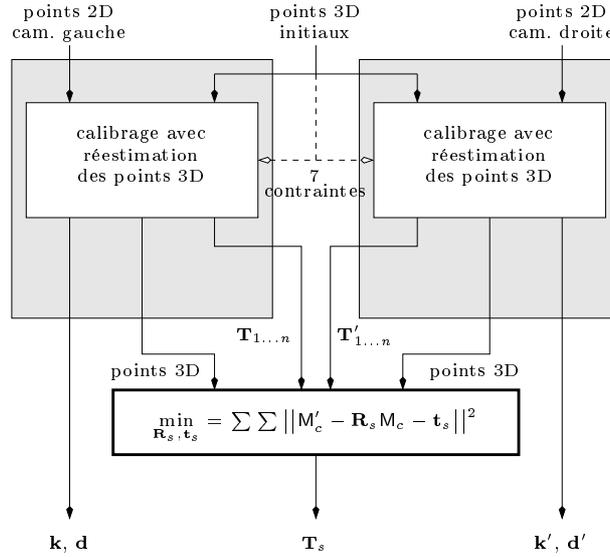


FIG. 1.15 : Méthode **LINn** : la transformation gauche-droite  $\mathbf{T}$  est calculée en utilisant toutes les paires de matrices des extrinsèques.

$$\min_{\mathbf{R}_s, \mathbf{t}_s} \sum_{i=1}^n \sum_{j=1}^p \left\| M'_{c_i}{}^j - (\mathbf{R}_s M_{c_i}^j + \mathbf{t}_s) \right\|^2 \quad (1.34)$$

Ce problème est bien connu et revient à estimer la meilleure transformation rigide entre deux nuages de points tridimensionnels. Plusieurs méthodes existent [LEF97], LORUSSO *et al.* ont comparé quatre algorithmes, permettant de déterminer cette transformation, basés sur :

- une décomposition en valeurs singulières (SVD) ;
- l'utilisation de matrices orthonormales ;
- la représentation de la rotation sous forme de quaternions unitaires ;
- l'utilisation de quaternions duaux.

Les auteurs montrent que ces quatre méthodes donnent des résultats similaires ne permettant pas de dégager un algorithme plus performant que les autres<sup>6</sup>. Nous avons utilisé la solution reposant sur la décomposition en valeurs singulières d'une matrice de corrélation pour sa simplicité de mise en œuvre. Nous nommerons cette deuxième méthode **LINn**.

<sup>6</sup>Les auteurs indiquent que le quatrième algorithme — utilisation de quaternions duaux pour représenter la rotation — est toutefois sensiblement moins stable numériquement que les trois autres.

### 1.4.2 Méthode GLOB

Dans les méthodes **LIN**,  $\mathbf{R}_s$  et  $\mathbf{t}_s$  sont calculés en utilisant les deux ensembles  $\{\mathbf{T}_i\}$  et  $\{\mathbf{T}'_i\}$  des matrices des paramètres extrinsèques. En utilisant l'équation (1.19),  $\mathbf{R}_s$  et  $\mathbf{t}_s$  peuvent être directement calculés en minimisant la fonctionnelle suivante :

$$\theta = \arg \min_{\theta} \sum_{i=1}^n \sum_{j=1}^p [\delta(i, j) + \delta'(i, j)] \quad (1.35)$$

avec :

$$\begin{aligned} \theta &= (\mathbf{k}, \mathbf{k}', \mathbf{d}, \mathbf{d}', \mathbf{R}_{1\dots n}, \mathbf{t}_{1\dots n}, \mathbf{R}_s, \mathbf{t}_s, M_{1\dots p}) \\ \delta(i, j) &= \left\| \hat{\mathbf{m}}_i^j - f(\mathbf{k}, \mathbf{d}, \mathbf{R}_i, \mathbf{t}_i, M_j) \right\|^2 \\ \delta'(i, j) &= \left\| \hat{\mathbf{m}}_i^j - g(\mathbf{k}', \mathbf{d}', \mathbf{R}_i, \mathbf{t}_i, \mathbf{R}_s, \mathbf{t}_s, M_j) \right\|^2 \end{aligned}$$

La fonction  $g$  est similaire à la fonction  $f$  à la différence près qu'elle détermine la matrice des paramètres extrinsèques  $\mathbf{T}'_i$  en utilisant l'équation (1.19) avec  $\mathbf{R}_i$ ,  $\mathbf{t}_i$ ,  $\mathbf{R}_s$  et  $\mathbf{t}_s$ .

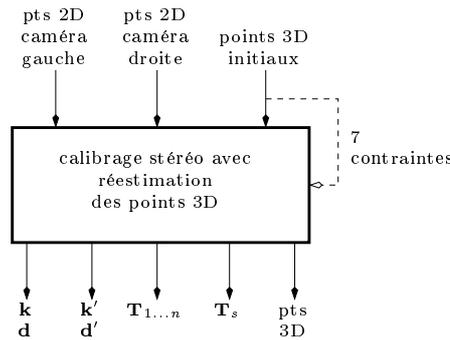


FIG. 1.16 : Méthode **GLOB**.

Nous avons aussi utilisé l'algorithme de LEVENBERG-MARQUARDT pour résoudre ce problème non linéaire. L'estimée initiale de la transformation  $\mathbf{T}_s$  est calculée par l'équation (1.34). Nous appellerons cette méthode **GLOB** — cf. figure 1.16.

### 1.4.3 Comparaison des différentes approches

Les trois méthodes proposées ont été testées sur des données de simulation. Travailler avec des données de synthèse permet une évaluation approfondie des différentes méthodes en comparant les paramètres estimés à la vérité.

#### Contexte expérimental

Pour évaluer l'influence du bruit dans les points image et dans les points tridimensionnels de la mire, nous avons simulé un capteur de stéréovision observant  $n = 9$  vues d'une grille régulière faite de  $p = 48$  points, espacés de cinq unités de longueur (en pratique le mm),

cf. figure 1.17. Le choix des données de synthèse provient d'un vrai calibrage : nous avons toutefois volontairement arrondi les paramètres ainsi calculés de sorte qu'il soit plus aisé de les présenter. Les paramètres intrinsèques et les coefficients de distorsion du couple de caméras sont indiqués table 1.1. Pour chaque vue, l'orientation et la position de la mire par rapport à la caméra gauche sont données table 1.2. Les angles  $y$  sont exprimés en degrés et correspondent à la séquence de CARDAN-BRYANT : rotations successives autour de l'axe  $x$  d'un angle  $\theta_x$ , puis autour du nouvel axe  $y'$  d'un angle  $\theta_y$ , et enfin autour du nouvel axe  $z''$  d'un angle  $\theta_z$  :

$$\mathbf{R} = \mathbf{R}_z(\theta_z) \cdot \mathbf{R}_y(\theta_y) \cdot \mathbf{R}_x(\theta_x)$$

Les positions données par les vecteurs de translation sont exprimées en unités de longueur. La rotation et la translation reliant le référentiel de la caméra droite à celui de la caméra

caméra	$u_0$	$v_0$	$\alpha_u$	$\alpha_v$	$r_1$	$r_2$	$r_3$
gauche	530	610	3800	-3800	-2	100	-1000
droite	520	630	3850	-3850	-2	100	-1000

**TAB. 1.1 :** Paramètres intrinsèques des deux caméras du capteur de stéréovision synthétique.

vue	rotation en °			translation		
	$\theta_x$	$\theta_y$	$\theta_z$	$t_x$	$t_y$	$t_z$
1	-4	-15	-52	-23	2	185
2	-19	-31	-47	-23	-5	183
3	12	7	-52	-20	7	188
4	-22	-1	-55	-24	11	209
5	6	-16	-91	-20	1	182
6	5	-13	-88	-20	25	180
7	-13	-5	4	-9	-30	194
8	-22	-27	14	-6	-12	184
9	-15	7	45	2	-13	215

**TAB. 1.2 :** Positions de la mire de calibrage par rapport à la caméra gauche du capteur de stéréovision synthétique.

gauche sont données table 1.3.

rotation en °			translation		
$\theta_{s_x}$	$\theta_{s_y}$	$\theta_{s_z}$	$t_{s_x}$	$t_{s_y}$	$t_{s_z}$
31	0	0	1	100	20

**TAB. 1.3 :** Position de la caméra droite par rapport à la caméra gauche du capteur de stéréovision synthétique.

Comment évaluer les performances des différentes méthodes? Tout d'abord, nous allons utiliser les vrais paramètres pour déterminer sur une série de calibrages :

- le biais des estimations ;

– l'écart-type de la dispersion des paramètres estimés.

Ensuite, nous utilisons deux critères de comparaison. Un capteur de stéréovision obéit à la contrainte épipolaire exposée section 1.3.3. Cette contrainte va être utilisée pour évaluer la qualité d'un calibrage stéréoscopique en calculant la distance d'un point de l'image droite, respectivement gauche, à sa droite épipolaire par l'équation (1.31), respectivement (1.32), nous l'appellerons **critère épipolaire** par la suite. Enfin, en utilisant un ensemble de points tridimensionnels aléatoirement répartis dans un volume donné à l'intérieur du champ de vue des deux caméras, nous avons calculé la distance euclidienne entre les vrais points 3D et ceux reconstruits par triangulation (*cf.* section 1.3.3), ce sera le **critère de reconstruction 3D**.

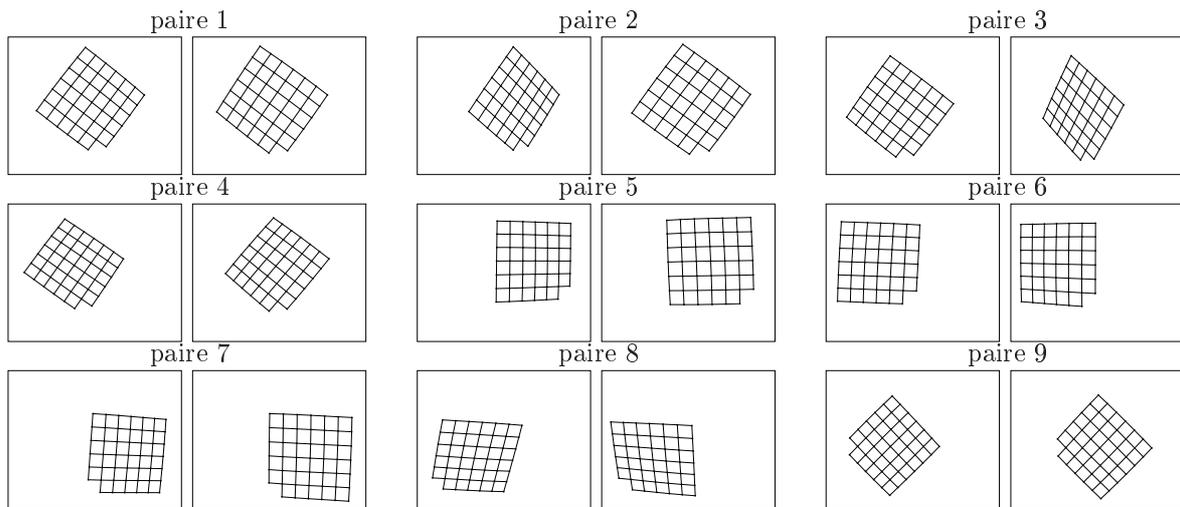


FIG. 1.17 : Vues utilisées pour le calibrage du capteur de stéréovision.

### Bruit dans les points image

Afin d'évaluer l'influence du bruit dans les points image, projections des points tridimensionnels de la mire, sur l'estimation des paramètres du modèle du capteur de stéréovision, nous avons contaminé ces points avec un bruit gaussien de moyenne nulle et d'écart-type  $\sigma$ . En pratique, le niveau de bruit de notre méthode d'extraction de points dans l'image est d'environ 0,05 pixel. Nous avons procédé à 200 calibrages pour chaque niveau de bruit  $\sigma$ . Comme nous nous y attendions, nous avons pu vérifier que les estimations sont non biaisées ; le seul paramètre statistique pertinent est donc l'écart-type de leur distribution, *cf.* figures 1.18 et 1.19.

Nous pouvons remarquer sur ces figures que notre méthode de calibrage **GLOB** fournit une estimation des paramètres affines  $\alpha_u$  et  $\alpha_v$  bien meilleure que les méthodes **LIN**. L'estimation du centre optique et des coefficients de distorsion n'est pas vraiment améliorée par **GLOB**.

Nous avons aussi évalué les performances de chacune des méthodes en utilisant les critères épipolaire et de reconstruction 3D. Les résultats sont montrés figure 1.20. Nous pouvons remarquer que notre méthode globale améliore significativement la précision de reconstruction tridimensionnelle par triangulation, ainsi que la qualité de la contrainte épipolaire, par rapport à la méthode **LIN1** classiquement employée. Nous pouvons aussi y voir que **LIN $n$**  améliore la contrainte épipolaire par rapport à **LIN1**.

Pour la méthode **GLOB**, sous l'hypothèse d'une précision d'extraction des points images de

$\frac{1}{20}$  pixel (facilement réalisable), la figure 1.20 nous montre que la précision de reconstruction tridimensionnelle est d'environ  $1,5 \mu\text{m}$ , ce qui conduit à une précision relative (*i.e.* la précision de détermination des coordonnées divisée par la taille de l'objet) d'environ  $\frac{1.5 \times 10^{-3}}{50\sqrt{2}} \approx \frac{1}{50\,000}$ , pour des points tridimensionnels distants de 165 mm à 215 mm des caméras.

### Bruit dans les points de la mire

Nous avons aussi cherché à déterminer l'influence du bruit dans les points tridimensionnels du modèle de mire. Nous avons procédé à de nouvelles séries de 200 calibrages chacune avec un bruit dans les points image de 0, 10 pixel, et un bruit sur les points du modèle de mire allant de 0, 10 à 1 unité de longueur. Nous avons pu constater que le bruit sur les points de la mire n'ont pas d'influence sur l'estimation des paramètres. Ce résultat remarquable a déjà été constaté par LAVEST dans [LVD99]. Ainsi, le modèle de la mire n'a pas besoin d'être connu précisément, ce qui est d'une grande souplesse pour le calibrage.

### Influence du nombre de points de la mire

Nous avons procédé à de nouvelles séries de 200 calibrages chacune en utilisant des grilles faites de 8, 15, 24, 35, 48 et 63 points, dans le but d'évaluer les performances de chacune des méthodes. Pour se placer dans des conditions d'expérimentation similaires entre chaque mire, nous avons adapté l'espacement entre chaque point de la grille, de sorte que la taille de celle-ci soit toujours la même (*cf.* figure 1.21).

Nous pouvons observer sur les figures 1.22 et 1.23 des résultats que l'utilisation de plus de 35 points pour la mire de calibrage n'améliore pas de manière significative les performances des différentes méthodes. Nous avons déjà remarqué cela lors de calibrages réels. Notons que la méthode **GLOB** donne les meilleurs résultats du point de vue des deux critères que toutes les autres méthodes, même en utilisant une mire de seulement huit points.

### Comparaison avec un calibrage réel

Notre capteur de vision stéréoscopique est composé de deux caméras donnant des images de dimension  $1280 \times 1024$ . Nous avons acquis 19 paires d'images d'une mire de 64 points : seulement neuf paires ont servi au calibrage de notre capteur (*cf.* figure 1.25), les dix autres nous permettant d'évaluer le critère épipolaire. Le résultat de l'estimation des paramètres intrinsèques est donné table 1.4. Les paramètres de la transformation rigide  $\mathbf{T}_s$  estimée sont donnés table 1.5. L'évaluation du critère épipolaire a été effectuée sur les  $10 \times 64 = 640$

méthode	caméra	$u_0$	$v_0$	$\alpha_u$	$\alpha_v$	$r_1$	$r_2$	$r_3$
<b>LIN</b>	gauche	509, 23	646, 86	3797, 55	-3798, 20	-0, 34677	0, 2727	-11, 780
	droite	509, 13	639, 84	3801, 83	-3803, 20	-0, 34744	0, 2261	-12, 237
<b>GLOB</b>	gauche	512, 91	649, 87	3795, 63	-3796, 23	-0, 35247	0, 2246	-11, 886
	droite	506, 95	639, 84	3801, 56	-3802, 00	-0, 34677	0, 0867	-7, 850

**TAB. 1.4 :** Paramètres intrinsèques des deux caméras calculés par les méthodes **LIN** et **GLOB**.

paires de points des paires d'images qui n'ont pas servi au calibrage : les résultats sont donnés table 1.6. On y remarque l'amélioration apportée par notre méthode **GLOB** sur le respect de la contrainte épipolaire par rapport à la méthode classique **LIN1** : environ d'un facteur 3. Si

méthode	rotation en °			translation (en mm)		
	$\theta_{s_x}$	$\theta_{s_y}$	$\theta_{s_z}$	$t_{s_x}$	$t_{s_y}$	$t_{s_z}$
<b>LIN1</b>	12,160	0,165	0,303	0,46	96,75	5,93
<b>LIN<math>n</math></b>	11,991	0,104	0,276	0,99	95,46	5,70
<b>GLOB</b>	12,001	0,164	0,286	1,21	95,18	5,60

**TAB. 1.5 :** *Position et orientation de la caméra droite par rapport à la caméra gauche.*

caméra	<b>LIN1</b>	<b>LIN<math>n</math></b>	<b>GLOB</b>
gauche	0,3445	0,1815	0,1261
droite	0,3397	0,1789	0,1244

**TAB. 1.6 :** *Écart-types de la distribution du critère épipolaire évalué sur les méthodes **LIN1**, **LIN $n$**  et **GLOB**.*

L'on compare ces résultats par rapport à la simulation, on peut constater que l'on n'atteint pas le même ordre de précision. Cette différence peut s'expliquer par la faible précision à laquelle nous pouvons localiser les points de mesure dans les images d'évaluation, au regard de celle attendue (les couples de points ayant servis à l'évaluation du critère épipolaire pour la simulation étaient exempts de bruit).

## 1.5 Conclusion

Lorsque des mesures dimensionnelles de précision sont nécessaires, le besoin d'une méthode de calibrage précis d'un capteur de stéréovision est de grande importance. Nous avons comparé trois méthodes (**LIN1**, **LIN $n$**  et **GLOB**), toutes basées sur une approche photogrammétrique, qui estiment en même temps les paramètres intrinsèques de chaque caméra, les points tridimensionnels de la mire et, pour la méthode **GLOB**, la position relative des deux caméras. Nous avons montré que l'estimation globale des paramètres du capteur de stéréovision améliore beaucoup la précision de reconstruction tridimensionnelle par rapport à la méthode classique **LIN1** utilisée par la plupart des chercheurs en photomécanique. La méthode a été entièrement automatisée et est à la fois souple et précise.

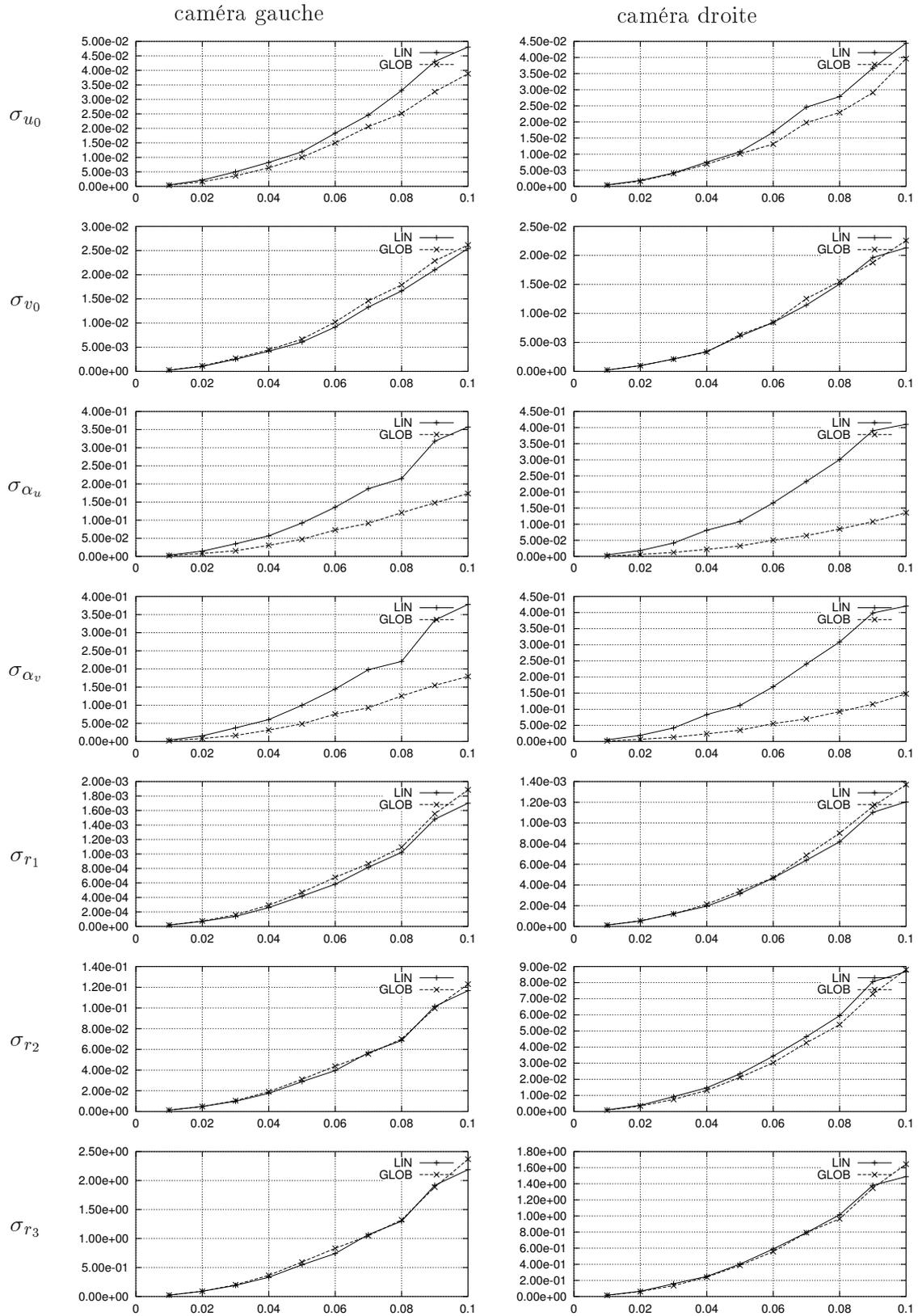


FIG. 1.18 : Écart-types des paramètres intrinsèques et de distorsion des deux caméras en fonction du niveau de bruit des points dans l'image.

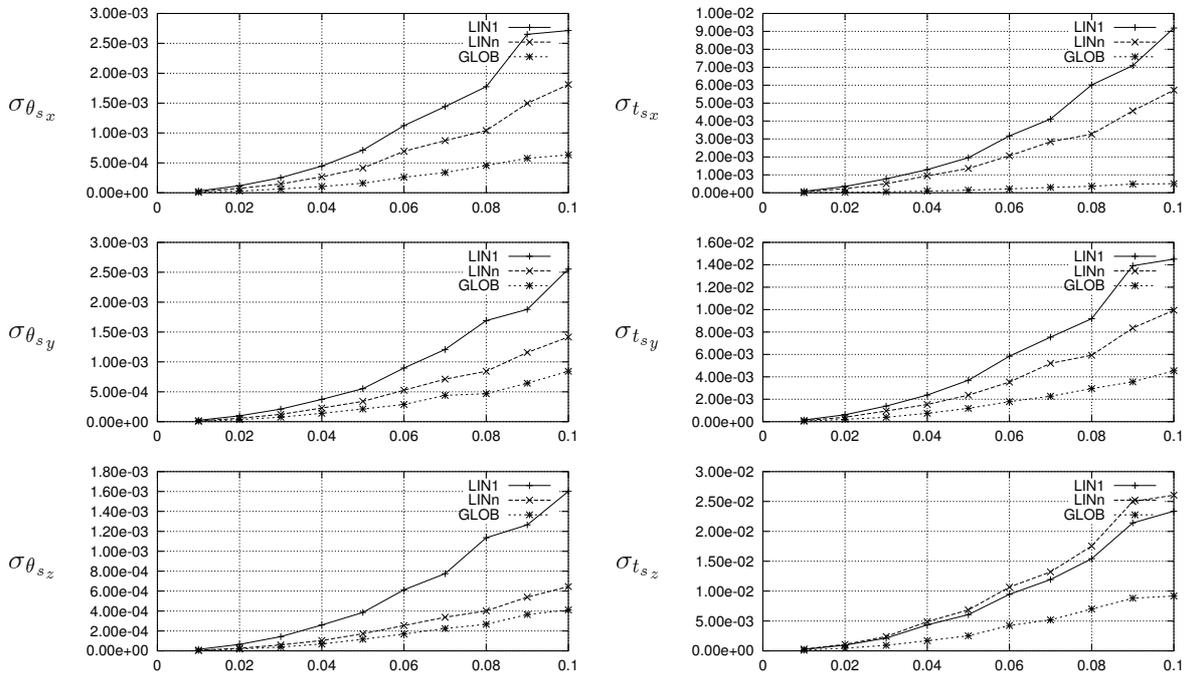


FIG. 1.19 : Écart-types des paramètres de la transformation rigide entre les deux caméras en fonction du niveau de bruit des points dans l'image.

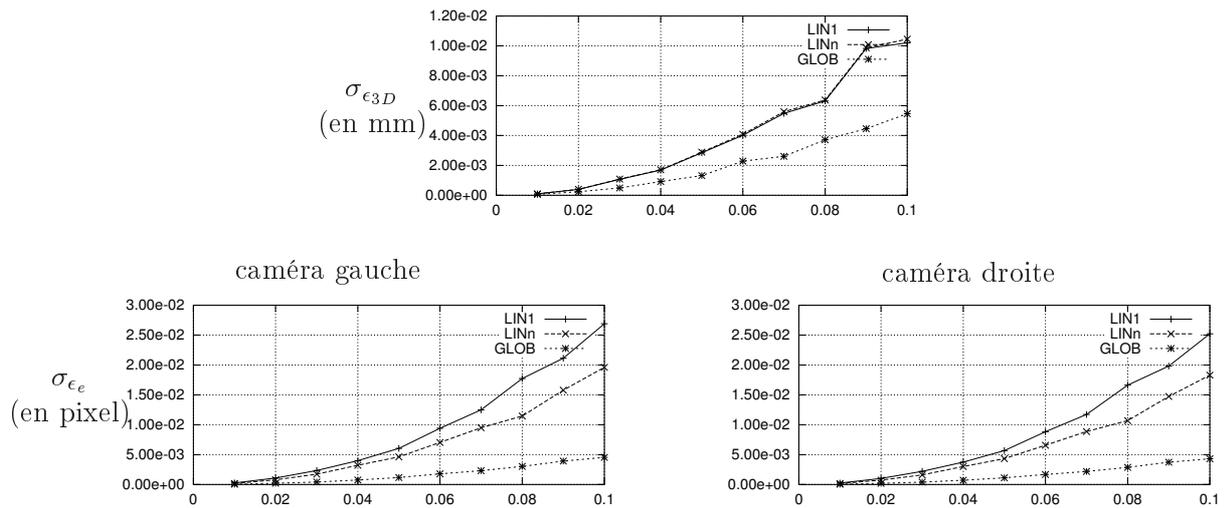
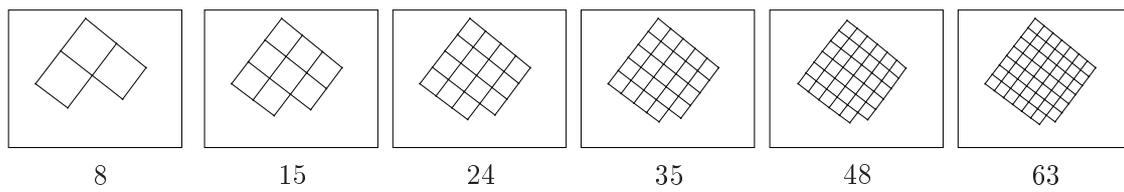


FIG. 1.20 : Écart-types des deux critères en fonction du niveau de bruit des points dans l'image.



**FIG. 1.21 :** *Les différentes mires utilisées pour évaluer l'influence du nombre de leurs points pour le calibrage. L'espacement entre deux points adjacents a été adapté de sorte que les mires soient toutes de même dimension.*

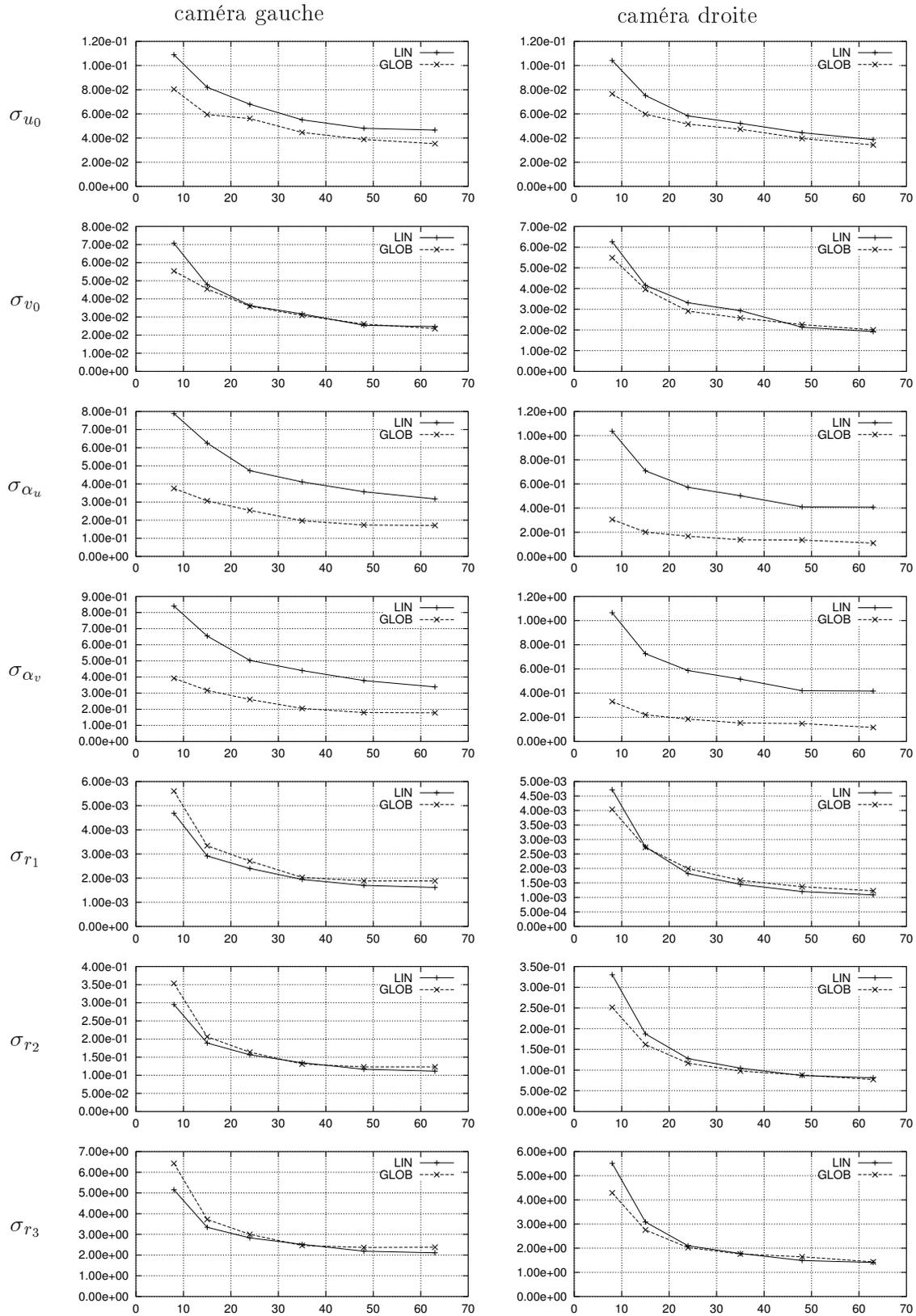


FIG. 1.22 : Écart-types des paramètres intrinsèques et de distorsion des deux caméras en fonction du nombre de points de la mire de calibration.

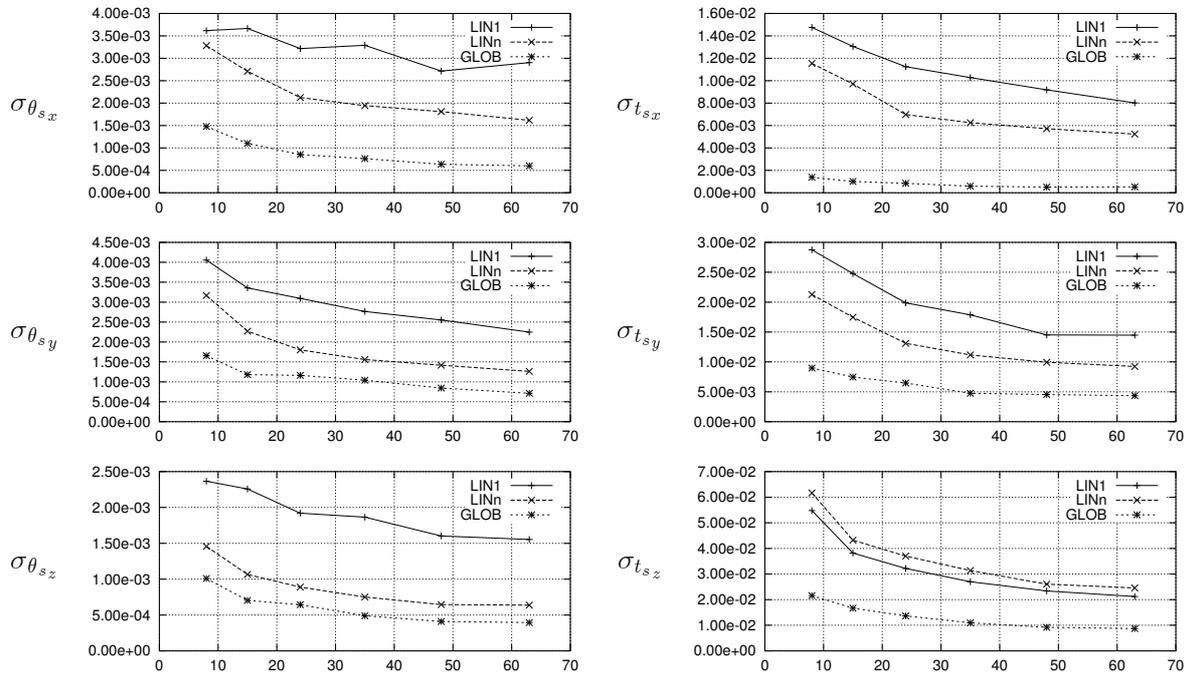


FIG. 1.23 : Écart-types des paramètres de la transformation rigide entre les deux caméras en fonction du nombre de points de la mire de calibrage.

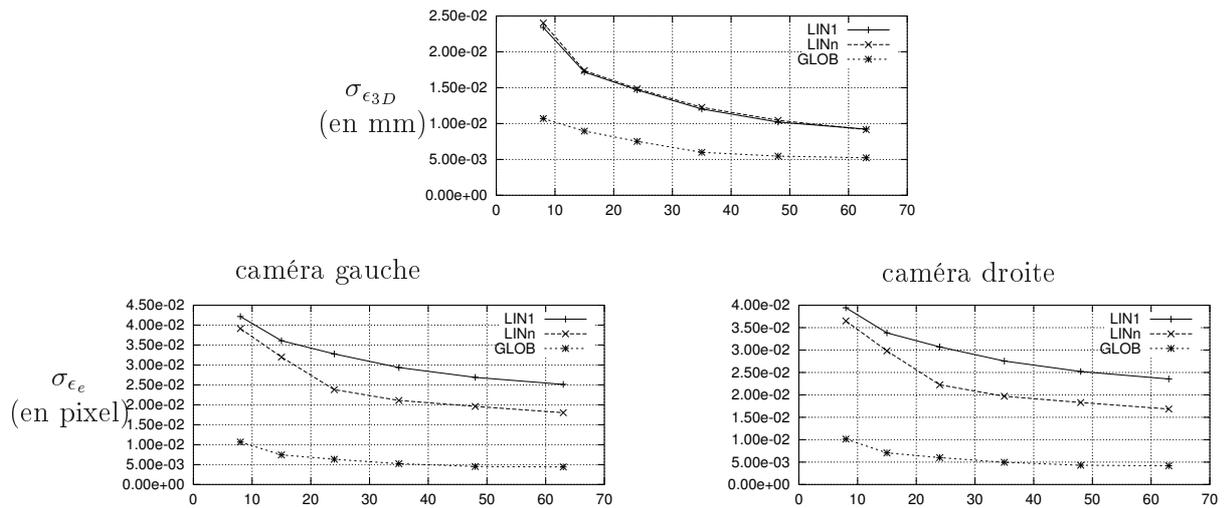


FIG. 1.24 : Écart-types des deux critères en fonction du nombre de points de la mire de calibrage.

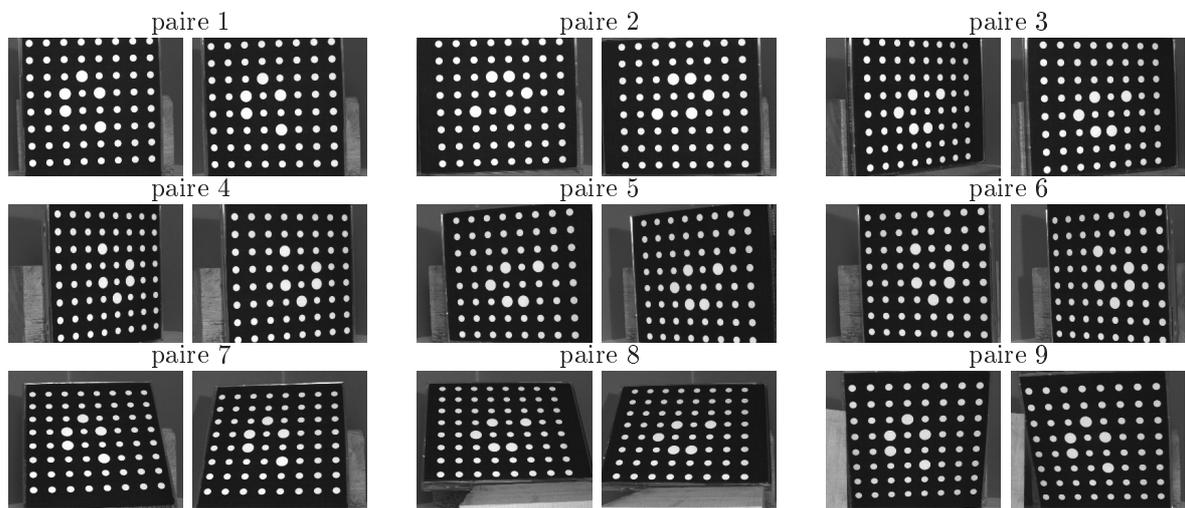


FIG. 1.25 : Vues utilisées pour le calibrage réel du capteur de stéréovision.



## Bibliographie du chapitre

- [Bey92] Horst A. BEYER. Accurate Calibration of CCD Cameras. Dans *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'92)*, 1992.
- [Bla97] Thierry BLASZKA. *Approches par modèles en vision précoce*. Thèse de doctorat, Université de Nice Sophia-Antipolis (France), février 1997.
- [Bra95] Pascal BRAND. *Reconstruction tridimensionnelle à partir d'une caméra en mouvement : de l'influence de la précision*. Thèse de doctorat, Université Claude Bernard – Lyon I (France), octobre 1995.
- [Dol96] Jürgen DOLD. Influence of large targets on the results of photogrammetric bundle adjustment. *International Archives of Photogrammetry and Remote Sensing*, 31(B5), 1996.
- [Fau93] Olivier FAUGERAS. *Three-Dimensional Computer Vision : A Geometric Viewpoint*. The MIT Press, 1993. ISBN 0-262-06158-9.
- [FT87] Olivier FAUGERAS et Giorgio TOSCANI. Camera Calibration for 3D Computer Vision. Dans *International Workshop on Machine Vision and Machine Intelligence*, pp. 240–247, Tokyo (Japon), février 1987.
- [Fus00] Andrea FUSIELLO. Uncalibrated Euclidean Reconstruction : A Review. *Image and Vision Computing*, 18:555–563, septembre 2000.
- [Gar96] Vincent GARRIC. *Vision pour la robotique de manipulation : calibration, localisation et saisie d'objets*. Thèse de doctorat, Institut National des Sciences Appliquées de Toulouse (France), mai 1996.
- [Har93] Richard I. HARTLEY. Computing Matched-epipolar Projections. Dans *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'93)*, pp. 549–555, New York (USA), juin 1993.
- [Har95] Richard I. HARTLEY. In defence of the 8-point algorithm. Dans *IEEE International Conference on Computer Vision (ICCV'95)*, pp. 1064–1070, Cambridge (USA), juin 1995.
- [HM95] R. HORAUD et O. MONGA. *Vision par ordinateur : outils fondamentaux*. Hermès, 1995. Deuxième édition.
- [HS97] R. I. HARTLEY et P. STURM. Triangulation. Dans *Computer Vision and Image Understanding (CVIU'97)*, volume 68 :2, pp. 146–157, novembre 1997.
- [KW97] K. KRAUS et P. WALDHÄUSL. *Manuel de photogrammétrie — principes et procédés fondamentaux*. Hermès, 1997. ISBN 2-8660-1656-4.
- [LD99] Manolis I.A. LOURAKIS et Rachid DERICHE. Camera self-calibration using the singular value decomposition of the fundamental matrix : From point correspondences to 3D measurements. Rapport de recherche 3748, INRIA, août 1999.
- [LD00] Manolis I.A. LOURAKIS et Rachid DERICHE. Camera self-calibration using the Kruppa equations and the SVD of the fundamental matrix : The case of varying intrinsic parameters. Rapport de recherche 3911, INRIA, 2000.
- [LDFP93] Q.-T. LUONG, R. DERICHE, O. FAUGERAS, et T. PAPADOPOULO. On determining the fundamental matrix : analysis of different methods and experimental results. Dans *Israeli Conf. on Artificial Intelligence and Computer Vision*, Tel-Aviv (Israel), 1993.

- [LEF97] A. LORUSSO, D. W. EGGERT, et R. B. FISHER. A Comparison of Four Algorithms for Estimating 3-D Rigid Transformations. Dans *Machine Vision and Applications*, volume 9, pp. 272–290, 1997.
- [LH81] H. C. LONGUET-HIGGINS. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, 293:133–135, 1981.
- [Luo92] Quang-Tuan LUONG. *Matrice fondamentale et calibration visuelle sur l'environnement — Vers une plus grande autonomie des systèmes robotiques*. Thèse de doctorat, Université de Paris-Sud, Centre d'Orsay (France), décembre 1992.
- [LVD99] J.M. LAVEST, M. VIALA, et M. DHOME. Quelle précision pour une mire d'étalonnage? *Traitement du Signal*, 16(3), 1999.
- [OGD97] J.J. ORTEU, V. GARRIC, et M. DEVEY. Camera calibration for 3D reconstruction : application to the measure of 3D deformations on sheet metal parts. Dans *European Symposium on Lasers, Optics and Vision in Manufacturing*, Munich (Allemagne), juin 1997.
- [Ort91] Jean-José ORTEU. *Application de la vision par ordinateur à l'automatisation de l'abattage dans les mines*. Thèse de doctorat, Université Paul Sabatier de Toulouse (France), novembre 1991.
- [Peu94] Bernard PEUCHOT. Utilisation de détecteurs subpixels dans la modélisation d'une caméra. Dans *Actes du 9ème Congrès AFCET (RFIA '94)*, pp. 691–695, Paris (France), janvier 1994.
- [Pér94] José-Philippe PÉREZ. *Optique géométrique et ondulatoire*. Masson, édition quatrième, 1994. ISBN 2-225-84270-1.
- [PTVF92] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, et B. P. FLANNERY. *Numerical Recipes in C — The Art of Scientific Computing*. Cambridge University Press, édition 2nd, 1992.
- [Tos87] Giorgio TOSCANI. *Système de calibration optique et perception du mouvement en vision artificielle*. Thèse de doctorat, Université de Paris-Sud, Centre d'Orsay (France), 1987.
- [Tsa87] R.Y. TSAI. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. *IEEE Journal of Robotics and Automation*, RA-3(4), août 1987.
- [WCH92] J. WENG, P. COHEN, et M. HERNIOU. Camera Calibration with Distorsion Models and Accuracy Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI'92)*, 14(10):965–980, octobre 1992.
- [Zel96] Cyril ZELLER. *Calibration projective, affine et euclidienne en vision par ordinateur et application à la perception tridimensionnelle*. Thèse de doctorat, École Polytechnique (France), février 1996.
- [Zha98a] Zhengyou ZHANG. A Flexible New Technique for Camera Calibration. Rapport technique MSR-TR-98-71, Microsoft Research, décembre 1998. Mis à jour en mars 1999.
- [Zha98b] Zhengyou ZHANG. Understanding the Relationship Between the Optimization Criteria in Two-View Motion Analysis. Dans *IEEE International Conference on Computer Vision (ICCV'98)*, pp. 231–236, Bombay (India), janvier 1998.

## Chapitre 2

# Suivi de points par corrélation

Dans ce chapitre nous nous intéressons à la mesure de champs de déplacements bidimensionnels en surface d'un objet, par analyse d'une séquence d'images acquises à des instants différents. Appliquée à la mécanique expérimentale pour la détermination de lois de comportement de matériaux, cette technique s'inscrit dans le cadre général des méthodes d'**extensométrie optique bidimensionnelle**.

Le but recherché par ces techniques est généralement la mesure de déformations. Habituellement, les points suivis au travers de la séquence d'images sont répartis suivant un maillage virtuel, ce qui fournit un champ de déplacements bidimensionnels en chacun des nœuds du maillage. Celui-ci est ensuite exploité pour le calcul des déformations au moyen d'un schéma numérique de dérivation discrète [SDC86, MD92, Dou00].

Pour être en mesure de déterminer un champ de déplacements d'un ensemble de points entre deux images d'une surface déformée, il faut résoudre le problème de l'appariement de point, ou mise en correspondance. Nous commencerons section 2.1 par un bref état de l'art sur les différentes approches existantes. Il y sera montré comment la méthode d'appariement par corrélation permet de résoudre efficacement ce problème. Nous présenterons en détail cette technique dans la section 2.2. La section 2.3 exposera la méthode que nous avons développée, puis nous analyserons ses performances dans la section 2.4.

### 2.1 État de l'art

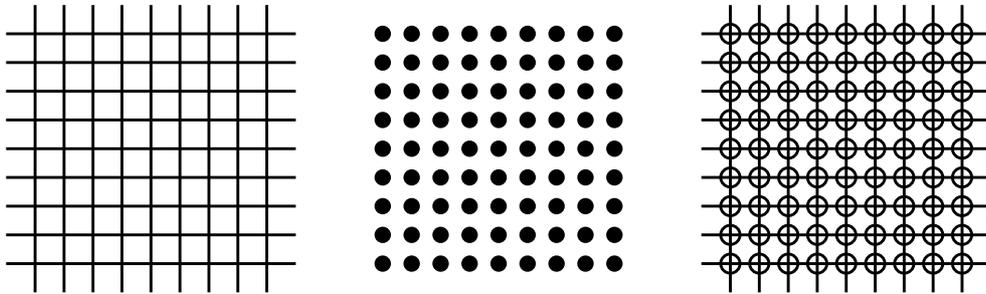
Il existe principalement trois méthodes de détermination du champ de déplacement d'un ensemble de points observé dans une séquence d'images. La première technique utilise des primitives, aussi appelées « marqueurs », dans les images. Une deuxième technique repose sur l'utilisation de dérivées partielles spatio-temporelles en chaque point des images pour déterminer leur vitesse de déplacement. Enfin, la troisième méthode repose sur la technique de corrélation en traitement du signal.

#### 2.1.1 Appariement de primitives

Cette technique d'appariement repose sur la recherche et l'extraction de primitives dans une séquence d'images, puis sur leur appariement. Ces primitives sont des attributs géométriques facilement identifiables et discriminants tels que des points caractéristiques, des segments de droite, des coniques, des régions, *etc.*

Dans le cadre général où le nombre et la répartition de ces primitives sont inconnus, leur mise en correspondance repose sur le choix d'un critère de similarité locale qui est fonction des attributs attachés à chaque primitives (par exemple l'intensité pour un point caractéristique), et d'un critère global permettant de contraindre la recherche parmi les primitives candidates potentielles. En revanche, lorsqu'on connaît approximativement la répartition des primitives dans les images, il est relativement facile de les apparier en opérant de proche en proche récursivement.

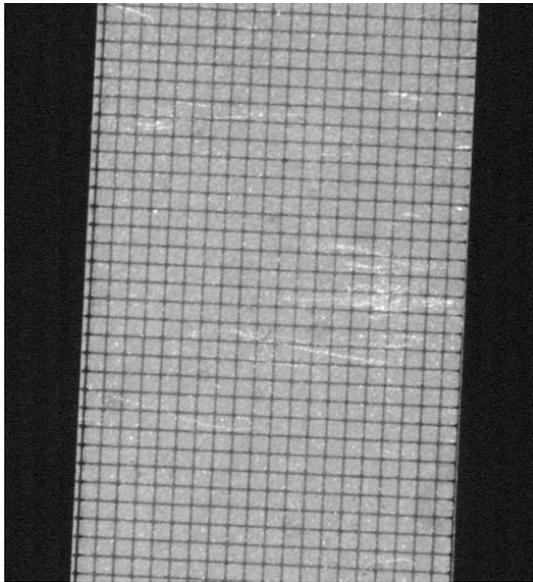
L'extensométrie optique bidimensionnelle a d'abord été développée autour de cette technique par l'utilisation de marqueurs de formes géométriques particulières et de répartition régulière [MBL93, BDL95, NGIG95, MBB95, GRMT98, HEG01, TBBD01]. Ainsi, ces primitives permettent par la seule analyse des variations de leurs attributs caractéristiques, de déterminer les directions principales et les valeurs du champ de déformation, sans utiliser de schéma numérique de dérivation du champ de déplacement. La plus répandue de ces formes géométriques est la composition d'un cercle et d'une croix, *cf.* figure 2.1. L'inconvénient de ce



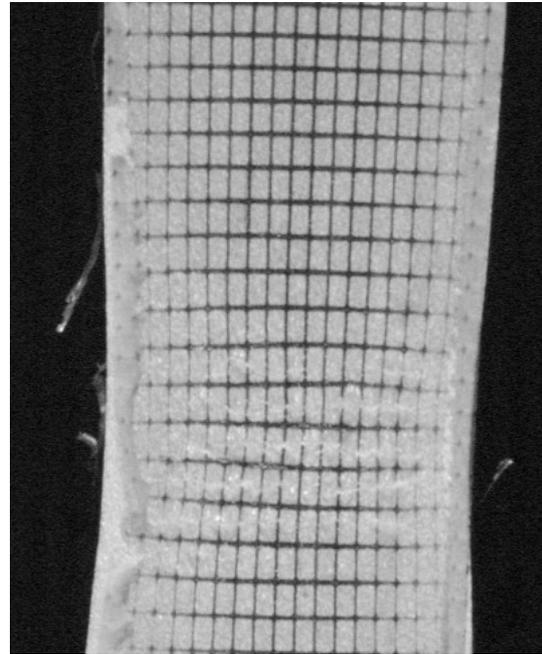
**FIG. 2.1 :** *Différents motifs de grilles utilisés par les méthodes de suivi de marqueurs (d'après [Dou00]).*

type d'approche réside dans la difficulté d'extraire précisément les attributs qui définissent ces motifs géométriques. Une autre solution consiste à utiliser une primitive élémentaire pouvant être localisée précisément par un extracteur approprié. C'est le cas notamment des croix et de l'extracteur de PEUCHOT [Peu94] ou de la méthode par courbes B-splines de BRAND [Bra95], ou bien des taches elliptiques avec l'extracteur de LAVEST [LVD99] ou de BRAND [Bra95], ou encore de coins [BD94]. Les champs de déformations peuvent alors être calculés en utilisant un schéma de dérivation discrète sur un voisinage de primitives [Dou00]. La figure 2.2 illustre cette méthode pour le cas de primitives de type croix, intersections de droites horizontales et verticales d'une grille.

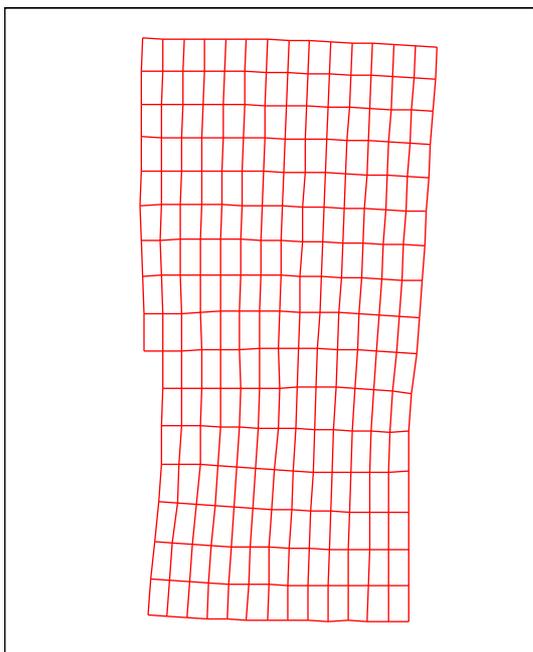
La détermination de champs de déplacements par cette technique présente deux inconvénients majeurs. Tout d'abord la surface du matériaux doit être marquée au préalable d'un réseau de motifs géométriques. La technique de gravage utilisée peut alors altérer le comportement de l'objet durant l'essai. De plus, la réalisation d'un marquage est une opération longue qui nécessite une préparation minutieuse de la surface de l'objet. Un deuxième inconvénient est lié à la densité des motifs géométriques suivis qui est fixée à la réalisation du marquage. Si celui-ci s'avère inadapté pour les conditions d'essais, il est nécessaire de réaliser une nouvelle éprouvette et de recommencer l'expérience. Ces deux problèmes sont fort contraignants et nous ont amené à étudier deux autres méthodes de mise en correspondance décrites dans les sections suivantes.



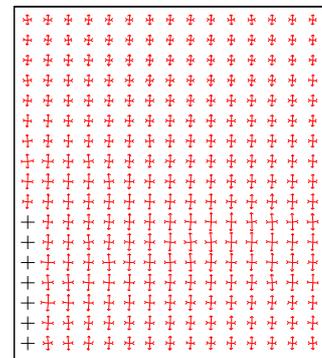
(1) éprouvette à l'état initial, marquée d'une grille



(2) éprouvette déformée



(3) grille déformée



(4) déformations principales calculées aux nœuds de la grille

**FIG. 2.2** : *Suivi des nœuds d'une grille lors d'un essai de traction sur éprouvette métallique.*

## 2.1.2 Méthodes différentielles

Les méthodes différentielles, appelées aussi **flot optique**, ont été initialement développées pour déterminer la vitesse de déplacement apparente d'une scène en mouvement par rapport à la caméra. Elles sont basées sur l'hypothèse de conservation de l'intensité lumineuse de tout point de la scène pendant leurs mouvements, c'est-à-dire qu'est considérée constante la réflectance en tout point de la scène et que l'illumination de celle-ci ne varie pas au cours du temps. Ces deux conditions sont les hypothèses fondamentales du flot optique.

Si l'on pose  $\mathcal{I}(\mathbf{m}, t)$  l'intensité lumineuse du point image  $\mathbf{m}$  à l'instant  $t$ , l'estimation du mouvement entre les instants  $t - \Delta_t$  et  $t$  revient à déterminer le vecteur déplacement  $\mathbf{d}$  tel que :

$$\mathcal{I}(\mathbf{m} - \mathbf{d}, t - \Delta_t) = \mathcal{I}(\mathbf{m}, t)$$

La contrainte de base définie par HORN et SCHUNCK dans leurs travaux fondateurs [HS81] est déduite d'une expansion en série de Taylor au premier ordre de l'égalité précédente :

$$\mathcal{I}(\mathbf{m} - \mathbf{d}, t - \Delta_t) = \mathcal{I}(\mathbf{m}, t) - \mathbf{d} \nabla \mathcal{I}(\mathbf{m}, t) - \Delta_t \frac{\partial \mathcal{I}(\mathbf{m}, t)}{\partial t} \quad \text{avec} \quad \nabla = \begin{pmatrix} \frac{\partial}{\partial u} \\ \frac{\partial}{\partial v} \end{pmatrix}$$

En posant le vecteur vitesse de déplacement  $\nu = \mathbf{d}/\Delta_t$ , il vient la contrainte spatio-temporelle :

$$\nu \nabla \mathcal{I}(\mathbf{m}, t) + \frac{\partial \mathcal{I}(\mathbf{m}, t)}{\partial t} = 0$$

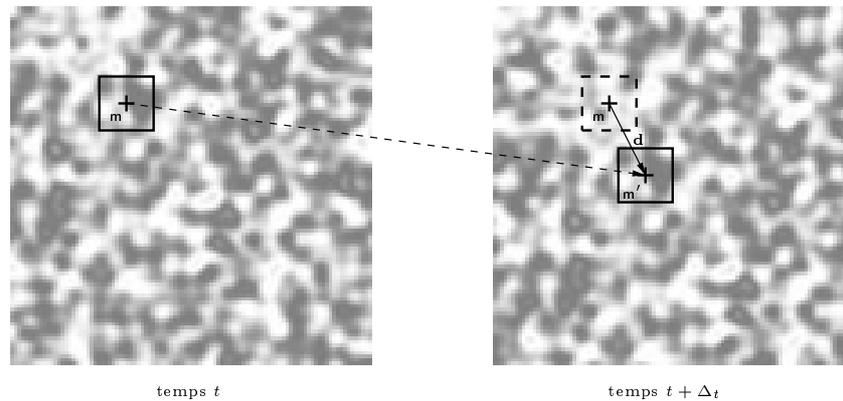
Cette équation introduit une seule contrainte, or le vecteur de mouvement d'un pixel possède deux composantes. Une deuxième contrainte est donc nécessaire pour estimer complètement chaque vecteur mouvement : cette condition est appelée **problème d'ouverture**.

De nombreuses solutions au problème d'ouverture ont été proposées et peuvent se classer en deux catégories : imposition de la régularité du flot optique et utilisation d'un modèle de déplacement. Cependant, la solution du flot optique dépend fortement de cette contrainte supplémentaire et tend à rendre difficile un choix objectif de celle-ci [Yan92]. De plus, les méthodes numériques mise en œuvre pour cette méthode ne fonctionnent que pour de faibles déplacements souvent inférieurs au pixel. Dans le cas de déplacements importants il faut recourir à une estimation préalable du champ de déplacement par suivi de marqueurs ou par corrélation (*cf.* section suivante) [Mon98]. Enfin, GALVIN *et al.* ont testé huit algorithmes de flot optique dans [GMN<sup>+</sup>98] et concluent que la meilleure des méthodes conduit à une précision d'environ 0,1 pixel seulement.

## 2.1.3 Méthodes par corrélation

Étant donné deux images correspondant à deux états de déformation d'un objet, pour déterminer l'homologue d'un pixel de la première image dans la seconde, les méthodes par corrélation mesurent la ressemblance entre deux pixels en calculant un score de corrélation (ou critère de similarité) déterminé sur leur voisinage. Le pixel correspondant dans la deuxième image est alors choisi comme celui qui maximise cette mesure de ressemblance — *cf.* figure 2.3. Le point et son voisinage, support de la mesure, sont appelés **fenêtre** ou **domaine de corrélation**.

L'intérêt majeur de cette classe de méthode réside dans la possibilité d'apparier *a priori* tous les points d'une image ayant un correspondant dans l'autre image, pour peu que l'objet



**FIG. 2.3 :** *Le suivi de pixels par corrélation consiste à déterminer le vecteur déplacement  $\mathbf{d}$  entre les positions  $m$  et  $m'$  de deux domaines de pixels pris dans deux images acquises à deux instants différents. La position  $m'$  du domaine de pixels au deuxième instant est déterminée par maximisation de la similarité du contenu des deux domaines via une mesure de corrélation.*

observé soit suffisamment texturé. Dès lors que l'on s'intéresse aux applications de l'extensométrie optique bidimensionnelle pour l'étude comportementale des matériaux, il s'avère que nombre d'entre eux possèdent déjà une texture suffisamment aléatoire et contrastée pour être utilisés sans préparation préalable (par exemple les bétons). Si un matériau n'est pas suffisamment texturé (tel qu'une éprouvette en métal poli), il est possible de pulvériser des gouttelettes de peinture sur sa surface [CRSP86]. La mise en correspondance par corrélation est donc d'utilisation plus simple, plus souple et moins coûteuse que les méthodes d'appariement de primitives. De plus, certains travaux sur la mise en correspondance précise par corrélation [BMSP89, SBS00, SMHS00] ont montré qu'il est possible d'atteindre une précision de 0,02 pixel, ce qui surclasse les méthodes de type flot optique. Pour ces raisons, l'appariement par corrélation paraît être le meilleur choix car cette approche combine à la fois souplesse, simplicité d'utilisation et précision, ce qui rejoint l'analyse de YANG [Yan92].

Les méthodes d'appariement de pixels par corrélation (y compris celle que l'on présentera dans le cadre stéréoscopique du chapitre 3) nécessitent le respect de trois contraintes majeures pour pouvoir fonctionner :

**limite de la déformation locale apparente** : la surface observée doit localement être similaire entre deux images à appairer pour présenter « suffisamment de similarité » aux mesures de corrélation ;

**contrainte de continuité** : la surface doit être et rester au moins localement continue pour que les voisinages de deux points en correspondance soient similaires. En particulier, la déformation de l'objet ne doit pas entraîner de rupture visible en surface, ou une hétérogénéité des déplacements ;

**contrainte lambertienne** : la surface doit être et rester globalement lambertienne, c'est-à-dire que la luminance en tout point de cette surface doit être indépendante de la position d'observation et de sa déformation (dans le cas contraire, les reflets spéculaires des sources d'éclairage mettent en échec la plupart des méthodes surfaciques).

## 2.2 Corrélation

Les méthodes d'appariement par corrélation reposent sur une mesure de similarité entre deux domaines de pixels à deux états différents de sollicitation de l'objet observé. Cela signifie que localement les domaines de pixels entre l'état initial et l'état déformé dans les images de l'objet sont eux-même déformés. L'approche classique de la mise en correspondance par corrélation consiste alors à faire l'hypothèse que ces déformations des domaines de pixels sont négligeables, et donc à utiliser des fenêtres de corrélation de même géométrie (souvent des carrés) entre tous les instants d'acquisition des images (exemple de la figure 2.3). Cette hypothèse conduit toutefois à une perte de précision dans l'appariement et l'on est alors amené à prendre en compte la transformation géométrique de chaque domaine entre les différents instants. Cet aspect est abordé dans la section suivante. Nous rappellerons ensuite les quatre principales mesures de corrélation utilisées en vision par ordinateur, puis nous verrons comment estimer les paramètres des transformations géométriques.

### 2.2.1 Approximation de la transformation locale

Entre les deux instants d'acquisition des images de l'objet, ce dernier peut avoir subi une déformation et un déplacement de solide rigide. La déformation de l'objet entraîne localement une distorsion de sa texture. Cela a pour conséquence de faire correspondre deux domaines de géométries différentes entre l'image de référence et celle déformée. Nous noterons  $\Phi$  la transformation matérielle qui relie les coordonnées  $\mathbf{m}_i$  et  $\mathbf{m}'_i$  des pixels entre ces deux images :

$$\mathbf{m}'_i = \Phi(\mathbf{m}_i)$$

Sans restreindre les applications envisagées, il est impossible de connaître l'expression de la transformation matérielle  $\Phi$ . Nous pouvons toutefois l'approcher localement car les domaines de corrélation sont de faibles dimensions.

Soit  $\mathbf{m}_c = (u_c \ v_c)^\top$  un point de l'image de référence, habituellement le centre du domaine de corrélation, et  $\mathbf{m} = (u \ v)^\top$  un point de ce domaine. En ce point  $\mathbf{m}$ , la transformation  $\Phi$  s'écrit :

$$\Phi(\mathbf{m}) = \mathbf{m} + \psi(\mathbf{m})$$

où  $\psi(\mathbf{m})$  est le déplacement au point  $\mathbf{m}$ . En écrivant le développement limité de  $\psi(\mathbf{m})$  au voisinage du point  $\mathbf{m}_c$ , on obtient une expression  $\Phi_c$  de  $\Phi$  en fonction des variations locales de  $\psi$  aux différents ordres :

$$\Phi_c(\mathbf{m}, \mathbf{m}_c) = \mathbf{m} + \psi(\mathbf{m}_c) + \frac{\partial \psi}{\partial \mathbf{m}}(\mathbf{m}_c) \cdot (\mathbf{m} - \mathbf{m}_c) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_c)^\top \cdot \frac{\partial^2 \psi}{\partial \mathbf{m}^2}(\mathbf{m}_c) \cdot (\mathbf{m} - \mathbf{m}_c) + \dots$$

Cette expression tronquée à différents ordres conduit à différentes approximations  $\Phi_c(\mathbf{m}, \mathbf{m}_c, \mathbf{p})$  de  $\Phi(\mathbf{m})$  autour du point  $\mathbf{m}_c$ , définissant diverses variantes de la méthode de corrélation. Le vecteur  $\mathbf{p}$  est le vecteur des paramètres de l'approximation de la transformation locale.

#### Ordre 0

À l'ordre 0, les variations de déplacement sur le domaine de corrélation sont supposées nulles. C'est la modélisation la plus simple qui conduit à considérer que  $\Phi$  est localement une translation autour du point  $\mathbf{m}_c$  :

$$\Phi_c(\mathbf{m}, \mathbf{m}_c, \mathbf{p}) = \mathbf{m} + \mathbf{t} \quad \text{avec} \quad \mathbf{t} = \begin{pmatrix} t_u \\ t_v \end{pmatrix} \quad (2.1)$$

Le vecteur  $\mathbf{p}$  des paramètres de l'approximation de la transformation locale à l'ordre 0 est égal au vecteur de translation  $\mathbf{t}$  :  $\mathbf{p} = (t_u \ t_v)^\top$ .

### Ordre 1

À l'ordre 1, la transformation est représentée par une translation et une déformation locale homogène sur le domaine de corrélation en supposant des variations linéaires du champ de déplacements. Cela conduit à une transformation affine d'expression :

$$\Phi_c(\mathbf{m}, \mathbf{m}_c, \mathbf{p}) = \mathbf{m} + \mathbf{t} + \begin{pmatrix} a_u & a_v \\ b_u & b_v \end{pmatrix} \cdot (\mathbf{m} - \mathbf{m}_c) \quad (2.2)$$

où les termes  $a_u$ ,  $a_v$ ,  $b_u$  et  $b_v$  représentent des termes d'élongation. Le vecteur  $\mathbf{p}$  est composé des six paramètres de l'approximation de la transformation locale à l'ordre 1 :  $\mathbf{p} = (t_u \ t_v \ a_u \ a_v \ b_u \ b_v)^\top$ .

### Ordre 2

À l'ordre 2, la transformation est une expression quadratique comprenant douze paramètres :

$$\begin{aligned} \Phi_c(\mathbf{m}, \mathbf{m}_c, \mathbf{p}) = & \mathbf{m} + \mathbf{t} + \begin{pmatrix} a_u & a_v \\ b_u & b_v \end{pmatrix} \cdot (\mathbf{m} - \mathbf{m}_c) \\ & + \begin{pmatrix} a_{uv}(u - u_c)(v - v_c) + a_{uu}(u - u_c)^2 + a_{vv}(v - v_c)^2 \\ b_{uv}(u - u_c)(v - v_c) + b_{uu}(u - u_c)^2 + b_{vv}(v - v_c)^2 \end{pmatrix} \end{aligned} \quad (2.3)$$

Le vecteur  $\mathbf{p}$  des paramètres de l'approximation de la transformation locale à l'ordre 2 est :

$$\mathbf{p} = (t_u \ t_v \ a_u \ a_v \ a_{uu} \ a_{uv} \ a_{vv} \ b_u \ b_v \ b_{uu} \ b_{uv} \ b_{vv})^\top$$

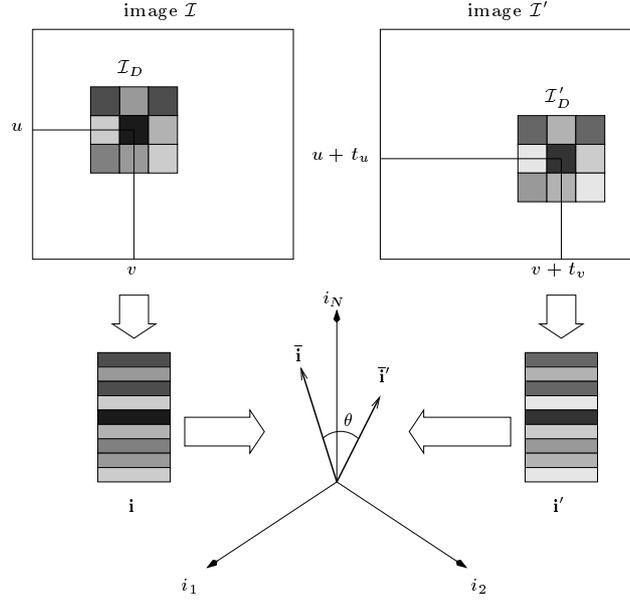
## 2.2.2 Coefficients de corrélation

Les méthodes d'appariement par corrélation sont basées sur une mesure de similarité entre deux domaines de pixels. Afin de rendre ce score de corrélation le plus discriminant possible, il est préférable d'exploiter de grands domaines de corrélation. Cependant, non seulement les temps de calculs deviennent élevés, mais aussi l'hypothèse d'homogénéité de la déformation sur le domaine s'affaiblit. Il y a donc là un compromis à faire pour la taille du domaine de corrélation. Généralement, ce choix est guidé par l'expérience et les longueurs des côtés de domaines carrés s'échelonnent de 7 à 25 pixels dans la majorité des cas.

Appelons  $\mathbf{m}_c$  un pixel de l'image  $\mathcal{I}$  de référence,  $V(\mathbf{m}_c)$  l'ensemble des  $N$  pixels voisins du point  $\mathbf{m}_c$ , et  $\mathcal{I}'$  l'image déformée. Nous notons  $\mathbf{p}$  le vecteur des paramètres de l'approximation de la transformation locale autour du point  $\mathbf{m}_c$ , *i.e.* selon l'équation (2.1), (2.2) ou (2.3), le point  $\mathbf{m}$  de l'image  $\mathcal{I}$  a pour homologue le point  $\Phi_c(\mathbf{m}, \mathbf{m}_c, \mathbf{p})$  dans l'image déformée  $\mathcal{I}'$ . Tous les critères de corrélation  $C$  que nous allons présenter peuvent alors se formuler selon l'équation suivante :

$$C(\mathbf{m}_c, \mathbf{p}) = \sum_{\mathbf{m}_i \in V(\mathbf{m}_c)} \epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p}) \quad (2.4)$$

Les différents critères de corrélation se différencient dans l'expression de la fonctionnelle  $\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})$ . Nous allons passer en revue les quatre expressions des critères classiques SSD, ZSSD, NCC et ZNCC [AG92, FHM<sup>+</sup>93]. Nous invitons le lecteur à se référer à la thèse de LAN [Lan97] pour la présentation d'autres critères de corrélation moins fréquemment utilisés.



**FIG. 2.4 :** Calcul du degré de corrélation entre deux domaines  $\mathcal{I}_D$  et  $\mathcal{I}'_D$  de dimension  $3 \times 3$ , encodés respectivement par les vecteurs  $\mathbf{i}$  et  $\mathbf{i}'$  dans  $\mathbb{R}^9$ . La mesure ZNCC correspond au cosinus de l'angle  $\theta$  entre les vecteurs  $\bar{\mathbf{i}}$  et  $\bar{\mathbf{i}'}$  obtenus par soustraction de l'intensité moyenne de chacun des domaines respectifs. La mesure ZSSD correspond à norme euclidienne du vecteur  $\bar{\mathbf{i}} - \bar{\mathbf{i}'}$ .

### Critère SSD

La mesure **Sum of Squared Differences** est définie par l'équation (2.4) dont la fonctionnelle  $\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})$  a pour expression :

$$\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p}) = \left[ \mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})) - \mathcal{I}(\mathbf{m}_i) \right]^2$$

Ce critère correspond au carré de la norme  $L_2$  évaluée sur les  $N$  points du voisinage de  $\mathbf{m}_c$ , entre les fonctions  $\mathcal{I}$  et  $\mathcal{I}'$ , respectivement l'image de référence et l'image déformée. Géométriquement, il correspond au carré de la distance euclidienne  $\|\mathbf{i} - \mathbf{i}'\|_2$  entre les vecteurs  $\mathbf{i}$  et  $\mathbf{i}'$  de la figure 2.4. Notons que deux domaines de corrélation identiques annulent ce critère.

### Critère ZSSD

La version centrée de SSD est le critère **Zero mean Sum of Squared Differences** et est défini par l'équation (2.4) dont la fonctionnelle  $\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})$  a pour expression :

$$\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p}) = \left[ (\mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})) - \bar{\mathcal{I}'}) - (\mathcal{I}(\mathbf{m}_i) - \bar{\mathcal{I}}) \right]^2$$

où  $\bar{\mathcal{I}}$  et  $\bar{\mathcal{I}'}$  sont les niveaux de gris moyens respectivement sur les domaines de corrélation de l'image de référence et de l'image déformée, et ont pour expressions :

$$\bar{\mathcal{I}} = \frac{1}{N} \sum_{\mathbf{m}_i \in V(\mathbf{m}_c)} \mathcal{I}(\mathbf{m}_i) \quad \text{et} \quad \bar{\mathcal{I}'} = \frac{1}{N} \sum_{\mathbf{m}_i \in V(\mathbf{m}_c)} \mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})) \quad (2.5)$$

Géométriquement, le critère ZSSD correspond au carré de la distance euclidienne  $\|\bar{\mathbf{i}} - \bar{\mathbf{i}}'\|_2$  entre les vecteurs  $\bar{\mathbf{i}}$  et  $\bar{\mathbf{i}}'$  de la figure 2.4. Notons que deux domaines de corrélation identiques à une valeur additive  $k$  près, c'est-à-dire que  $\mathcal{I}(\mathbf{m}_i) = \mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})) + k$ ,  $\forall \mathbf{m}_i \in V(\mathbf{m}_c)$ , annulent ce critère.

### Critère NCC

La mesure **Normalized Cross Correlation** est définie par l'équation (2.4) dont la fonctionnelle  $\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})$  a pour expression :

$$\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p}) = \frac{\mathcal{I}(\mathbf{m}_i) \cdot \mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p}))}{\sum_{\mathbf{m}_j \in V(\mathbf{m}_c)} \left(\mathcal{I}(\mathbf{m}_j)\right)^2 \cdot \sum_{\mathbf{m}_j \in V(\mathbf{m}_c)} \left(\mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_j, \mathbf{p}))\right)^2}$$

Ce critère correspond géométriquement à  $\cos(\theta)$  de la figure 2.4. À la différence des deux critères précédents, celui-ci prend la valeur unitaire lorsque les deux domaines de corrélation sont identiques à un facteur multiplicatif non nul près. Les critères de corrélation présentés seront utilisés dans une procédure d'optimisation dans les sections suivantes. Afin de se ramener au seul cas de la minimisation, nous pourrions optimiser le critère  $1 - C_{\text{NCC}}(\mathbf{m}_c, \mathbf{p})$  qui s'annule pour deux domaines identiques à un facteur multiplicatif non nul près.

### Critère ZNCC

Le critère **Zero mean Normalized Cross Correlation** est définie par l'équation (2.4) dont la fonctionnelle  $\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})$  a pour expression :

$$\epsilon(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p}) = \frac{\left(\mathcal{I}(\mathbf{m}_i) - \bar{\mathcal{I}}\right) \cdot \left(\mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_i, \mathbf{p})) - \bar{\mathcal{I}}'\right)}{\sum_{\mathbf{m}_j \in V(\mathbf{m}_c)} \left(\mathcal{I}(\mathbf{m}_j) - \bar{\mathcal{I}}\right)^2 \cdot \sum_{\mathbf{m}_j \in V(\mathbf{m}_c)} \left(\mathcal{I}'(\Phi_c(\mathbf{m}_c, \mathbf{m}_j, \mathbf{p})) - \bar{\mathcal{I}}'\right)^2}$$

où  $\bar{\mathcal{I}}$  et  $\bar{\mathcal{I}}'$  sont les niveaux de gris moyens respectivement sur les domaines de corrélation de l'image de référence et de l'image déformée, voir équation (2.5). La soustraction de chaque niveau de gris par leur moyenne sur leur domaine respectif permet de rendre ce critère invariant à une valeur additive près des niveaux de gris d'un des deux domaines. De plus, ce critère est, comme NCC, invariant à un facteur multiplicatif non nul près des niveaux de gris d'un des deux domaines ce qui le rend particulièrement robuste aux changements d'illumination de l'objet observé entre l'image de référence et l'image déformée. Ici encore, pour se ramener à une mesure nulle indiquant une parfaite corrélation entre les deux domaines, on peut utiliser le critère  $1 - C_{\text{ZNCC}}(\mathbf{m}_c, \mathbf{p})$ .

### 2.2.3 Techniques d'estimation de la transformation locale

Le problème de l'estimation des paramètres de la transformation locale passe par la résolution de problèmes de minimisation de la forme :

$$\mathbf{p} = \arg \min_{\mathbf{p}} C(\mathbf{m}_c, \mathbf{p})$$

où le vecteur de paramètres  $\mathbf{p}$  est celui du modèle choisi de la transformation locale, et  $C$  un critère de corrélation. Les deux critères  $C_{\text{SSD}}$  et  $C_{\text{ZSSD}}$  sont sous la forme d'une somme

d'erreurs quadratiques et l'on peut donc utiliser une méthode de minimisation telle que celle de LEVENBERG-MARQUARDT. En revanche, les critères  $C_{NCC}$  et  $C_{ZNCC}$  ne s'expriment pas sous cette forme, et nécessitent donc l'utilisation d'une technique d'optimisation de fonction objective telle que la méthode de descente du gradient utilisée par DOUMALIN dans sa thèse [Dou00], ou la méthode du gradient conjugué [PTVF92].

Ces méthodes numériques nécessitent la connaissance d'une estimée initiale de chacun des paramètres à calculer, notamment la position approximative de l'appariement (c'est-à-dire une connaissance des paramètres de translation  $t_u$  et  $t_v$  du modèle de la transformation locale utilisé). Cette dernière peut-être obtenue en utilisant une méthode classique d'appariement par corrélation : recherche bidimensionnelle en coordonnées entières dans une région de l'image suffisamment grande pour être certain qu'elle contient l'appariement. L'estimation de ce dernier est alors le point (et son domaine de pixel associé) qui maximise la mesure de ressemblance. Les autres paramètres peuvent être approximés en se basant sur les estimations obtenus dans l'image précédente de la séquence, ou bien en les considérant nuls si la déformation des domaines de corrélation est peu importante.

#### 2.2.4 Corrélation subpixel : modèle de luminance

Les techniques d'optimisation, telles que celles de LEVENBERG-MARQUARD ou du gradient conjugué, nécessitent la connaissance des gradients de niveaux de gris dans les images. De plus, du fait de l'approche sous-pixellique de l'appariement, on est conduit à évaluer des niveaux de gris dans des images en des coordonnées non entières. Or une image n'est qu'une représentation discrète de la scène réellement observée. Il est donc indispensable de définir un modèle continue de la luminance de la scène observée à partir de sa seule représentation discrète : l'image.

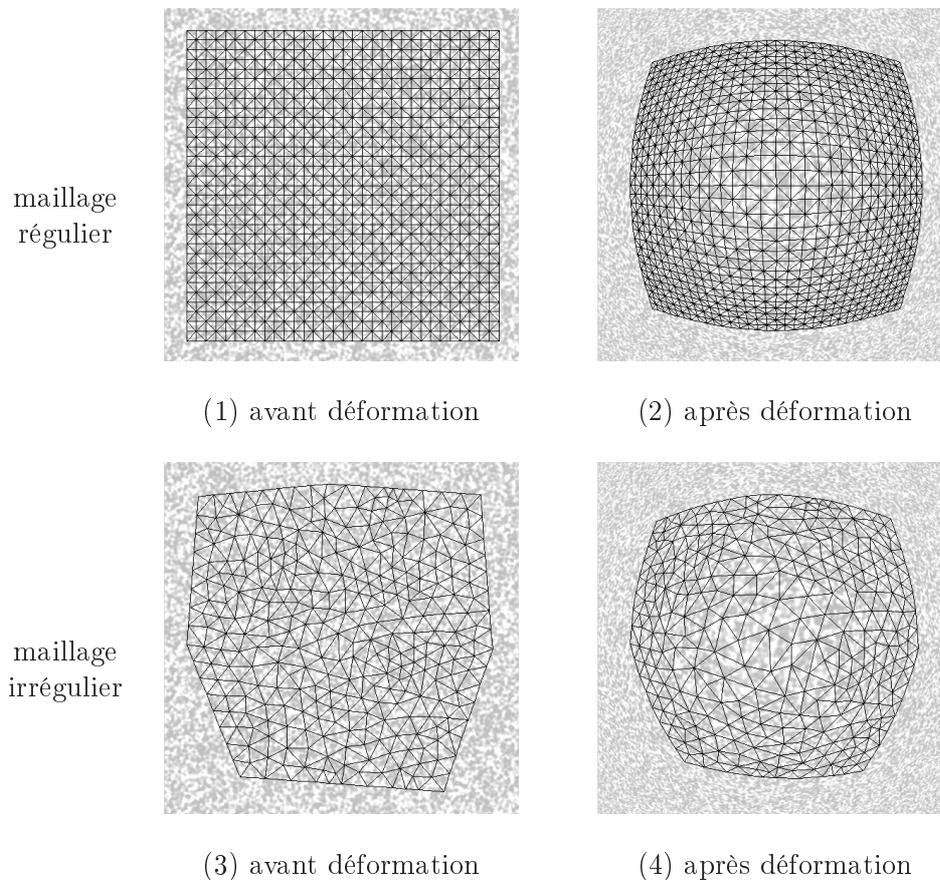
Cette modélisation revient à interpoler chacun des niveaux de gris par une fonction bidimensionnelle appropriée. De plus en plus de chercheurs utilisent un interpolant de type surface B-spline pour sa simplicité de mise en œuvre, et ses bonnes propriétés de fidélité au signal avant numérisation [Uns99]. De plus, il existe une implantation très efficace permettant une évaluation très rapide de ce modèle [UAE93a, UAE93b]. Par ailleurs, SCHREIER *et al.* ont montré dans [SBS00] que l'utilisation de surfaces B-spline d'ordre supérieur ou égal à trois permet de réduire significativement l'erreur systématique d'appariement, ce que nous observerons et confirmerons dans la section 2.4, notamment par rapport à une simple interpolation bilinéaire des niveaux de gris. Enfin, il est aisé de déterminer le gradient des niveaux de gris en tout point de ce modèle par sa dérivation analytique. Dès lors que l'ordre du modèle B-spline est supérieur ou égal deux, ce dernier est continuellement dérivable en tout point, ce qui n'est pas le cas d'une simple interpolation bilinéaire. Ceci est un avantage important pour les méthodes d'optimisation qui ont de meilleures propriétés de convergence lorsque la fonctionnelle à minimiser est continuellement dérivable.

### 2.3 Méthode développée

La méthode que nous avons développée met en œuvre un modèle de transformation locale jusqu'à l'ordre deux, et propose l'utilisation d'un critère de corrélation parmi quatre : SDD, ZSSD, NCC et ZNCC. Nous modélisons la luminance dans chacune des image en les interpolant par des surfaces B-spline dont l'ordre peut être choisi entre 1 et 5.

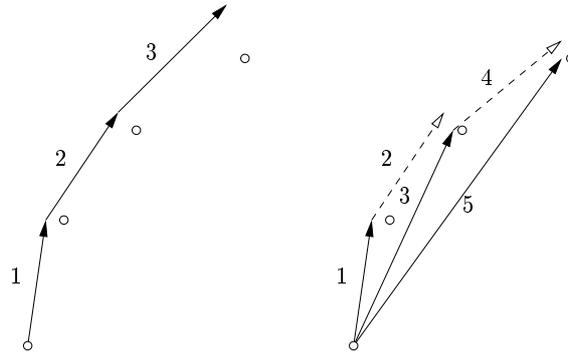
La plupart des applications d'extensométrie optique bidimensionnelle visent l'obtention d'un champ de déformations. Celui-ci est calculé à partir du champ de déplacements mesuré.

Ce calcul, qui s'effectue par dérivation numérique [SDC86, MD92, Dou00], est facilité si le champ de déplacements mesuré est obtenu pour un ensemble de points répartis suivant un maillage. Pour obtenir ce maillage, nous avons choisi d'implanter la méthode de triangulation de DELAUNAY contrainte, proposée par SHEWCHUK [She96], pour générer des maillages triangulaires. Les contraintes pouvant être imposées concernent notamment la taille maximale des triangles générés, ainsi que leur angle minimum. La figure 2.5 montre le résultat de deux maillages générés selon cet algorithme : le maillage régulier est obtenu par sélection de quatre points, sommets d'un carré, et par triangulation sous la contrainte de surface maximale de chaque triangle généré ; le maillage irrégulier est obtenu par triangulation de sept points selon la même technique.



**FIG. 2.5 :** Exemple du suivi des nœuds d'un maillage (1) régulier, et (3) irrégulier, générés par un algorithme de triangulation de DELAUNAY contrainte, sur une image ayant subi une déformation de type punch (cf. page 66).

Le suivi de points par corrélation dans une séquence d'images peut être conduit selon deux approches. La première consiste à suivre chaque point entre l'image d'indice  $i$  et l'image d'indice  $i+1$  dans la séquence, puis de réitérer le processus pour les points entre les images  $i+1$  et  $i+2$ , et ainsi de suite jusqu'au traitement complet de la séquence. Cette approche, utilisée par LARABA-ABBES dans [LA98] a l'inconvénient majeur de cumuler les erreurs d'appariement tout au long de la séquence. La figure 2.6, schéma de gauche, illustre ce propos : le suivi du point entre les deux premiers instants de la séquence conduit à déterminer les coordonnées



**FIG. 2.6 :** Deux approches pour le suivi d'un pixel dans une séquence d'images (les cercles représentent la position réelle d'un point dans la séquence d'images) : à gauche, l'approche incrémentale conduit à cumuler les erreurs d'appariement au travers de la séquence ; à droite, cet inconvénient est contourné en se ramenant systématiquement à un problème d'appariement avec la première image de la séquence, les appariements entre deux images successives servent seulement d'estimations.

d'un point entaché d'erreur, lui-même suivi à l'instant suivant et conduisant à une localisation encore plus erronée, et ainsi de suite tout au long de la séquence.

Une deuxième approche consiste à suivre les points entre la première image et l'image d'indice  $i$  dans la séquence. Les points de l'image d'indice  $i$  sont suivis dans l'image d'indice  $i + 1$  afin de fournir des estimées pour le suivi des points entre la première image et l'image d'indice  $i + 1$ . La figure 2.6, schéma de droite, explicite cette approche :

1. le point est apparié dans la seconde image ;
2. une estimation de son homologue dans la troisième image est déterminée par le même procédé ;
3. cette estimation est utilisée pour appairer le point de la première image dans la troisième et, ainsi, s'affranchir de l'influence de l'erreur commise dans l'appariement entre les deux premières images ;
4. *etc.*

Bien que cet algorithme nécessite l'estimation de près de deux fois plus d'appariements (et donc de temps de calcul), il évite le cumul des erreurs au travers de la séquence et constitue donc une méthode de choix pour des applications où la meilleure précision possible est l'objectif recherché. C'est la raison pour laquelle nous avons retenu cette approche dans notre implantation. Toutefois, cette technique peut montrer ses limites pour de longues séquences d'images où, entre la première et la dernière image, la déformation locale peut présenter de fortes hétérogénéités, et donc ne pas suivre fidèlement le modèle de la transformation locale utilisé. Cet inconvénient est d'un impact moindre avec la première approche où, entre deux instants successifs d'acquisition des images, l'hypothèse de déformation locale homogène est plus facilement vérifiée. Préalablement à l'utilisation de notre méthode, il faudra donc s'assurer que l'objet étudié en déformation ne présente pas fortes hétérogénéités de son champ de déplacement pour la taille des domaines de corrélation utilisée.

Nous avons présenté dans la section 2.2.3 une manière d'obtenir une estimée initiale de la position d'un appariement dans une image, nécessaire aux méthodes numériques d'estimation précise des paramètres de la transformation locale choisie. Elle repose sur une recherche

bidimensionnelle de l'appariement dans une région de l'image. Si cette région est de grande dimension, cela entraîne un fort ralentissement de la rapidité de calcul, et une augmentation potentielle du taux de faux appariements. Pour palier à ces inconvénients, nous réduisons la taille de cette zone de recherche et la centrons sur une prédiction de la position de chaque appariement par interpolation des déplacements déjà estimés au voisinage de chaque point. Expérimentalement, nous avons constaté que cette méthode fonctionne bien si le nombre de point à apparier est dense, *i.e.* si les points servant à la prédiction sont suffisamment proche du point à apparier. YANG utilise cette méthode dans sa thèse [Yan92] mais propose aussi une approche plus complexe basée sur l'utilisation d'un filtre de KALMAN.

## 2.4 Évaluation des performances

Il est difficile d'évaluer quantitativement les performances des algorithmes de suivi de pixels par corrélation, en vision monoculaire, à partir d'images réelles. En effet, la « vérité terrain » est délicate à obtenir et demande la mise en place d'un dispositif expérimental complexe pour maîtriser le positionnement de l'objet texturé visualisé, et un calibrage précis du capteur.

TOUCHAL-MGUIL [TM97] utilise une table micrométrique à déplacements croisés  $x$ - $y$  et une source laser, fixée à la caméra, lui permettant de positionner l'objet plan observé parallèlement au plan rétinien de la caméra (*cf.* figure 2.7). L'auteur n'indique pas de quelle manière est calibrée la caméra pour lui permettre d'estimer l'influence dans l'image d'un déplacement de l'objet par la table micrométrique. Toutefois, le positionnement opto-mécanique de la caméra ne peut être parfait, y compris l'estimation des paramètres du modèle de caméra. L'évaluation des performances de son algorithme de suivi de pixels par corrélation est donc biaisée par les incertitudes citées (précision annoncée de l'ordre de 0,02 pixels).

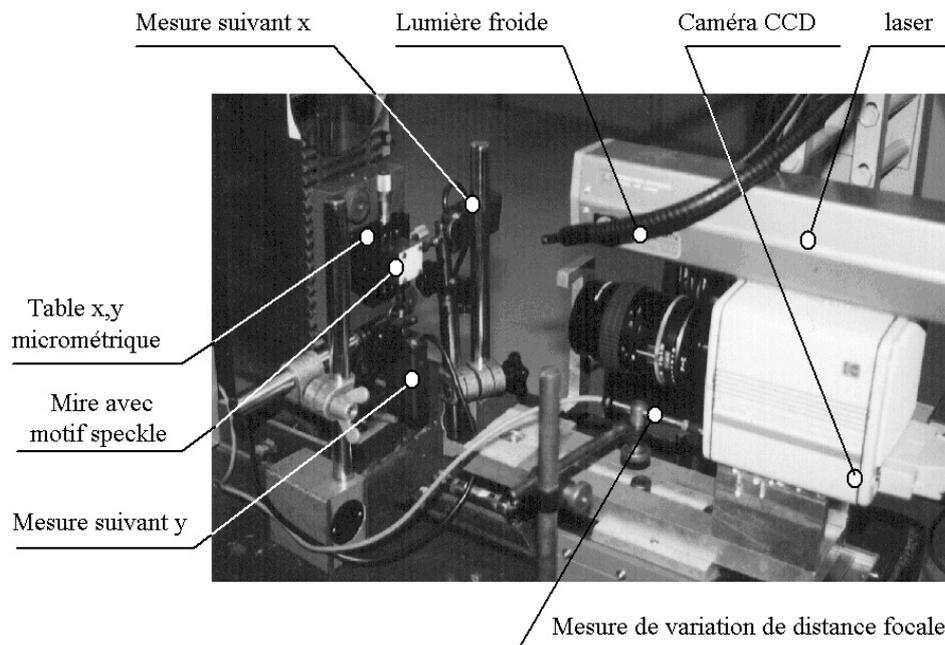
LARABA-ABBES [LA98] a étudié les performances de son système d'extensométrie optique bidimensionnelle en translatant un objet plan présentant une texture obtenue par granularité laser (effet SPECKLE). Des déplacements de 10 micromètres ont été imposés et mesurés via un capteur interférométrique de précision  $\frac{1}{10}$  micromètres. Lors de ses essais, l'auteur a pu déterminer une erreur maximale de  $\pm \frac{1}{16}$  pixel. Il faut toutefois nuancer ces résultats avec l'incertitude de 1% des mesures du capteur interférométrique pour les déplacements imposés.

Afin d'être en mesure d'évaluer sans biais les performances théoriques de nos algorithmes de suivi de pixels par corrélation, nous avons donc opté pour une approche par simulation. Cette solution nous permet de maîtriser parfaitement les mouvements imposés à l'objet plan virtuel observé, ainsi que de déterminer l'influence de nombreux paramètres tels le niveau de bruit des niveaux de gris dans l'image, les différences de dynamique d'acquisition, etc. Plusieurs auteurs utilisent la simulation comme moyen de caractérisation des performances de leurs algorithmes d'appariement :

- SUTTON *et al.* dans [SMJB88] génèrent des images de synthèse en se basant sur une fonction sinus amorti mono-dimensionnelle d'expression :

$$S(x) = ae^{-bx} \sin(\omega x) + c$$

où  $a$  est l'amplitude de la variation d'intensité,  $b$  est un coefficient d'amortissement et  $c$  l'intensité moyenne. Ils ont choisi cette fonction car elle correspondait approximativement aux variations des niveaux de gris observés le long des lignes de leurs images expérimentales. Cette solution ne nous a pas semblé satisfaisante en terme de fidélité aux images de mouchetis que nous avons rencontrées expérimentalement. De plus, elle ne permet pas une souplesse suffisante pour la synthèse d'une texture réelle ;



**FIG. 2.7 :** Montage expérimental d'extensométrie optique bidimensionnelle utilisé par Siham TOUCHAL-MGUIL dans sa thèse [TM97] (photo reproduite avec l'aimable autorisation de Fabrice MORESTIN).

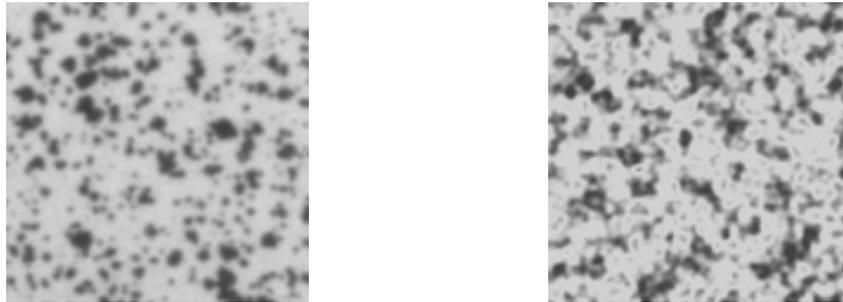
- WATTRISSE *et al.* utilisent des images « analytiques » calculées de telle manière qu'elles aient une densité de probabilité et une densité spectrale de puissance similaires de celles d'images réelles [Wat99]. Leur méthode repose sur une combinaison linéaire de fonctions gaussiennes bidimensionnelles réparties aléatoirement ;
- TARDIF *et al.* [Tar98] utilisent la méthode de THÉSING décrite dans [Thé96] pour générer des images de synthèse afin d'évaluer les performances de leur système. Les images sont produites par modélisation de l'effet SPECKLE dans l'espace de FOURIER. L'inconvénient de cette approche est qu'elle ne permet pas une grande souplesse dans la génération d'autres textures aléatoires ne provenant pas d'un effet de granularité laser ;
- SCHREIER *et al.* dans [SBS00] créent des séquences d'images artificielles en utilisant une image réelle de mouchetis qu'ils translatent en utilisant la transformée de FOURIER. Si cette approche autorise la génération d'une séquence d'images translatées précisément, elle ne permet pas en revanche d'appliquer une transformation donnée pour simuler une déformation ;
- DOUMALIN utilise dans sa thèse [Dou00] la corrélation directe pour apparier des motifs géométriques tels des croix, ou des cercles, déposées par microélectrolithographie, dans des images acquises au microscope électronique à balayage. Lors de ses évaluations de la précision d'appariement, il a utilisé plusieurs formes de motifs, y compris un mouchetis, en générant des images de synthèses construites par sur-échantillonnage d'un motif blanc sur fond noir. Ces résultats montrent par ailleurs que le motif mouchetis conduit à la meilleure précision d'appariement. L'auteur ne détaille pas la façon dont il génère le motif aléatoire.

Nous souhaitons réaliser nos simulations sur des images synthétiques, en ayant une grande

liberté sur le rendu de la texture, de sorte qu'il soit aisé de simuler le plus grand nombre de types de mouchetis. Les méthodes que nous venons de passer en revue ne nous ont pas semblé satisfaisantes par rapport à ces critères. La section suivante présente notre technique de génération d'images de synthèse de textures aléatoires de type mouchetis.

### 2.4.1 Génération des images de synthèse

Nous souhaitons générer des images d'une texture similaire à celle obtenue par projection de peintures blanche et noire — *cf.* figure 2.8 — ou, plus généralement, ressemblant à une texture granulaire qui peut être naturelle ou artificielle. Nous avons utilisé l'algorithme de

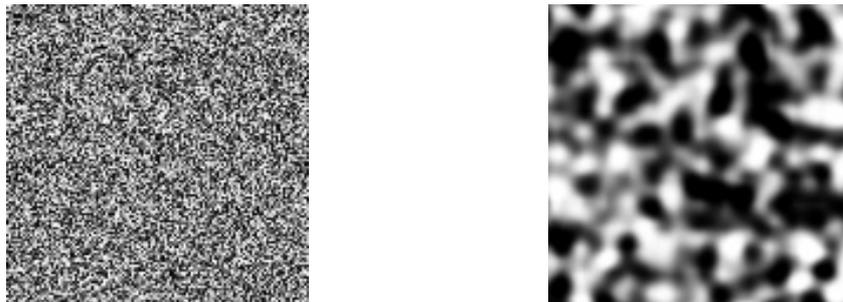


**FIG. 2.8 :** *Texture obtenue par projection de peintures blanche et noire (à gauche) et image de synthèse similaire obtenue par l'algorithme de PERLIN (à droite)*

PERLIN, décrit dans [Per85]<sup>1</sup>, qui définit une fonction *noise* de génération de bruit cohérent pour tout point  $\mathbf{m} \in \mathbb{R}^n$  :

$$\text{noise}(\mathbf{m}) : \mathbb{R}^n \mapsto [-1, 1]$$

Un bruit est dit « cohérent » si pour deux points de l'espace, la valeur de la fonction bruit varie sans discontinuité d'un point à l'autre, *i.e.* si *noise* est continuellement dérivable, *cf.* figure 2.9. Nous nous intéressons dans cette discussion aux images de textures, nous allons donc présenter



**FIG. 2.9 :** *Bruit incohérent (à gauche) et bruit de PERLIN (à droite)*

l'algorithme de PERLIN seulement dans le cas bidimensionnel ( $n = 2$ ).

---

<sup>1</sup>Il existe de très nombreuses implantations de cet algorithme. Nous nous sommes inspiré de celles des logiciels de synthèse d'images **Persistence Of Vision** et **Rayshade**, dont les codes sources sont disponibles sur internet.

Un vecteur  $\mathbf{n}$  unitaire d'orientation pseudo-aléatoire est associé à chaque point du plan  $\mathbb{Z}^2$ . Soient un point  $\mathbf{m} = (x \ y)^\top \in \mathbb{R}^2$ , et quatre points  $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4 \in \mathbb{Z}^2$ , voisins les plus proches de  $\mathbf{m}$ , avec leurs vecteurs  $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4$  associés respectifs. Soient les vecteurs  $\mathbf{d}_1 = (\delta_x \ \delta_y)^\top$ ,  $\mathbf{d}_2 = (1 - \delta_x \ \delta_y)^\top$ ,  $\mathbf{d}_3 = (\delta_x \ 1 - \delta_y)^\top$  et  $\mathbf{d}_4 = (1 - \delta_x \ 1 - \delta_y)^\top$  (cf. figure 2.10).

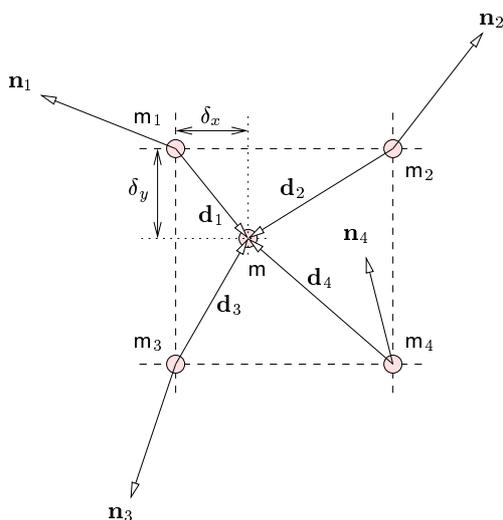


FIG. 2.10 : Notations utilisées pour la définition de la fonction de bruit.

La valeur de la fonction de bruit au point  $\mathbf{m}$  est déterminée par interpolation des quatre produits scalaires  $s_i = \mathbf{n}_i \times \mathbf{d}_i$ ,  $i = 1 \dots 4$ . PERLIN pondère les coefficients  $s_i$  avant interpolation de sorte que la fonction bruit soit d'autant plus « influencée » par un coefficient  $s_i$  lorsque le point  $\mathbf{m}$  est proche de  $\mathbf{m}_i$ . La fonction de pondération qu'il utilise est  $w(\delta) = 3\delta^2 - 2\delta^3$  (cf. figure 2.11), et son schéma d'interpolation devient :

$$\begin{aligned} s_{12} &= s_1 + w(\delta_x)(s_2 - s_1) \\ s_{34} &= s_3 + w(\delta_x)(s_4 - s_3) \\ s &= s_{12} + w(\delta_y)(s_{34} - s_{12}) \end{aligned}$$

La fonction de bruit  $noise(\mathbf{m})$  prend alors pour valeur celle du scalaire  $s$ .

La création d'une image en utilisant cette fonction de bruit conduit à un résultat similaire à celui montré figure 2.9, image droite. Pour simuler un mouchetis particulier, il est possible d'appliquer une transformation  $T$  de la fonction de bruit (cf. tableau 2.1) :

$$T(noise(\mathbf{m})) \tag{2.6}$$

Afin d'éviter le phénomène de crénelage (*aliasing*) lié au sous échantillonnage de la fonction texture, nous sur-échantillons celle-ci et calculons, pour chaque pixel, la valeur du niveau de gris moyen. Cette méthode s'apparente au principe physique d'intégration par un photo-détecteur (d'une matrice CCD) du flux de photons incidents. La figure 2.12 montre les

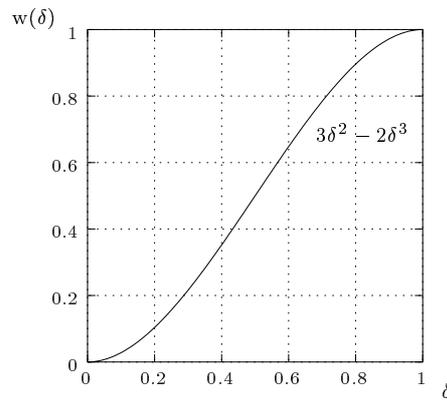


FIG. 2.11 : Fonction de pondération des produit scalaires.

coordonnées des évaluations pour un sur-échantillonnage  $4 \times 4$ , autour d'un point de coordonnées  $(u, v)$ . Pour une grille de sur-échantillonnage  $n \times n$ , le point au nœud  $(i, j)$  a pour coordonnées :

$$u_i = u + \frac{2i + 1 - n}{2n}$$

$$v_j = v + \frac{2j + 1 - n}{2n}$$

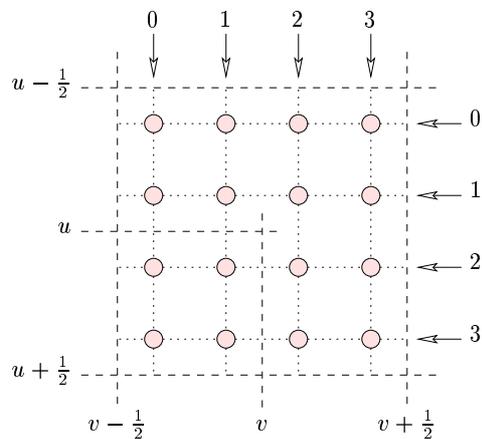
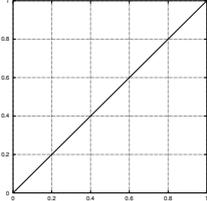
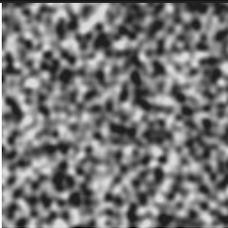
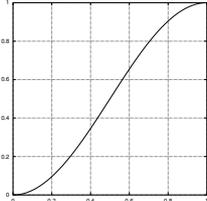
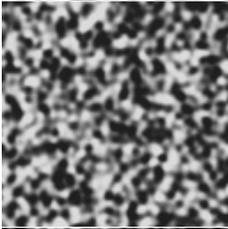
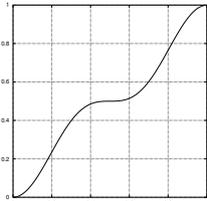
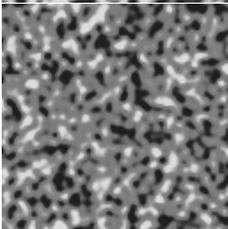
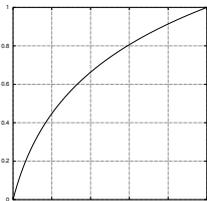
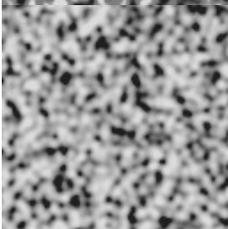
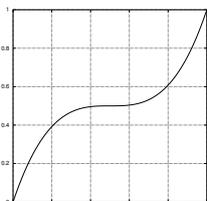
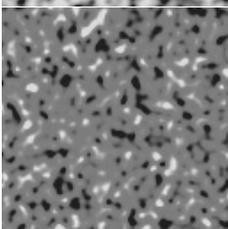
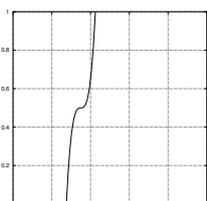
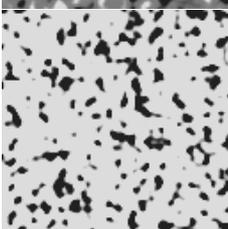


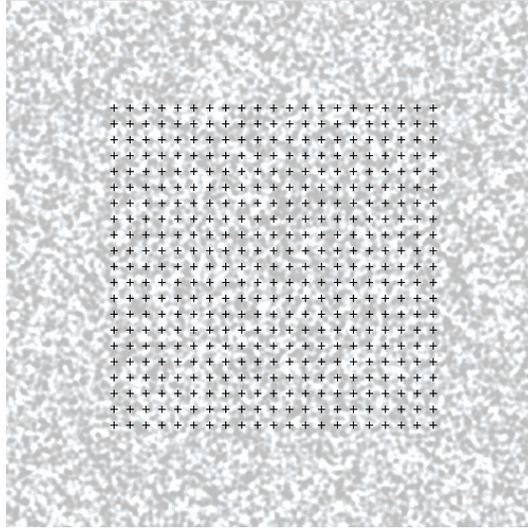
FIG. 2.12 : Sur-échantillonnage au nœud d'une grille régulière  $4 \times 4$  de la fonction texture autour du point de coordonnées  $(u, v)$ .

	$T(s)$	$T(\text{noise}(m))$
$s$		
$\frac{1}{2} [1 - \cos(\pi s)]$		
$\frac{1}{2} [1 - \cos(\pi s)^3]$		
$\log(9s + 1)$		
$4 \left(s - \frac{1}{2}\right)^3 + \frac{1}{2}$		
$\frac{4}{0.15^3} (s - 0.35)^3 + \frac{1}{2}$		

**TAB. 2.1 :** Exemples de transformations de la fonction de bruit permettant de simuler divers types de mouchetis.

### 2.4.2 Évaluation de l'incertitude sur la mesure de déplacements

L'utilisation d'images de synthèse nous permet de caractériser finement les performances théoriques de nos algorithmes de suivi de pixels par corrélation. Cette section présente les résultats obtenus pour le suivi de 441 points dans des séquences d'images synthétiques de dimensions  $500 \times 500$  pixels (*cf.* figure 2.13).



**FIG. 2.13** : Répartition des  $21 \times 21 = 441$  points utilisés pour l'évaluation des performances des algorithmes de suivi de pixels par corrélation.

Les transformations globales sont indirectement imposées par transformation des coordonnées d'évaluation de la fonction texture (2.6). Cela signifie qu'à tout point  $\mathbf{m}$  de l'image à générer correspond un niveau de gris calculé par l'équation suivante :

$$\mathbf{m} \mapsto T(\text{noise}(\mathbf{m}')) \quad \text{avec} \quad \mathbf{m}' = \Theta(\mathbf{m})$$

où  $\Theta(\mathbf{m})$  est une fonction de transformation des coordonnées  $\mathbf{m}$ . La transformation globale  $\Phi$  de l'image est donc égale à la transformation  $\Theta$  inverse :

$$\begin{aligned} \mathbf{m}' &= \Theta(\mathbf{m}) \\ \mathbf{m} &= \Phi(\mathbf{m}') \\ &\Downarrow \\ \Phi &= \Theta^{-1} \end{aligned}$$

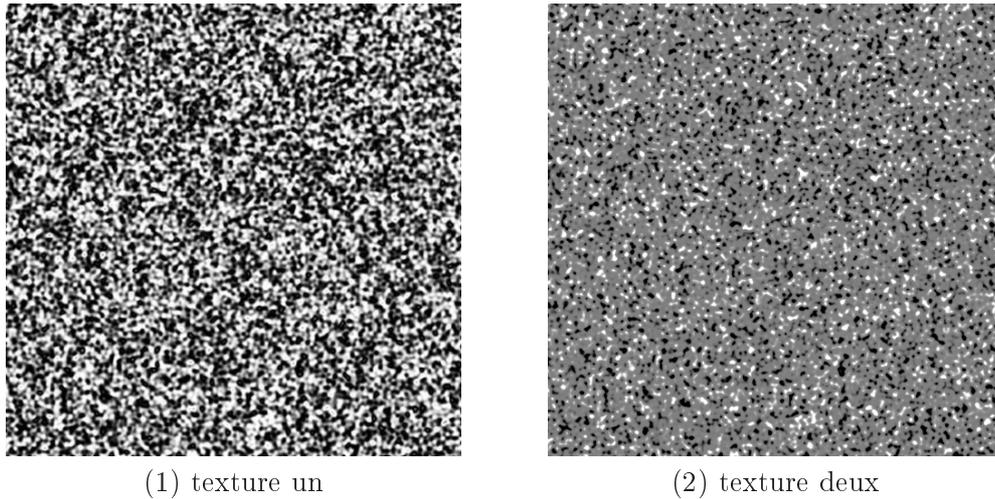
Nous allons étudier dans cette section les performances de notre méthode d'appariement par corrélation sur plusieurs séquences d'images ayant subi une transformation : translation, rotation et déformation de type « punch ». Nous étudierons ensuite l'influence du facteur de remplissage optique (problème des pixels non jointifs), paramètre technologique des capteurs de type CCD, et du bruit dans la chaîne d'acquisition des images.

#### Translations

Afin d'évaluer les performances de la méthode de suivi de pixels par corrélation dans le cas d'une translation, nous avons généré deux séquences, composée de quarante images

chacune, d'un mouchetis synthétique. Nous avons imposé un déplacement en translation selon l'axe vertical  $u$ , de  $\frac{1}{40} = 0,025$  pixels entre deux images successives. Enfin, nous avons sur-échantillonné la fonction texture par  $40 \times 40$  évaluations par pixels selon le schéma de la figure 2.12.

La première séquence d'images (*cf.* figure 2.14-1) a une distribution uniforme de ses niveaux de gris, alors que la deuxième séquence possède une distribution tri-modale (*cf.* figure 2.14-2) : trois pics correspondants aux niveaux noir (0), gris (128) et blanc (255).

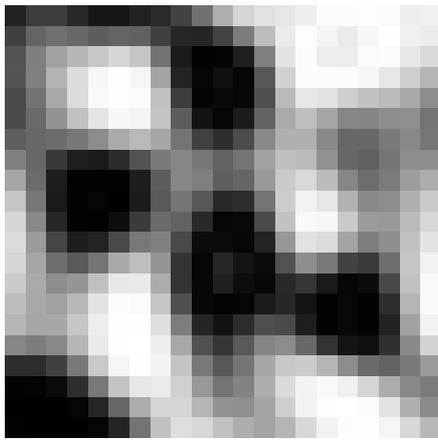


**FIG. 2.14 :** Première image des deux séquences générées : distribution des niveaux de gris (1) uniforme et (2) tri-modale.

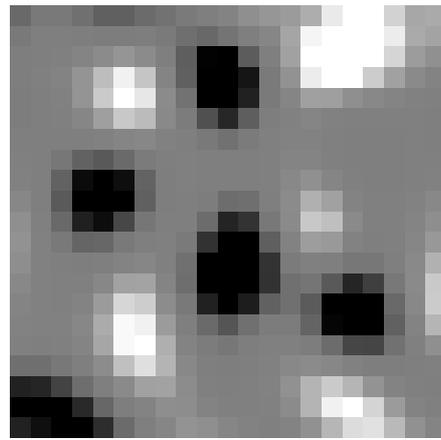
Nous avons utilisé le critère de corrélation SSD avec une taille de fenêtre de  $21 \times 21$  pixels, et un processus de minimisation basé sur l'algorithme de LEVENBERG-MARQUARDT pour estimer précisément les seuls paramètres de translation (modèle de la transformation locale d'ordre 0). La figure 2.15 montre en détail le contenu d'une des fenêtres de corrélation pour les deux textures.

Le tableau 2.2 montre les résultats obtenus pour les deux séquences d'images des textures 1 et 2. Hormis le cas de l'interpolation bilinéaire, l'erreur moyenne (ou erreur systématique) présente une symétrie autour du déplacement de 0,5 pixel. Si l'on compare les résultats obtenus pour chacune des deux textures, il apparaît clairement que la texture 2 conduit à des résultats moins précis que ceux obtenus pour la texture 1. Récemment, SCHREIER *et al.* ont montré dans une étude approfondie [SBS00] que cette symétrie et la dégradation de la précision dans le cas de la texture 2 ont pour origine l'interpolant des niveaux de gris utilisé : l'utilisation d'un interpolant spline altère à la fois l'amplitude et la phase du signal en fonction de la position subpixel et de l'harmonique du signal considérés. DOUMALIN dans sa thèse [Dou00] observe le même phénomène périodique et symétrique pour des motifs géométriques (cercles, croix, ...), en utilisant une interpolation bilinéaire. Il en va de même pour LARABA-ABBES qui, dans sa thèse [LA98], utilise une interpolation polynomiale bicubique et montre, lors d'un essai expérimental, la périodicité unitaire et le caractère symétrique de cette erreur systématique.

Nous retrouvons donc les résultats observés par SCHREIER *et al.* : l'augmentation de l'ordre de l'interpolant spline utilisé permet de réduire l'erreur systématique du suivi de pixels par corrélation. Cette réduction est la plus significative lorsque l'on passe d'un interpolant bilinéaire



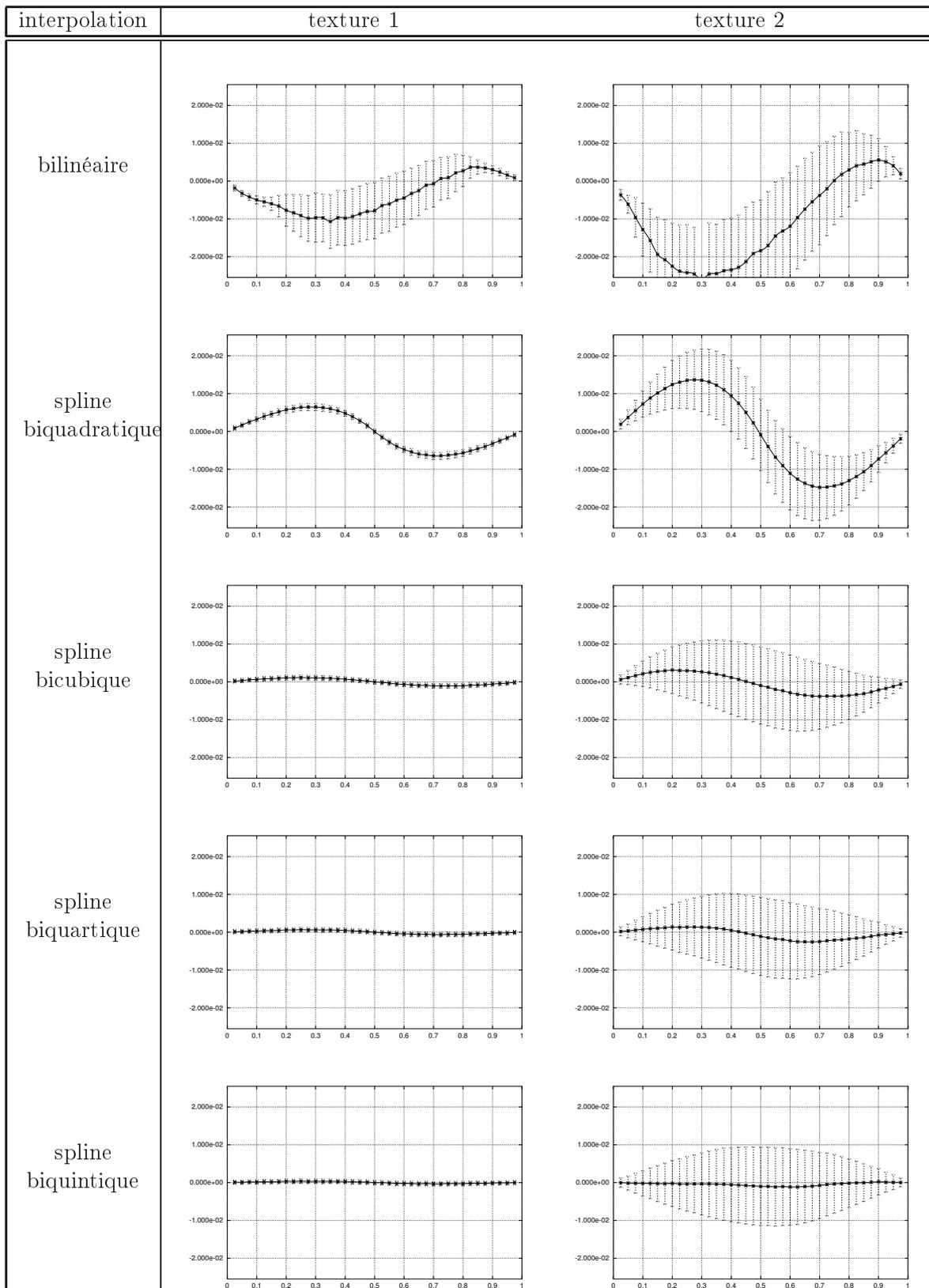
(1) texture un



(2) texture deux

**FIG. 2.15 :** *Exemple de contenu des fenêtres de corrélation de taille  $21 \times 21$  pixels pour (1) la première texture, et pour (2) la deuxième texture.*

à un interpolant spline bicubique.



**TAB. 2.2 :** Précision du suivi de pixels par corrélation pour chacune des méthodes d'interpolation. L'axe horizontal des courbes représente la translation (en pixel) imposée, et l'axe vertical indique l'erreur moyenne d'estimation de ce déplacement (en pixel), ainsi que son écart-type.

## Rotations

L'essai comporte dix-huit images de synthèse, pour lesquelles nous avons appliqué une rotation de cinq degrés entre deux images successives, conduisant à une rotation maximale de 85 degrés par rapport à la première image (identique à celle montrée figure 2.14-1). Toutes les images ont été sur-échantillonnées par  $40 \times 40$  évaluations par pixel.

Le critère de corrélation SSD avec une taille de fenêtre de  $21 \times 21$  pixels a été utilisé. Le modèle de transformation locale utilisé est d'ordre 1 et l'estimation de ses paramètres est réalisée par l'algorithme de minimisation de LEVENBERG-MARQUARDT. Le tableau 2.3 présente

interpolant spline	paramètres de la transformation locale d'ordre 1						rotation (deg.)
	$t_u$ (pix.)	$t_v$ (pix.)	$a_u$	$a_v$	$b_u$	$b_v$	
bilinéaire	3,767e-03	3,882e-03	8,014e-04	8,038e-04	8,087e-04	8,195e-04	3,186e-02
biquadratique	7,441e-04	7,662e-04	1,428e-04	1,436e-04	1,425e-04	1,443e-04	5,943e-03
bicubique	5,821e-04	5,990e-04	1,060e-04	1,069e-04	1,057e-04	1,057e-04	4,132e-03
biquartique	5,587e-04	5,795e-04	9,981e-05	1,004e-04	9,982e-05	9,909e-05	3,901e-03
biquintique	5,540e-04	5,712e-04	9,844e-05	9,895e-05	9,736e-05	9,704e-05	3,833e-03

**TAB. 2.3 :** Écart-types moyens des paramètres de la transformation locale d'ordre 1 et des angles de rotation estimés, en fonction de l'interpolant utilisé.

les résultats de cette simulation. Les erreurs moyennes n'y sont pas reportées car elles sont nulles. Cela était prévisible étant donné que le seul biais attendu provient des erreurs systématiques liées aux positions sub-pixels des points homologues réels (*cf.* section précédente), or la rotation des 441 points conduit à des positions sub-pixels symétriquement réparties entre  $-0,5$  et  $0,5$  pixel. De plus, les écarts-types des erreurs d'estimation sont très similaires quel que soit l'angle de rotation. Nous avons donc reporté dans ce tableau les moyennes, sur la séquence d'images, des écarts-types des paramètres estimés. Afin d'être en mesure d'apprécier plus finement ces résultats, nous avons cherché à estimer les angles de rotation en fonction des paramètres  $a_u$ ,  $a_v$ ,  $b_u$  et  $b_v$ . La transformation locale au premier ordre a pour expression :

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} t_u \\ t_v \end{pmatrix} + \begin{pmatrix} a_u & a_v \\ b_u & b_v \end{pmatrix} \begin{pmatrix} u - u_c \\ v - v_c \end{pmatrix}$$

Identifions les termes de cette équation avec l'expression de la transformation globale :

$$\begin{aligned} \begin{pmatrix} u' \\ v' \end{pmatrix} &= \mathbf{R} \begin{pmatrix} u \\ v \end{pmatrix} \\ &= \begin{pmatrix} u \\ v \end{pmatrix} + (\mathbf{I} - \mathbf{R}) \begin{pmatrix} u_c \\ v_c \end{pmatrix} + (\mathbf{R} - \mathbf{I}) \begin{pmatrix} u - u_c \\ v - v_c \end{pmatrix} \end{aligned}$$

Il vient :

$$\underbrace{\begin{pmatrix} 1 + a_u & a_v \\ b_u & 1 + b_v \end{pmatrix}}_{\hat{\mathbf{R}}} = \mathbf{R}$$

L'estimation des paramètres  $a_u$ ,  $a_v$ ,  $b_u$  et  $b_v$  est effectuée sans prise en compte de la contrainte d'orthonormalité d'une matrice de rotation. La matrice  $\hat{\mathbf{R}}$  ainsi estimée n'est donc pas orthonormale. Nous avons cherché à déterminer la matrice de rotation  $\mathbf{R}$  la plus « proche » de  $\hat{\mathbf{R}}$

au sens de la norme de FROBENIUS de leur différence, *i.e.* en résolvant  $\min_{\mathbf{R}} \|\mathbf{R} - \widehat{\mathbf{R}}\|_F$  sous les contraintes d'orthonormalité de  $\mathbf{R}$ . Ce problème classique trouve une solution élégante en utilisant la décomposition en valeurs singulières de  $\widehat{\mathbf{R}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ . La matrice  $\mathbf{R}$  est alors déterminée par :  $\mathbf{R} = \mathbf{U}\mathbf{V}^\top$ . Il est alors aisé de calculer l'angle de rotation en fonction des éléments de  $\mathbf{R}$ . La colonne de droite du tableau 2.3 présente la moyenne des écarts-types des estimations des angles de rotations, en utilisant cette méthode.

Le tableau 2.3 montre que l'utilisation d'un interpolant spline biquintique permet d'améliorer d'un facteur 8 environ la précision d'estimation des paramètres de la transformation locale par rapport à une simple interpolation bilinéaire des niveaux de gris. On retrouve ce même facteur d'amélioration dans l'estimation de l'angle de rotation par la méthode exposée ci-avant. Il est alors intéressant de constater que l'on peut obtenir une estimation de cet angle avec un écart-type de l'erreur aussi faible que  $\frac{1}{250}$  degré.

### Déformations de type *punch*

Nous définissons une déformation de type *punch* par la transformation  $\Theta$  suivante :

$$\Theta(\mathbf{m}) = \frac{1 + p(u^2 + v^2)}{1 + ph^2} \begin{pmatrix} u \\ v \end{pmatrix}$$

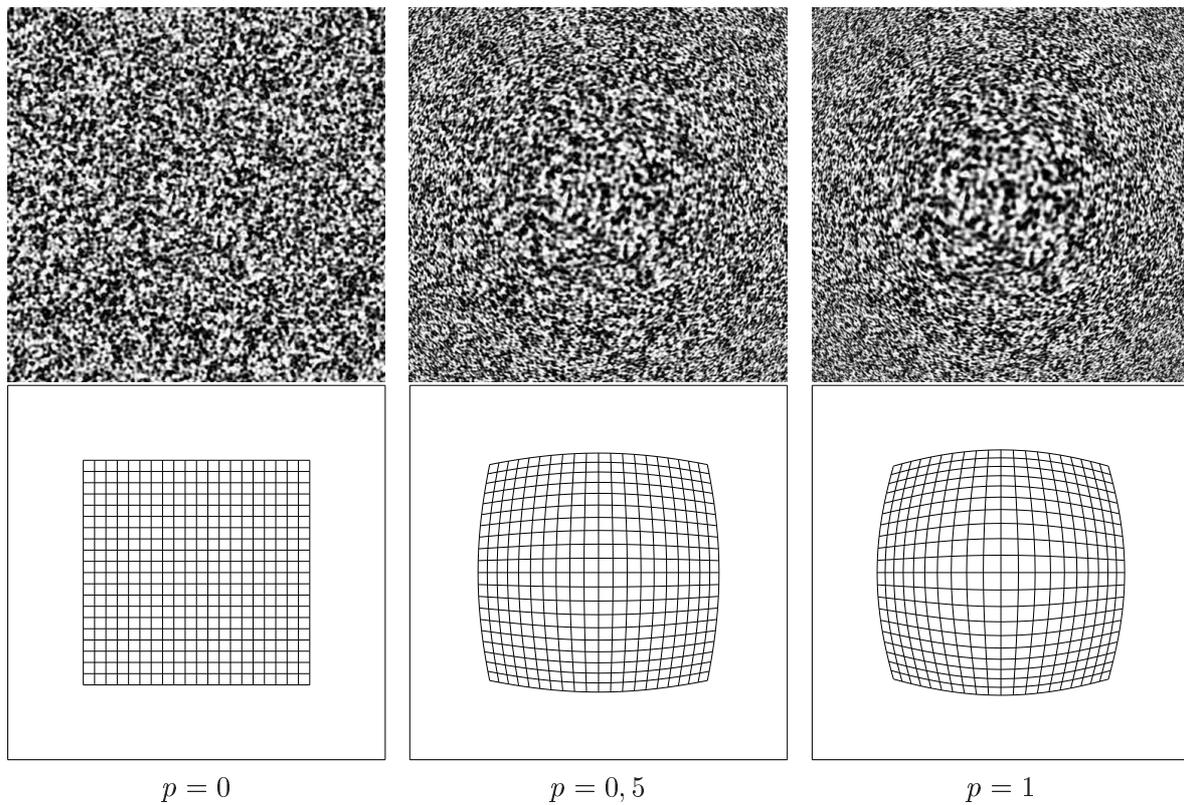
où les coordonnées  $u$  et  $v$  sont normalisées dans l'intervalle  $[-1 \dots 1]$ . Le paramètre  $p$  définit l'importance de la déformation et  $h$  est déterminé de telle manière que tout point  $\mathbf{m}$  distant de l'origine d'une valeur  $h$  soit invariant par la transformation  $\Theta$ .

La séquence d'images correspond à une variation du paramètre  $p$  de 0 à 1 par incrément de  $\frac{50}{1000}$ , soit 21 images, et à une valeur fixe  $h = 0,75$ . La figure 2.16 montre les images et la position des 441 vrais points<sup>2</sup> pour les valeurs  $p = 0$ ,  $p = 0,5$  et  $p = 1$ .

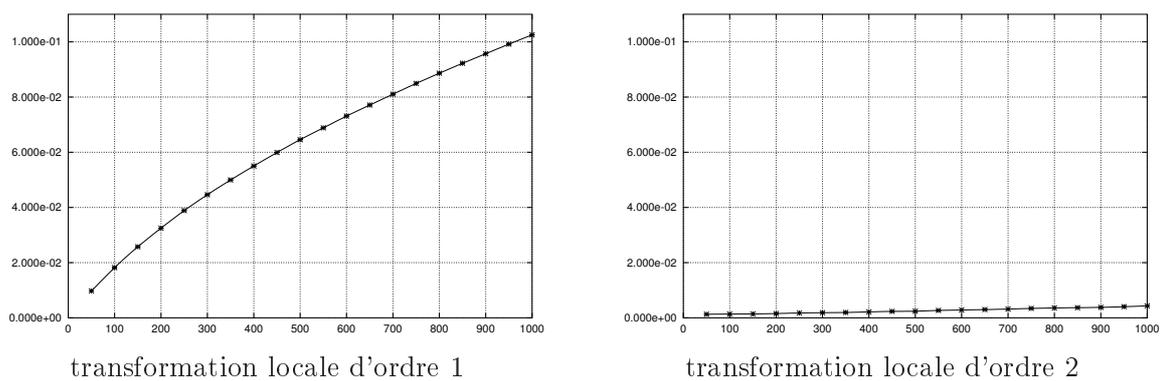
La figure 2.17 montre les écarts-types des erreurs d'appariement pour la coordonnée  $u$  (les erreurs selon l'axe  $v$  sont similaires). La moyenne des erreurs n'est pas représentée car elle est nulle pour les mêmes raisons citées section précédente. Les gradients de déplacements étant élevés, une approximation de la transformation locale par un modèle du premier ordre conduit naturellement à une précision dégradée par rapport à une approximation au second ordre (d'un facteur 25 environ). La figure 2.18 illustre ce propos en montrant les erreurs d'appariement pour chacun des points suivis, en fonction de l'ordre du modèle de la transformation géométrique locale. Il y apparaît clairement l'amélioration apportée par l'utilisation d'un modèle d'ordre 2 par rapport à un modèle d'ordre 1 pour ce type de déformation.

---

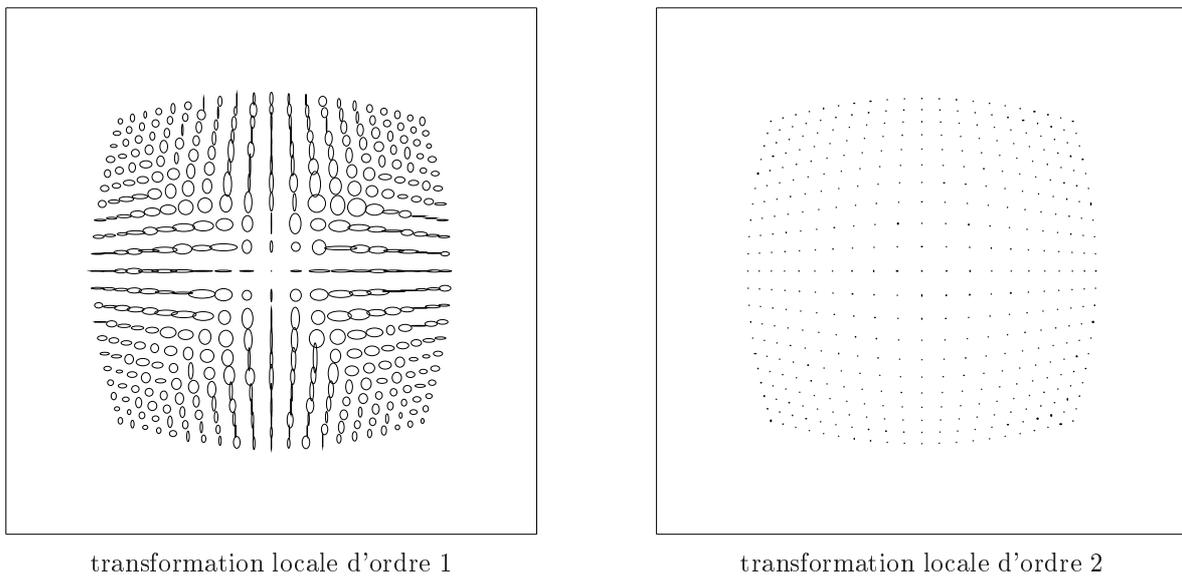
<sup>2</sup>Ces points ont été déterminés par l'évaluation de la transformation  $\Theta$  inverse sur les 441 points de l'image de référence. L'inversion de  $\Theta$ , et donc la détermination de la transformation globale  $\Phi$ , passe par la recherche des racines d'un polynôme cubique. Voir [Gar96] pour la résolution d'un problème similaire.



**FIG. 2.16 :** Séquence d'images punch. En haut : image de référence ( $p = 0$ ), onzième image ( $p = 0,5$ ) et dernière image ( $p = 1$ ) de la séquence. En bas : positions des 441 points suivis dans la séquence.



**FIG. 2.17 :** Comparaison des écarts-types des erreurs, selon l'axe  $u$ , de suivi de pixels par corrélation pour une déformation de type punch selon l'ordre du modèle de la transformation locale (l'interpolation des niveaux de gris est faite par spline biquintique). L'échelle horizontale représente les valeurs du paramètres  $p$  ( $\times 1000$ ), l'échelle verticale est en pixel.



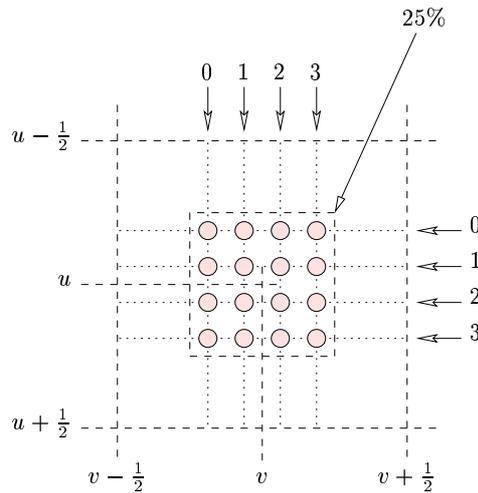
transformation locale d'ordre 1

transformation locale d'ordre 2

**FIG. 2.18** : Erreurs de suivi des 441 points par corrélation selon l'ordre du modèle de la transformation locale, pour la dernière image de la séquence punch. Les hauteurs, respectivement les largeurs, des ellipses représentent les valeurs absolues multipliées par 50 des erreurs selon l'axe  $u$ , respectivement  $v$ .

### Facteur de remplissage optique des CCD

La technologie CCD à transfert interligne induit un facteur de remplissage optique d'environ 20% à 25% seulement (source : Olympus America, Inc.), et jusqu'à 70% dès lors que des micro-lentilles sont adjointes à chaque photodiode. Le facteur de remplissage optique n'est jamais égal à l'unité car les photo-détecteurs ne sont jamais parfaitement jointifs.



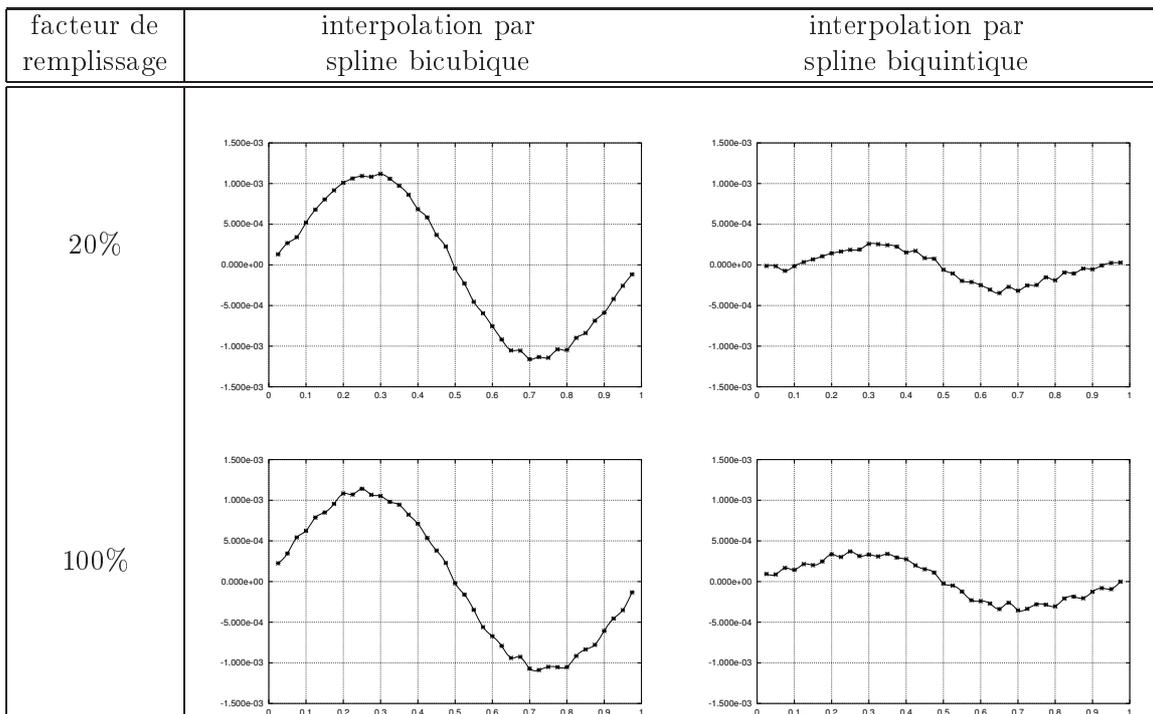
**FIG. 2.19** : Sur-échantillonnage au nœud d'une grille régulière  $4 \times 4$  de la fonction texture autour du point de coordonnées  $(u, v)$  pour un facteur de remplissage optique de 25%.

Nous avons simulé l'influence de ce facteur de remplissage optique sur les incertitudes de la corrélation dans le cas d'une translation. Les séquences d'images utilisées sont similaires à celle de l'essai en translation, en début de section, mais avec un facteur de remplissage variant de 20% à 100% par incrément de 20%. La figure 2.19 montre le schéma d'intégration utilisé pour simuler le facteur de remplissage optique. Pour une valeur de ce dernier égal à  $f$ , et une grille de sur-échantillonnage  $n \times n$ , le point au nœud  $(i, j)$  a pour coordonnées :

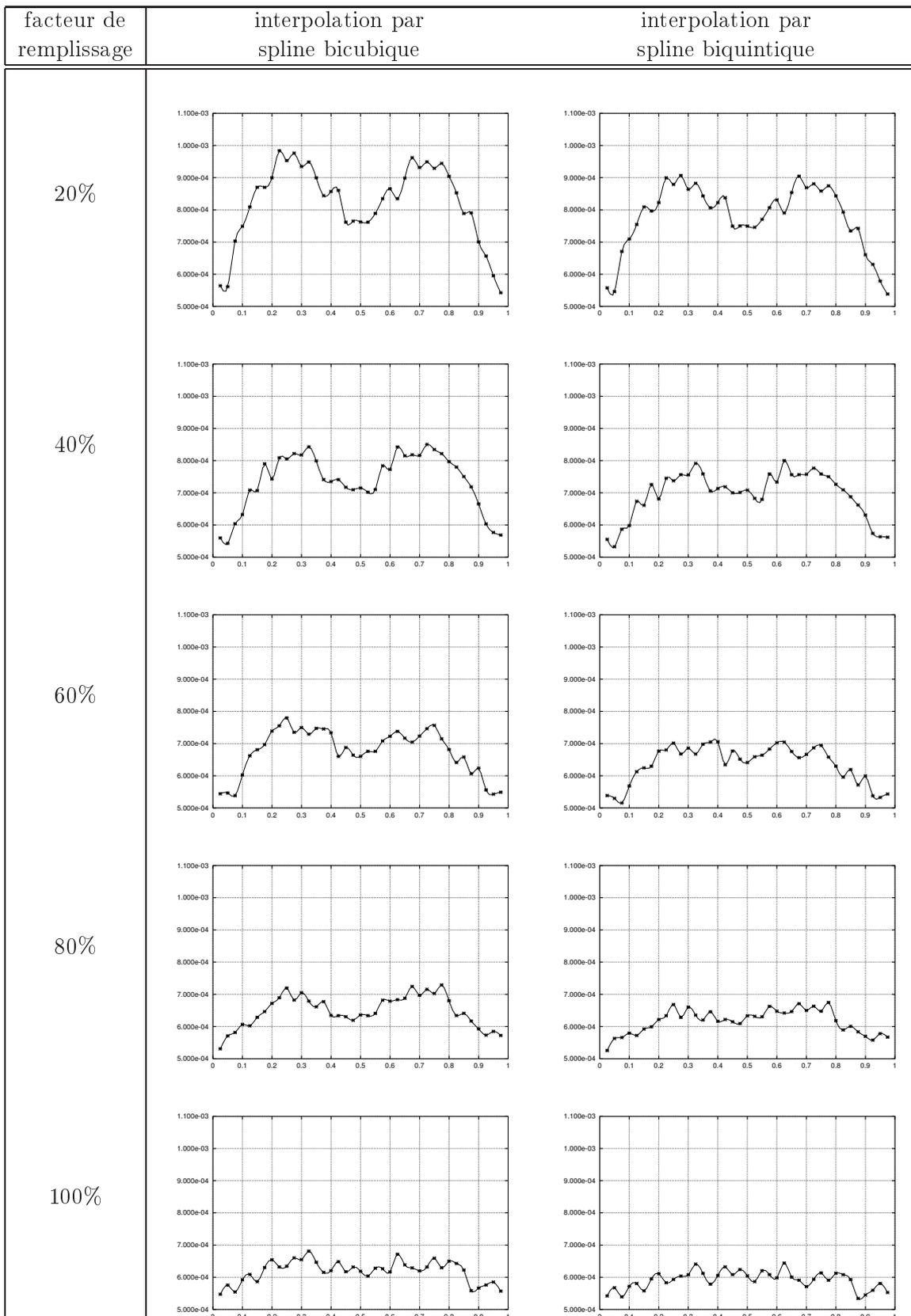
$$u_i = u + \sqrt{f} \frac{2i + 1 - n}{2n}$$

$$v_j = v + \sqrt{f} \frac{2j + 1 - n}{2n}$$

La table 2.4 montre l'erreur moyenne de corrélation, ou erreur systématique, pour deux facteurs de remplissage : 20% et 100%. Ces résultats montrent que le facteur de remplissage optique ne semble pas influencer l'erreur systématique. En revanche, nous pouvons observer table 2.5 que celui-ci dégrade légèrement l'écart-type de l'erreur, ou erreur aléatoire, de l'appariement par corrélation pour des facteurs de remplissage optique de plus en plus faibles : près de 10% plus élevé lorsque le facteur de remplissage optique passe de 100% à 20%. Il est donc souhaitable d'utiliser des caméras ayant un facteur de remplissage optique le plus élevé possible afin de réduire l'erreur aléatoire d'appariement liée à ce paramètre technologique.



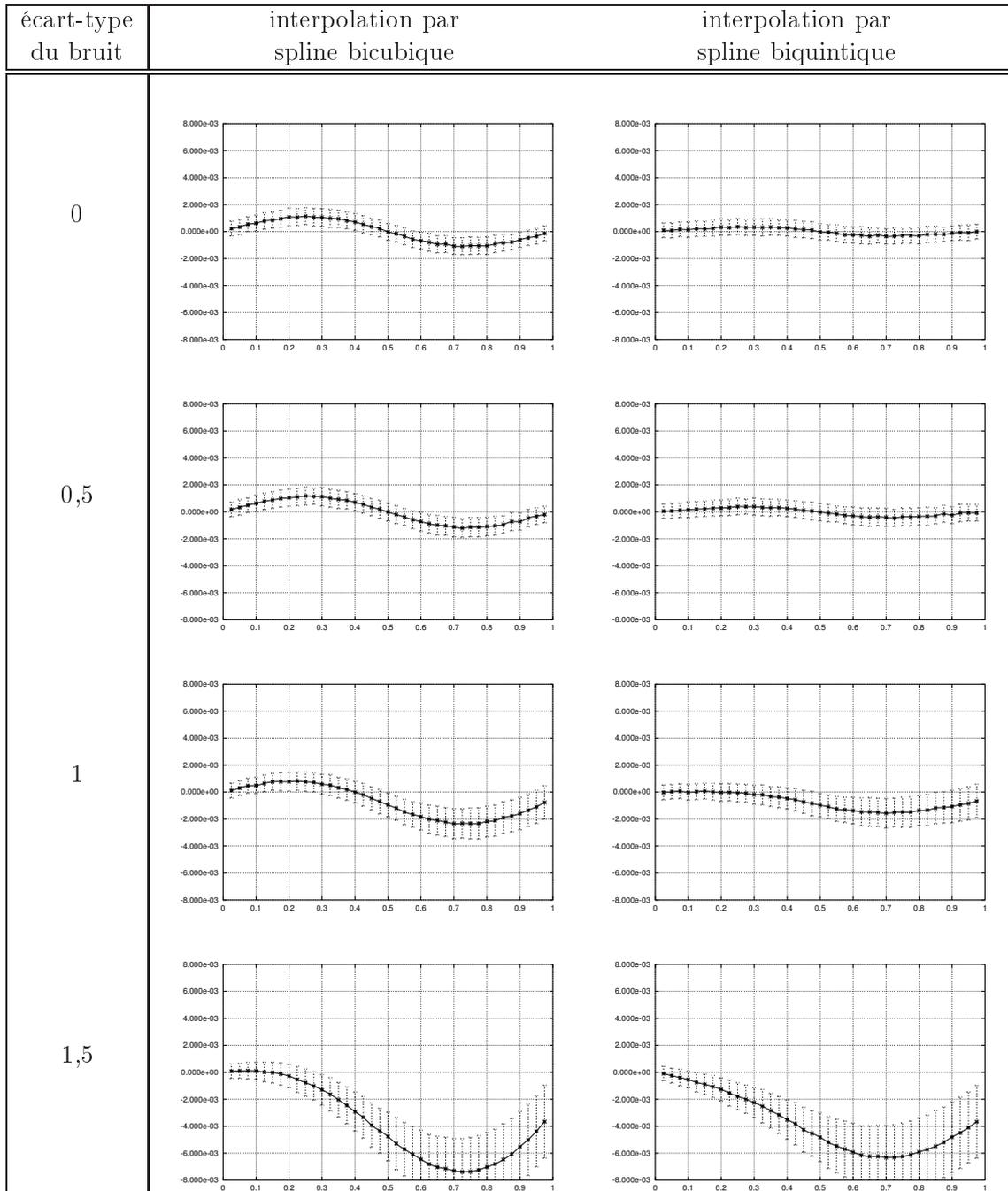
**TAB. 2.4 :** Influence du facteur de remplissage optique sur l'erreur systématique d'appariement pour une séquence d'images en translation. Quel que soit le degré de l'interpolant des niveaux de gris utilisé, on constate que l'erreur systématique est indépendante du facteur de remplissage optique.



**TAB. 2.5 :** Influence du facteur de remplissage optique sur l'erreur aléatoire d'appariement pour une séquence d'images en translation. L'analyse de ces courbes nous montre une augmentation de près de 10% de l'erreur aléatoire lorsque ce paramètre technologique passe de 100% à 20%.

## Bruit

Nous avons étudié également l'influence d'un bruit gaussien additif dans les niveaux de gris des images, sur la précision d'appariement par corrélation dans le cas d'une séquence d'images en translation. Le tableau 2.6 illustre les résultats obtenus pour les erreurs systématique et aléatoire d'appariement. Lorsque le bruit gaussien a un écart-type de 0,5 niveau de gris, on observe le comportement classique déjà mentionné dans les sections précédentes, à savoir une erreur aléatoire indépendante et une erreur systématique fonction de la translation imposée. En revanche, dès lors que le bruit dans les niveaux de gris atteint ou dépasse un écart-type unitaire, il apparaît que l'erreur aléatoire d'appariement ne présente plus un caractère symétrique autour du déplacement de 0,5 pixel. Ce phénomène s'explique par une convergence des algorithmes d'optimisation utilisés de plus en plus difficile au fur et à mesure que la translation imposée augmente. En effet, pour chaque translation imposée, chacune des estimées initiales, fournie à l'algorithme d'optimisation, est déterminée en considérant une translation nulle. Ces estimées initiales sont donc de plus en plus éloignées de la translation réelle au fur et à mesure que celle-ci augmente. Il faut toutefois nuancer ces résultats en considérant que dans le cas d'une séquence d'images réelles en translation, les estimées initiales seraient fournies via un algorithme de corrélation classique au pixel près qui estimerait une translation nulle lorsque celle-ci est comprise entre 0 et 0,5 pixel, et une translation d'un pixel lorsqu'elle est comprise réellement entre 0,5 et 1 pixel. Cela conduirait à une symétrie des résultats du tableau 2.6 par rapport à la translation imposée de 0,5 pixel.



**TAB. 2.6 :** Influence de l'écart-type d'un bruit gaussien additif dans les images sur la précision de corrélation pour une séquence de translation d'un pixel le long de l'axe vertical.

## 2.5 Conclusion

Nous venons d'étudier une méthode précise d'appariement par corrélation qui utilise des domaines de corrélation dont la forme s'adapte dynamiquement à l'information présente dans les images par la prise en compte de la transformation locale de la texture. Elle met en œuvre une technique d'interpolation des niveaux de gris dans les images par surfaces B-splines et l'on a pu montrer que l'utilisation de surfaces B-splines d'ordre supérieur ou égal à trois améliore significativement la précision d'appariement.

Par rapport aux méthodes classiques de suivi de marqueurs, l'appariement par corrélation présente de nombreux avantages :

- la facilité de préparation de la surface de l'objet (projection de peinture ou pulvérisation de *toner* d'imprimante en quelques secondes) lorsque cela est nécessaire ;
- en théorie, chacun des pixels de l'image de référence peut être suivi, ce qui conduit à l'obtention d'un champ de déplacements dense.

Toutefois, l'extensométrie optique bidimensionnelle, qui est maintenant couramment utilisée pour des mesures de champs de déplacements bidimensionnels, présente deux inconvénients majeurs : d'abord, elle ne permet de mesurer que des déplacements liés à une déformation plane, ensuite, d'un point de vue expérimental, elle nécessite que le plan de déformation et le plan image de la caméra soient parallèles et le restent pendant l'essai, ce qui est difficile à garantir. Ce sont les raisons qui nous ont amené à nous intéresser à la technique de stéréovision présentée dans le chapitre suivant.

## Bibliographie du chapitre

- [AG92] P. ASCHWANDEN et W. GUGGENBÜHL. Experimental results from a comparative study on correlation-type registration algorithms. Dans *Robust Computer Vision*, 1992.
- [BD94] T. BLASZKA et R. DERICHE. Recovering and Characterizing Image Features Using an Efficient Model Based Approach. Rapport de recherche 2422, INRIA, novembre 1994.
- [BDL95] F. BRÉMAND, J. C. DUPRÉ, et A. LAGARDE. Mesure de déformations sans contact par analyse d'images. Dans *Photomécanique'95*, ENS Cachan (France), mars 1995.
- [BMSP89] H. BRUCK, S. MCNEILL, M. SUTTON, et W. PETERS. Digital Image Correlation Using Newton-Raphson Method of Partial Differential Correction. Dans *Experimental Mechanics*, volume 29, pp. 261–267, avril 1989.
- [Bra95] Pascal BRAND. *Reconstruction tridimensionnelle à partir d'une caméra en mouvement : de l'influence de la précision*. Thèse de doctorat, Université Claude Bernard – Lyon I, octobre 1995.
- [CRSP86] T. C. CHU, W. F. RANSON, M. A. SUTTON, et W. H. PETERS. Application of Digital-Image-Correlation Techniques to Experimental Mechanics. Dans *Experimental Mechanics*, volume 25, pp. 232–244, 1986.
- [Dou00] Pascal DOUMALIN. *Microextensométrie locale par corrélation d'images numériques. Application aux études micromécaniques par microscopie électronique à balayage*. Thèse de doctorat, École Polytechnique (France), juin 2000.
- [FHM<sup>+</sup>93] O. FAUGERAS, B. HOTZ, H. MATHIEU, T. VIÉVILLE, Z. ZHANG, P. FUA, E. THÉRON, L. MOLL, G. BERRY, J. VUILLEMIN, P. BERTIN, et C. PROY. Real time correlation-based stereo : algorithm, implementations and applications. Rapport de recherche 2013, INRIA, août 1993.
- [Gar96] Vincent GARRIC. *Vision pour la robotique de manipulation : calibration, localisation et saisie d'objets*. Thèse de doctorat, Institut National des Sciences Appliquées de Toulouse (France), mai 1996.
- [GMN<sup>+</sup>98] B. GALVIN, B. MCCANE, K. NOVINS, D. MASON, et S. MILLS. Recovering Motion Fields : An Evaluation of Eight Optical Flow Algorithms. Dans *British Machine Vision Conference (BMVC'98)*, septembre 1998.
- [GRMT98] M. GASPÉRINI, A. RAZAKAINAVO, M. MOREL, et C. TEODOSIU. Mesure par microgrilles des déformations en étirage plan dans l'épaisseur de tôles minces d'alliage d'aluminium. Dans *Photomécanique'98*, Marne la Vallée (France), avril 1998.
- [HEG01] J. M. HIVER, J. ETIENNE, et C. G'SELL. Mesure de l'endommagement plastique dans une striction par un système d'analyse d'image. Dans *Photomécanique'2001*, Futuroscope, Poitiers (France), avril 2001.
- [HS81] K. P. HORN et B. G. SCHUNCK. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [LA98] Fazilay LARABA-ABBES. *Étude des comportements hyperélastique et viscohyperélastique de deux élastomères de type NR et PDMS par extensométrie optique bidimensionnelle*. Thèse de doctorat, École Centrale de Paris (France), février 1998.

- [Lan97] Zhong-Dan LAN. *Méthodes robustes en vision : application aux appariements visuels*. Thèse de doctorat, Institut National Polytechnique de Grenoble (France), mai 1997.
- [LVD99] J.M. LAVEST, M. VIALA, et M. DHOME. Quelle précision pour une mire d'éta-  
lonnage ? *Traitement du Signal*, 16(3), 1999.
- [MBB95] F. MORESTIN, P. BOGAERT, et M. BOIVIN. Mesure de déformations par imagerie :  
détermination des courbes limites de formage. Dans *Photomécanique '95*, ENS  
Cachan (France), mars 1995.
- [MBL93] D. W. MANTHEY, R. M. BASSETTE, et D. LEE. Comparison of Different Surface  
Strain Measurement Techniques Used for Stamped Sheet Metal Parts. Dans *In-  
ternational Body Engineering Conference : Body Assembly & Manufacturing*, pp.  
106–111, Detroit, MI (USA), septembre 1993.
- [MD92] Z. MARCINIAK et J. L. DUNCAN. *The Mechanics of Sheet Metal Forming*. Edward  
Arnold, 1992.
- [Mon98] Jérôme MONTEIL. *Traitement et analyse par une technique adaptée de flux optique  
du champ de déformation de matériaux à partir de séquence d'images de micro-  
scopie*. Thèse de doctorat, Université Denis Diderot — Paris VII (France), mars  
1998.
- [NGIG95] M. NÉMOZ-GAILLARD, P. IENNY, et J. GERMAIN. Extensométrie bidimension-  
nelle par analyse d'images. Dans *Photomécanique '95*, ENS Cachan (France), mars  
1995.
- [Per85] Ken PERLIN. An Image Synthesizer. Dans *SIGGRAPH'85*, pp. 287–296, 1985.
- [Peu94] Bernard PEUCHOT. Utilisation de détecteurs subpixels dans la modélisation d'une  
caméra. Dans *Actes du 9ème Congrès AFCET (RFIA '94)*, pp. 691–695, Paris  
(France), janvier 1994.
- [PTVF92] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, et B. P. FLANNERY. *Nu-  
merical Recipes in C — The Art of Scientific Computing*. Cambridge University  
Press, édition 2nd, 1992.
- [SBS00] H. SCHREIER, J. BRAASCH, et M. SUTTON. Systematic errors in digital image  
correlation caused by intensity interpolation. *Optical Engineering*, 39(11):2915–  
2921, novembre 2000.
- [SDC86] R. SOWERBY, J. L. DUNCAN, et E. CHU. The Modelling of Sheet Metal Stamping.  
*International Journal of Mechanical Science*, 28(7):415–430, 1986.
- [She96] Jonathan Richard SHEWCHUK. Triangle : Engineering a 2D Quality Mesh Genera-  
tor and Delaunay Triangulator. Dans *Applied Computational Geometry : Towards  
Geometric Engineering*, volume 1148 de *Lecture Notes in Computer Science*, pp.  
203–222. Springer-Verlag, mai 1996.
- [SMHS00] M. A. SUTTON, S. R. MCNEILL, J. D. HELM, et H. W. SCHREIER. Compu-  
ter vision applied to shape and deformation measurement. Dans *International  
Conference on Trends in Optical Nondestructive Testing and Inspection, Elsevier  
Science*, pp. 571–589, Lugano (Suisse), mai 2000.
- [SMJB88] M. SUTTON, S. MCNEILL, J. JANG, et M. BABAI. Effects of subpixel image  
restoration on digital correlation error estimates. *Optical Engineering*, 27(10):870–  
877, octobre 1988.

- [Tar98] Fabrice TARDIF. *Analyse des mécanismes d'endommagement dans les composites 3D C/C — Mesures optiques des champs de déformation par corrélation d'images*. Thèse de doctorat, Université Bordeaux I (France), novembre 1998.
- [TBBD01] R. TIE BI, N. BRETAGNE, et J. C. DUPRÉ. Mesures par suivi de marqueurs et thermographie infrarouge de paramètres thermomécaniques. Dans *Photomécanique'2001*, Futuroscope, Poitiers (France), avril 2001.
- [Thé96] Jan THÉSING. *Mesure par corrélation des déplacements et des rotations à partir de figures granulaires*. Mémoire de fin d'études, LMT Cachan (France), 1996.
- [TM97] Siham TOUCHAL-MGUIL. *Une technique de corrélation directe d'images numériques : application à la détermination de courbes limites de formage et proposition d'un critère de striction*. Thèse de doctorat, Institut National des Sciences Appliquées de Lyon (France), juillet 1997.
- [UAE93a] M. UNSER, A. ALDROUBI, et M. EDEN. B-spline processing : part I — theory. *IEEE Transactions on Signal Processing*, 41(2):821–833, 1993.
- [UAE93b] M. UNSER, A. ALDROUBI, et M. EDEN. B-spline processing : part II — efficient design and applications. *IEEE Transactions on Signal Processing*, 41(2):834–847, 1993.
- [Uns99] Michael UNSER. Splines, A perfect Fit for Signal and Image Processing. *IEEE Signal Processing Magazine*, 16(6):22–28, novembre 1999.
- [Wat99] Bertrand WATTRISE. *Étude cinématique des phénomènes de localisation dans des aciers par intercorrélation d'image*. Thèse de doctorat, Université Montpellier II (France), février 1999.
- [Yan92] Ying YANG. *Mesure du champ de déplacement des matériaux sous charge mécanique par analyse des images*. Thèse de doctorat, École Nationale des Ponts et Chaussées (France), septembre 1992.



## Chapitre 3

# Mesure de formes tridimensionnelles par stéréo-corrélation

Nous allons décrire dans ce chapitre une méthode permettant de mesurer la forme tridimensionnelle d'objets par stéréovision et en particulier par stéréo-corrélation. Nous supposons le capteur stéréoscopique fortement calibré (*cf.* chapitre 1). Ce problème a déjà fait l'objet de nombreux travaux de recherche depuis ces trois dernières décennies, mais peu se sont véritablement intéressés à l'obtention d'une mesure de grande précision, ce qui constitue l'objet de ce chapitre.

La mesure de forme par stéréovision est basée sur le principe de triangulation déjà exposé au premier chapitre : étant donné deux points homologues dans les deux images de la paire stéréoscopique, il leur correspond deux rayons optiques dans l'espace dont l'intersection conduit aux coordonnées tridimensionnelles du point de la scène correspondant. La triangulation exige par conséquent de résoudre au préalable deux problèmes fondamentaux : déterminer deux points en correspondance stéréoscopique (appariement) et calculer l'intersection de leur deux rayons optiques associés. Le premier des problèmes va être longuement discuté dans ce chapitre, le second est déjà résolu par l'étape de calibrage du capteur stéréoscopique exposée au premier chapitre.

Nous commençons ce chapitre par le principe de l'appariement de points par stéréo-corrélation selon une approche classique opérant sur une paire d'images rectifiées. Nous étudierons comment améliorer la précision d'appariement en utilisant une approche similaire à celle du chapitre 2, puis nous exposerons une nouvelle méthode de stéréo-corrélation précise basée sur une paire d'images non rectifiées, et donc non dégradées. Les équations de la triangulation seront ensuite détaillées puis nous comparerons les performances de notre méthode par rapport à l'algorithme classique utilisant une paire d'images rectifiées.

### 3.1 Appariement d'images stéréoscopiques

Nous allons aborder le problème de la mise en correspondance de points dans une paire d'images stéréoscopiques. Cette phase d'appariement est préalable à la reconstruction tridimensionnelle par triangulation que l'on exposera dans la section 3.2, page 93. Dans le chapitre 2, nous avons étudié une technique d'appariement temporel de points répartis selon un maillage obtenu par une méthode de triangulation de DELAUNAY contrainte. Nous ne souhaitons pas restreindre le champ d'application de ce chapitre à la seule reconstruction tridimensionnelle

de points épars d'un maillage, c'est la raison pour laquelle nous présentons dans cette section une technique d'appariement dense.

Nous commençons par rappeler la notion de droite épipolaire dans la section 3.1.1, propriété des capteurs de stéréovision. Puis nous présentons une technique classique dite de **rectification**, basée sur cette propriété, permettant de simplifier le problème de l'appariement stéréoscopique en une simple recherche monodimensionnelle — section 3.1.2. Une technique classique d'appariement, reposant sur la mise en correspondance d'une pyramide d'images, est présentée et discutée dans la section 3.1.3. Puis, section 3.1.4, une méthode de corrélation précise est exposée en se basant sur la même approche que celle utilisée pour le suivi précis de pixel du chapitre 2. Enfin, nous détaillons dans la section 3.1.5 la méthode originale que nous avons développée dans le but d'atteindre les meilleures performances en précision possible.

### 3.1.1 Géométrie épipolaire

Nous complétons, dans cette section, l'étude de la géométrie d'un capteur de stéréovision présentée section 1.3.3, page 22, par les notions d'**épipole** et de **droite épipolaire**. Ces résultats sont nécessaires pour aborder la section suivante.

Pour toute paire de points homologues (corrégés de leur distorsion)  $\mathbf{m}$  et  $\mathbf{m}'$ , il existe une contrainte dite **contrainte épipolaire** qui est définie par l'équation implicite :

$$\tilde{\mathbf{m}}'^{\top} \mathbf{F} \tilde{\mathbf{m}} = 0 \quad (3.1)$$

où  $\mathbf{F}$  est appelée **matrice fondamentale**. C'est une matrice  $3 \times 3$  de rang 2 qui fait correspondre à tout point de  $\mathcal{I}$  une droite dans  $\mathcal{I}'$ , et à tout point de  $\mathcal{I}'$  une droite dans  $\mathcal{I}$ . Donc si  $\mathbf{m}$  est un point de  $\mathcal{I}$ , alors  $\mathbf{F}\tilde{\mathbf{m}} = l'$  est une droite épipolaire dans  $\mathcal{I}'$  puisque d'après l'équation (3.1) nous avons la relation  $\tilde{\mathbf{m}}'^{\top} l' = 0$ . Par conséquent un point  $\mathbf{m}'$  qui correspond à  $\mathbf{m}$  doit appartenir à sa droite épipolaire  $\mathbf{F}\tilde{\mathbf{m}}$ . Ceci est une contrainte forte qui est généralement exploitée pour simplifier la stéréoscopie car elle conduit à une recherche monodimensionnelle d'un stéréo-correspondant le long de sa droite épipolaire plutôt qu'une recherche bidimensionnelle dans une zone de l'image (comme c'est le cas pour la méthode de suivi de pixel présentée au chapitre 2). Cela a pour effet d'accélérer la mise en correspondance et de rendre la méthode plus robuste (moins de faux appariements potentiels).

Pour une matrice fondamentale  $\mathbf{F}$  donnée, il existe une paire de points uniques  $\mathbf{e} \in \mathcal{I}$  et  $\mathbf{e}' \in \mathcal{I}'$  tels que :

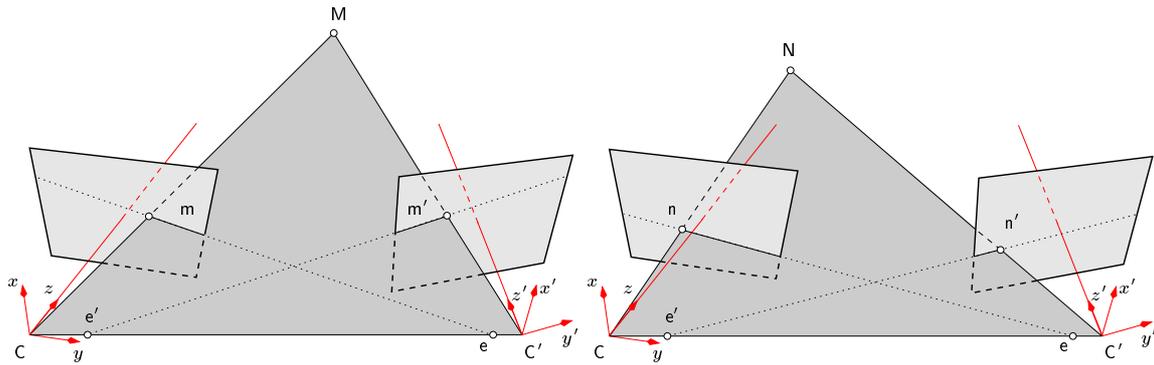
$$\mathbf{F}\tilde{\mathbf{e}} = \mathbf{0} = \mathbf{F}^{\top}\tilde{\mathbf{e}}' \quad (3.2)$$

où  $\mathbf{0} = (0 \ 0 \ 0)^{\top}$  est un vecteur nul. Ces points  $\mathbf{e}$  et  $\mathbf{e}'$  sont appelés respectivement épipole de l'image  $\mathcal{I}$  et épipole de l'image  $\mathcal{I}'$ . Les épipoles ont la propriété que toutes les droites épipolaires de  $\mathcal{I}$  passent par  $\mathbf{e}$  et toutes les droites épipolaires de  $\mathcal{I}'$  passent par  $\mathbf{e}'$  (cf. figure 3.1).

### 3.1.2 Rectification

Dans cette section, nous allons considérer chacune des deux caméras exempte de distorsion : il est en effet toujours possible de se ramener à une telle hypothèse en corrigeant les coordonnées par une méthode telle que celle décrite section 1.2.3, page 18.

Lorsque la transformation rigide  $\mathbf{T}_s$  (cf. figure 1.11, page 22) qui relie les deux repères des caméras est réduite à une seule translation le long de l'axe  $y$ , *i.e.* la droite  $\mathbf{CC}'$  est confondue avec les axes  $y$  et  $y'$ , alors les droites épipolaires sont horizontales et parallèles (cf. figure 3.2).

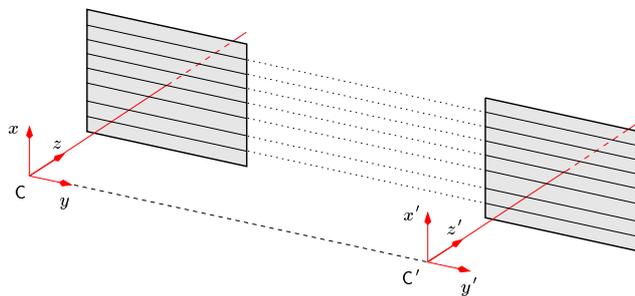


**FIG. 3.1 :** Pour une configuration géométrique donnée, un point de la scène et les deux centres optiques  $C$  et  $C'$  forment un plan dont les intersections avec chacun des deux plans images sont des droites épipolaires. L'ensemble des droites épipolaires d'une caméra forme un faisceau de droites dont le sommet, appartenant à la droite  $CC'$ , est appelé épipole.

Cela signifie que les deux épipoles  $e$  et  $e'$ , sommets des deux faisceaux de droites épipolaires, sont à l'infini. Cette configuration géométrique parfaite est particulièrement intéressante pour les algorithmes de stéréoscopie exploitant la contrainte épipolaire, car la recherche monodimensionnelle des points homologues est effectuée simplement sur une ligne horizontale de pixels d'ordonnées entières (*i.e.* la coordonnée  $u$  est constante et confondue avec une ligne de l'image). On dit alors que la paire d'images stéréoscopiques est **rectifiée**.

La différence entre les coordonnées du point homologue et celles dans l'image de référence est appelée **disparité**. Généralement c'est un vecteur à deux dimensions, mais dans le cas d'une paire d'images rectifiées, la disparité est définie par un seul scalaire : la différence des abscisses des deux points en correspondance stéréoscopique. Lorsqu'une paire d'images stéréoscopiques est rectifiée, la matrice fondamentale  $\mathbf{F}$  et la matrice essentielle  $\mathbf{E}$  prennent pour expression :

$$\mathbf{F} \cong \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \text{et} \quad \mathbf{E} \cong \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$



**FIG. 3.2 :** Capteur de stéréovision rectifié : les droites épipolaires sont horizontales et parallèles. Dans cette configuration géométrique les deux épipoles sont à l'infini.

L'équation (1.30), page 25, nous donne la relation entre la matrice fondamentale, la matrice

essentielle et les deux matrices des paramètres intrinsèques :

$$\begin{aligned} \mathbf{F} &\cong \mathbf{A}'^{-\top} \mathbf{E} \mathbf{A}^{-1} \\ &\cong \begin{pmatrix} 0 & 0 & \frac{1}{\alpha'_u} \\ 0 & 0 & 0 \\ -\frac{1}{\alpha_u} & 0 & \frac{u_0}{\alpha_u} - \frac{u'_0}{\alpha'_u} \end{pmatrix} \end{aligned}$$

Nous pouvons observer que la configuration géométrique rectifiée nécessite, en plus du positionnement en translation pure de la deuxième caméra par rapport à la première, que les deux capteurs aient leur point principal à la même ordonnée, et que les coefficients  $\alpha_u$  et  $\alpha'_u$  soient identiques. Expérimentalement, il est donc très difficile de garantir parfaitement toutes ces conditions.

Il est toutefois possible de se ramener à un système équivalent en procédant à une rectification numérique du capteur [Aya89, Fau93]. Il s'agit d'appliquer une transformation homographique  $\mathbf{H}$ , respectivement  $\mathbf{H}'$ , sur l'image de la caméra gauche, respectivement droite, de sorte que les épipoles soient rejetés à l'infini, *i.e.* en un point  $\tilde{\mathbf{e}}_\infty \cong (0 \ 1 \ 0)^\top$ . Notons  $\bar{\mathbf{F}}$  la matrice fondamentale rectifiée :

$$\bar{\mathbf{F}} \cong (\tilde{\mathbf{e}}_\infty)_\times \cong \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

où  $(\tilde{\mathbf{e}}_\infty)_\times$  est la matrice antisymétrique formée des éléments du point  $\mathbf{e}_\infty$ . Soient  $\mathbf{m} \in \mathcal{I}$  et  $\mathbf{m}' \in \mathcal{I}'$  une paire de points homologues,  $\tilde{\mathbf{m}}$  et  $\tilde{\mathbf{m}}'$  ces points rectifiés :

$$\tilde{\mathbf{m}} \cong \mathbf{H} \mathbf{m} \tag{3.3}$$

$$\tilde{\mathbf{m}}' \cong \mathbf{H}' \mathbf{m}' \tag{3.4}$$

D'après ces deux équations et la contrainte épipolaire (3.1), il vient :

$$\begin{aligned} \tilde{\mathbf{m}}'^\top \bar{\mathbf{F}} \tilde{\mathbf{m}} &= 0 \\ &\Downarrow \\ \tilde{\mathbf{m}}'^\top \underbrace{\mathbf{H}'^\top \bar{\mathbf{F}} \mathbf{H}}_{\mathbf{F}} \tilde{\mathbf{m}} &= 0 \end{aligned}$$

Ce qui conduit à la factorisation :

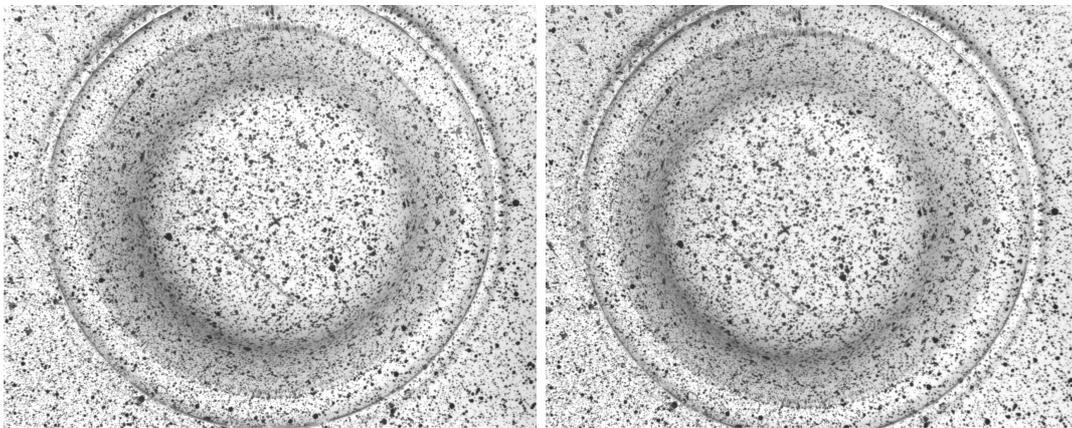
$$\mathbf{F} \cong \mathbf{H}'^\top (\tilde{\mathbf{e}}_\infty)_\times \mathbf{H}$$

L'ensemble des transformations homographiques  $\mathbf{H}$  et  $\mathbf{H}'$  satisfaisant cette relation forme une famille à neuf paramètres [Dev97]. Il est alors intéressant de déterminer cette paire de matrices de rectification de telle sorte qu'elle minimise la distorsion induite dans les deux images [LG95, Zel96, HS97, Dev97].

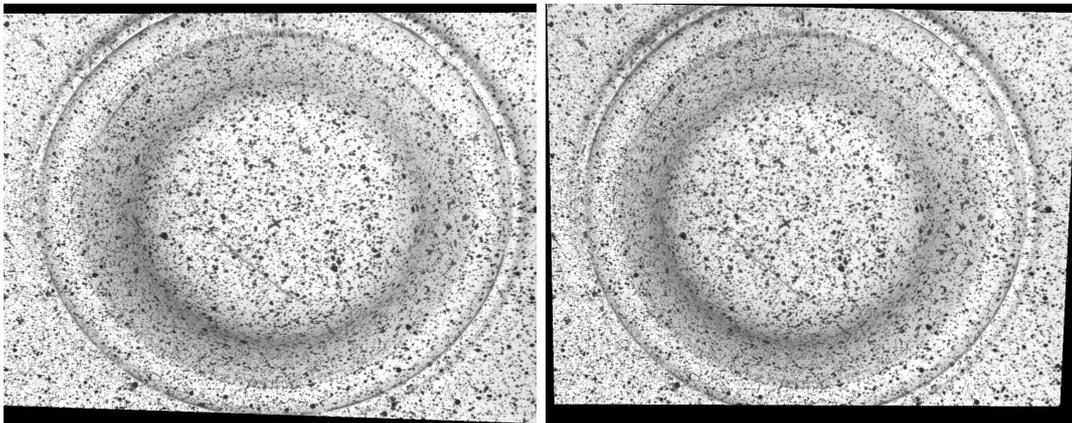
Nous avons utilisé la méthode de LOOP et ZHANG décrite dans [LZ99]. Ils proposent une approche originale basée sur une décomposition des deux homographies  $\mathbf{H}$  et  $\mathbf{H}'$  en transformations élémentaires :

$$\mathbf{H} = \mathbf{H}_a \mathbf{H}_p \quad \text{avec} \quad \mathbf{H}_a = \mathbf{H}_s \mathbf{H}_r$$

où  $\mathbf{H}_p$  est une transformation projective et  $\mathbf{H}_a$  une transformation affine. Cette dernière est le résultat de la composition de deux transformations :  $\mathbf{H}_r$  est une similitude et  $\mathbf{H}_s$  une transformation d'étirement. Un grand intérêt de cette décomposition réside dans l'interprétation géométrique immédiate de chacun des paramètres de la rectification. LOOP et ZHANG notent que les distorsions induites par la rectification ne peuvent être dues qu'aux transformations projectives  $\mathbf{H}_p$  et  $\mathbf{H}'_p$  et aux transformations d'étirement  $\mathbf{H}_s$  et  $\mathbf{H}'_s$ . Pour minimiser ces distorsions, ils proposent de déterminer les matrices  $\mathbf{H}_p$  et  $\mathbf{H}'_p$  les « plus affines » possibles, et tentent de préserver l'orthogonalité et le rapport d'aspect des deux segments horizontaux et verticaux médians dans les images pour le calcul de  $\mathbf{H}_s$  et  $\mathbf{H}'_s$  (voir un exemple de résultat de cet algorithme figure 3.3).



paire d'images stéréoscopiques originale



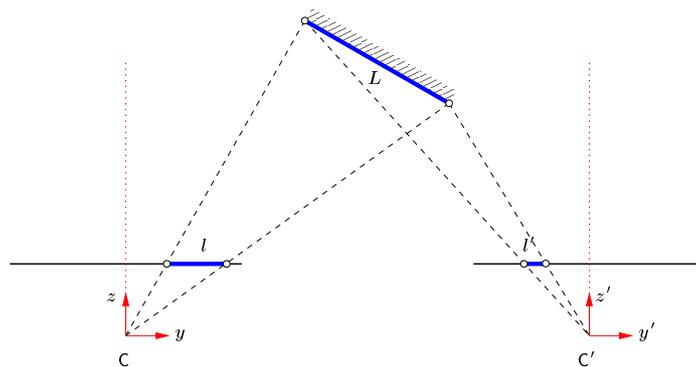
paire d'images stéréoscopiques rectifiées selon la méthode de LOOP et ZHANG

**FIG. 3.3 :** Exemple d'une paire d'images stéréoscopiques rectifiées d'une tôle mince emboutie : les bords d'images sont courbes car notre implémentation de la rectification corrige aussi la distorsion (cf. chapitre 1).

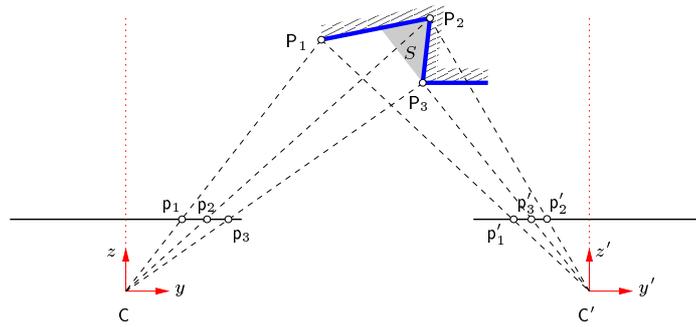
### 3.1.3 Technique d'appariement classique

La mise en correspondance de points dans une paire d'images stéréoscopiques est un problème majeur en vision par ordinateur ayant conduit à une abondante littérature ces trois dernières décennies (voir [BF82, DA89, Zha93a] pour un historique de l'évolution des différentes approches proposées). Une taxinomie couramment employée des différents algorithmes d'appariement consiste à distinguer les méthodes appariant des primitives des méthodes surfaciques. Pour les mêmes raisons énoncées au chapitre 2, section 2.1, page 43, nous choisissons de baser nos travaux sur les méthodes surfaciques plutôt que celles appariant des primitives. En effet, nous souhaitons mesurer la forme tridimensionnelle d'un objet en tout ou partie de sa surface observée : il est donc nécessaire d'utiliser une méthode de stéréoscopie pouvant conduire à l'obtention d'une carte dense de disparités, ce qui porte naturellement notre choix vers les méthodes surfaciques. Nous allons rappeler brièvement le principe de l'appariement par corrélation en l'appliquant au cas d'une paire d'images stéréoscopique.

Les méthodes surfaciques considèrent que deux images sont localement semblables autour de deux points en correspondance stéréoscopique. Une mesure de similarité (ou critère de corrélation) est alors utilisée pour rechercher dans une image le point homologue dont le voisinage est le plus similaire à celui d'un point de l'autre image, considérée comme celle de référence. Cette approche conduit naturellement à l'obtention d'un champ dense de disparité car cette recherche peut être effectuée en chaque point de l'image de référence. La condition préalable au bon fonctionnement des méthodes surfaciques (similarité des fonctions d'intensités au voisinage de deux points en correspondance) se traduit par les trois contraintes classiques, sur la surface d'un objet observé, déjà présentées page 2.1.3. La contrainte de limite de la déformation locale apparente est appelée, en stéréovision, **contrainte fronto-parallèle** et traduit la nécessité que la surface observée soit parallèle aux plans rétinien des deux caméras pour éviter qu'elle soit géométriquement déformée par les deux transformations projectives (*cf.* figure 3.4). Nous étudierons en section 3.1.4 une méthode de corrélation permettant d'assouplir cette contrainte. La contrainte de continuité doit aussi être vérifiée dans le cadre stéréoscopique. La figure 3.5 montre un cas extrême du non respect de cette contrainte : l'occlusion.



**FIG. 3.4 :** *Lorsqu'un élément de la surface observée n'est pas fronto-parallèle aux plans images, il lui correspond deux motifs de tailles différentes dans les images (dans cet exemple, le segment de longueur  $L$  se projette en deux segments de longueur  $l \neq l'$ ).*



**FIG. 3.5 :** Le problème d'occlusion : en ne considérant que le cas d'objets opaques, la zone grisée étiquetée  $S$  visible par la caméra gauche ne l'est pas par celle de droite. Le point  $p_2$ , par exemple, ne peut donc pas être apparié dans l'image de droite.

### Stéréo-corrélation classique

La détermination de l'homologue stéréoscopique d'un point donné par corrélation classique est similaire à la technique présentée section 2.2, page 48, pour laquelle la transformation locale est approximée à l'ordre 0. De plus la géométrie épipolaire indique le lieu de projection possible d'un point d'une image dans l'autre image. Lorsque les images sont rectifiées, nous savons, d'après la section précédente, que l'homologue d'un point donné est nécessairement sur une droite horizontale, de même ordonnée que le point à appairer. La transformation locale du domaine de corrélation de l'image de base dans l'autre image est donc réduite à une seule translation (la disparité) le long de l'axe horizontal de l'image, et non plus un vecteur à deux dimensions<sup>1</sup>. Cette contrainte géométrique forte permet de réduire la complexité des algorithmes d'appariement à une recherche monodimensionnelle sur cette droite horizontale, au lieu d'une recherche bidimensionnelle.

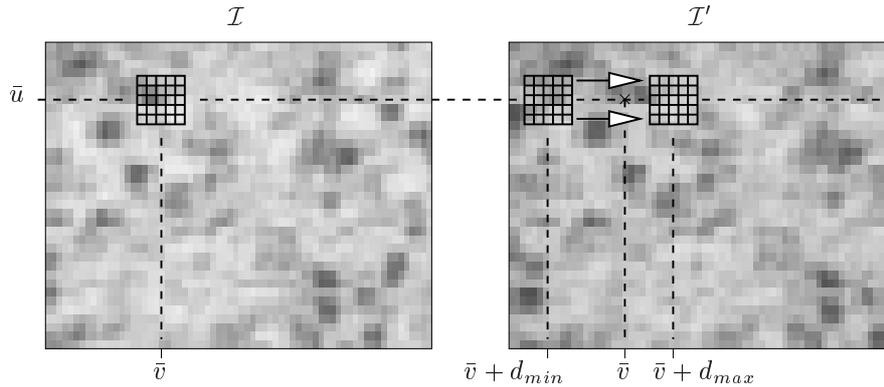
La recherche monodimensionnelle d'un stéréo-correspondant peut être effectuée sur un intervalle restreint de points plutôt que sur la ligne entière. Ceci permet de réduire le taux de faux appariements (*i.e.* échecs dans la mise en correspondance) et d'accélérer simultanément la recherche de l'homologue. Cette réduction de l'espace de recherche par un intervalle de disparité, *i.e.* un segment de la droite épipolaire où chercher le point homologue, peut être déterminé de plusieurs manières. La première consiste à déterminer les disparités minimale et maximale de la projection de l'objet à partir de la connaissance de sa distance au capteur [Dev97]. Une deuxième méthode consiste à déterminer cet intervalle en mesurant la disparité sur certains points des images.

L'algorithme classique de stéréoscopie par corrélation sur images rectifiées se décline selon les étapes suivantes :

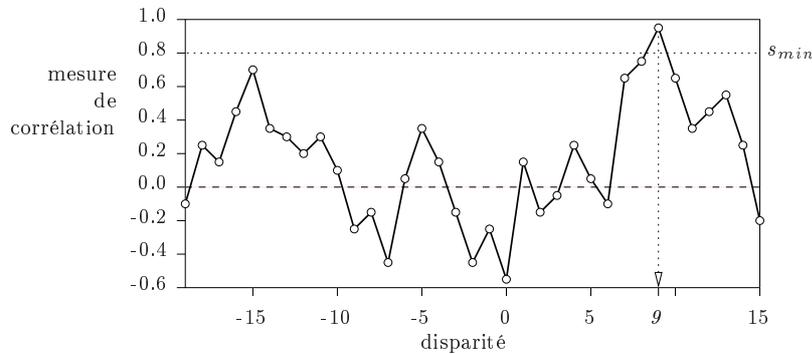
- pour chaque point à appairer de l'image de référence, un intervalle de recherche dans l'autre image est déterminé : l'intervalle de disparités  $[d_{min} \dots d_{max}]$  ;
- le degré de corrélation entre le motif de l'image de référence et les motifs centrés sur chaque point de cet intervalle dans l'autre image est calculé (*cf.* figure 3.6) ;
- l'homologue est alors déterminé comme celui conduisant à la plus forte valeur de corrélation (*cf.* figure 3.7).

Afin d'augmenter la robustesse des algorithmes de stéréoscopie par corrélation, il est né-

<sup>1</sup>C'est le cas du problème de suivi temporel de pixels dans une séquence d'images présenté au chapitre 2.



**FIG. 3.6 :** Appariement par corrélation d'un point dans une paire d'images stéréoscopiques rectifiées : des mesures de corrélation sont effectuées entre un domaine de l'image  $\mathcal{I}$  et chacun des domaines de l'image  $\mathcal{I}'$  défini sur un intervalle de disparités restreint.



**FIG. 3.7 :** L'homologue stéréoscopique d'un point est déterminé comme celui conduisant au meilleur score de corrélation sur un intervalle de disparité donné. Afin de rendre plus robuste cette recherche, un seuil de corrélation minimal  $s_{min}$  peut être imposé pour considérer l'appariement valide. Pour cet exemple, l'homologue est déterminé avec une disparité égale à 9 (valeur entière ; voir plus loin une discussion sur l'obtention d'une valeur subpixel).

cessaire de mettre en œuvre certaines contraintes pour la mise en correspondance [HM95] :

**contrainte épipolaire** : la recherche de l'homologue d'un point doit être effectuée le long de sa droite épipolaire associée. Cette contrainte est implicitement prise en compte lorsque la recherche est effectuée sur une seule ligne d'un couple d'images rectifiées ;

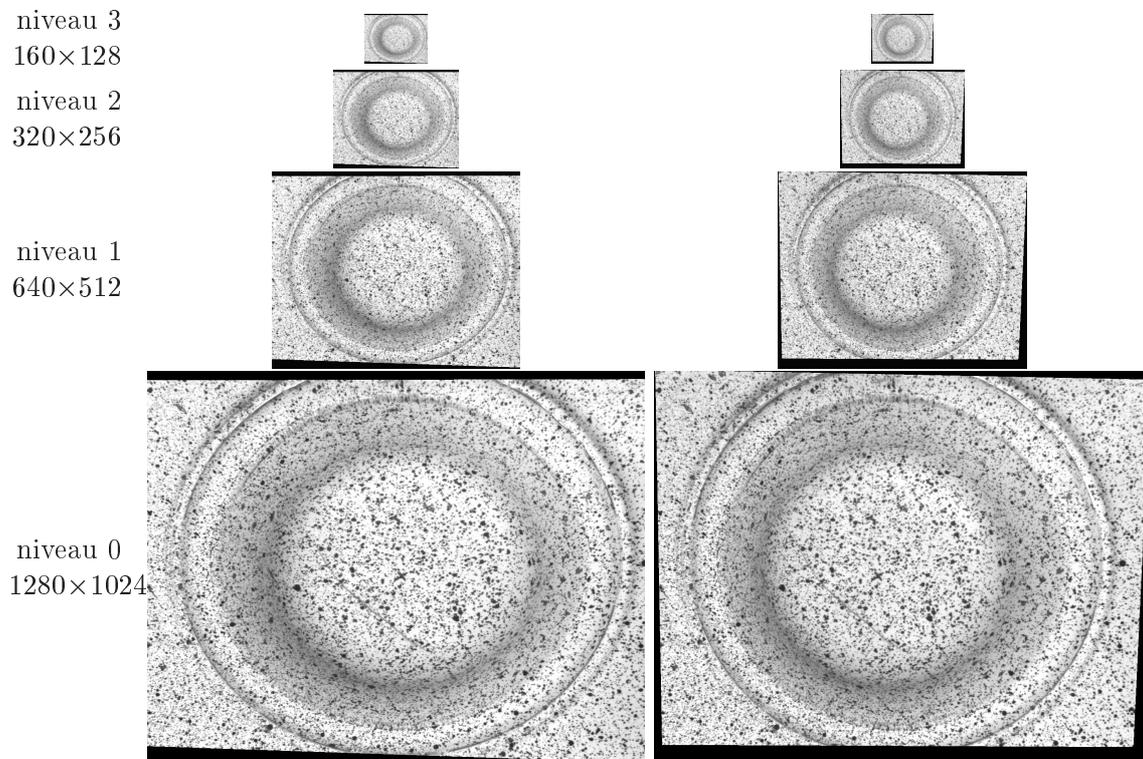
**contrainte d'unicité** : un point d'une image ne peut avoir au plus qu'un seul homologue dans l'autre image ;

**contrainte d'ordre** : lorsqu'un objet observé est opaque, l'ordre des points de l'image de référence doit être identique à celui des homologues appariés dans l'autre image ;

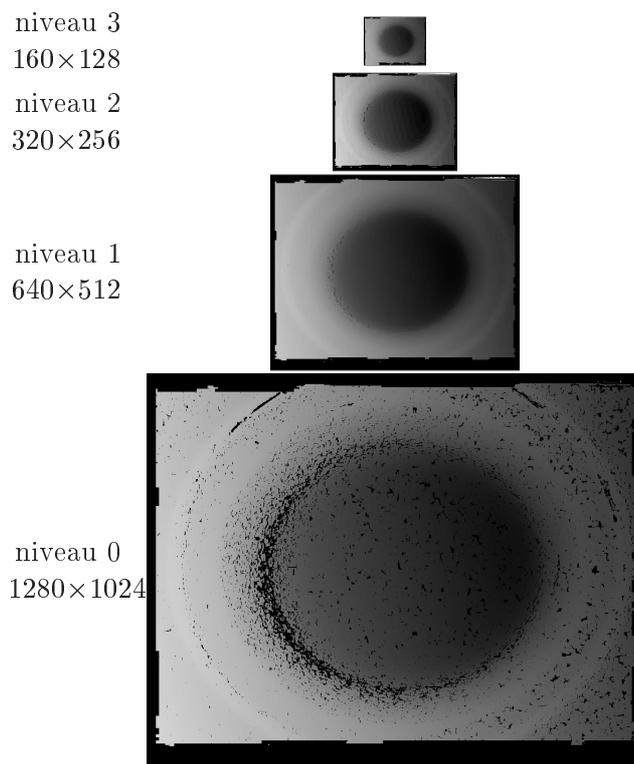
**validation croisée** : lorsqu'un point d'une image est apparié avec un point de l'autre image, alors ce dernier doit être apparié avec le même point initial [Fua91]. Cette vérification nécessite une double évaluation puisqu'elle fait jouer un rôle symétrique aux deux images. Toutefois elle s'avère indispensable en pratique car elle permet l'élimination de nombreux faux appariements.

### Approche pyramidale

La stéréoscopie par corrélation est généralement un processus nécessitant un temps de calcul important. Afin de réduire ce délai nécessaire à l'obtention de la carte de disparité, il est possible d'utiliser une approche multi-résolution. Elle consiste à effectuer la stéréoscopie par corrélation sur chaque niveau d'une paire de pyramides d'images, en utilisant l'information calculée au niveau inférieur pour limiter la zone de recherche des homologues à chaque niveau. La pyramide d'une image est construite par diminution progressive de la résolution par filtrage et sous-échantillonnage [Zha93b]. La figure 3.8 montre les deux pyramides d'images obtenues pour les niveaux zéro à trois (le niveau zéro correspondant aux images initiales) de la paire d'images stéréoscopiques rectifiée de la figure 3.3, en utilisant un filtre gaussien. La figure 3.9 donne le résultat de la stéréoscopie par corrélation sur chaque niveau de cette pyramide d'images. Outre le gain de temps procuré par cette approche, l'utilisation d'une pyramide d'images conduit généralement à une estimation encore plus robuste des disparités [Zha93b].



**FIG. 3.8 :** Paire de pyramides d'images obtenues par filtrage gaussien et sous-échantillonnage de la paire d'images initiale de dimensions  $1280 \times 1024$  du niveau 0.



**FIG. 3.9 :** Résultat de la stéréoscopie par corrélation classique sur la pyramide d'images de la paire stéréoscopique rectifiée de la figure 3.8. Le critère ZNCC a été utilisé sur des domaines de corrélation carrés de sept pixels de côté.

### Limitation des performances de l'approche classique

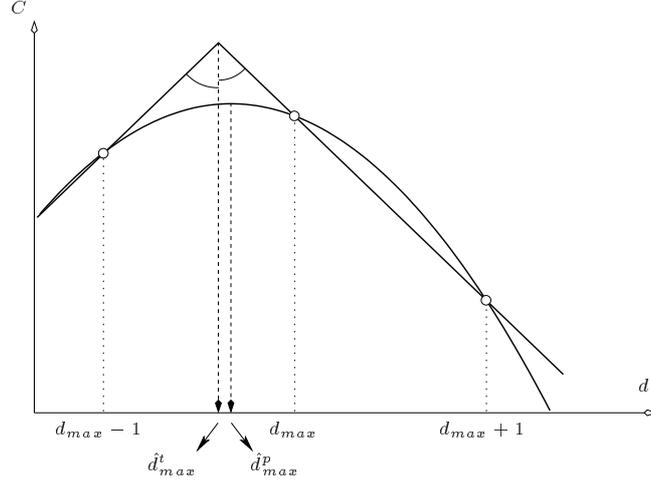
La stéréoscopie par corrélation classique est basée sur une mesure de ressemblance entre deux domaines de pixels de même formes et dimensions. Cette approche a un sens seulement si la contrainte fronto-parallèle est respectée. Dans le cas contraire, il apparaît une distorsion des domaines de pixels entre les deux images, conduisant à une dégradation de la mesure de ressemblance. Une solution consiste alors à opérer une deuxième passe pour affiner les disparités de l'approche classique, en estimant simultanément la géométrie de chaque domaine de corrélation ainsi que la disparité au centre du domaine. Cette amélioration fait l'objet de la section suivante.

#### 3.1.4 Appariement par corrélation précise

Lorsqu'on recherche l'homologue stéréoscopique d'un point par corrélation dans une paire d'images rectifiées, nous venons de voir que l'algorithme classique consiste à déterminer la disparité  $d$  qui maximise la ressemblance entre deux domaines rectangulaires de même dimension : l'un centré autour du point à appairer dans une image, et l'autre dans la deuxième image dont la position est fonction de  $d$ . Cependant, cette simplicité a pour contrepartie de conduire à l'estimation de disparités au pixel près, et de dégrader encore davantage la précision lorsque localement la surface n'est pas fronto-parallèle aux plans image.

Afin d'atteindre une précision sous-pixelique des disparités, de nombreux auteurs ont

proposé d'interpoler la fonction de corrélation autour de l'extremum trouvé, par une fonction (parabole, linéaire par morceaux, ...) dont l'extremum est censé fournir une précision sous-pixellique de la disparité. Cette méthode semble conduire qualitativement à une précision meilleure que le pixel, mais il est difficile de la confondre à une « réalité physique » [Dev97], cf. figure 3.10.



**FIG. 3.10 :** Approximation sous-pixellique d'une disparité par recherche de l'extremum d'une fonction d'interpolation des scores de corrélation au voisinage du score maximal  $d_{max}$ . L'interpolant utilisé peut être une fonction linéaire par morceaux (méthode dite « du toit ») ou une parabole conduisant respectivement à l'approximation sous-pixellique de la disparité  $\hat{d}_{max}^t$  et  $\hat{d}_{max}^p$  (d'après [Dev97]).

Il est aussi possible d'appliquer une méthode similaire à celle présentée au chapitre 2 en considérant la transformation locale stéréoscopique dans les critères de corrélation. Reprenons les raisonnements exposés section 2.2.1, page 48 pour le cas d'une paire d'images rectifiée : posons  $\bar{\mathbf{m}}_c = (\bar{u}_c \ \bar{v}_c)^\top$  un point de l'image de référence, habituellement le centre du domaine de corrélation, et  $\bar{\mathbf{m}} = (\bar{u} \ \bar{v})^\top$  un point de ce domaine. En ce point  $\bar{\mathbf{m}}$ , la transformation locale stéréoscopique, que nous noterons  $\zeta(\bar{\mathbf{m}})$  s'écrit :

$$\zeta(\bar{\mathbf{m}}) = \bar{\mathbf{m}} + \eta(\bar{\mathbf{m}})$$

où  $\eta(\bar{\mathbf{m}})$  est le déplacement au point  $\bar{\mathbf{m}}$ . Puisque la paire d'images est rectifiée, deux points en correspondance stéréoscopique ont la même ordonnée  $\bar{u}$ . La fonction déplacement  $\eta(\bar{\mathbf{m}})$  peut donc s'écrire en utilisant la fonction disparité  $d(\bar{\mathbf{m}})$  :

$$\eta(\bar{\mathbf{m}}) = \begin{pmatrix} 0 \\ d(\bar{\mathbf{m}}) \end{pmatrix}$$

En posant  $\delta_u = \bar{u} - \bar{u}_c$  et  $\delta_v = \bar{v} - \bar{v}_c$  pour simplifier l'écriture, le développement limité autour du point  $\bar{\mathbf{m}}_c$  de la fonction disparité  $d(\bar{\mathbf{m}})$  jusqu'à l'ordre deux est :

$$d(\bar{\mathbf{m}}) \approx \underbrace{d}_{\text{ordre 0}} + \underbrace{d_u \delta_u + d_v \delta_v}_{\text{ordre 1}} + \underbrace{d_{u^2} \delta_u^2 + d_{uv} \delta_u \delta_v + d_{v^2} \delta_v^2}_{\text{ordre 2}}$$

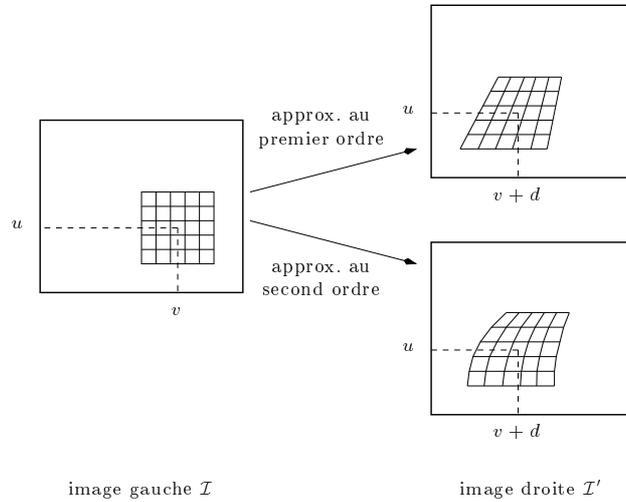
avec :

$$\begin{aligned} d &= d(\bar{\mathbf{m}}_c), \\ d_u &= \frac{\partial d}{\partial \bar{u}}(\bar{\mathbf{m}}_c), \quad d_v = \frac{\partial d}{\partial \bar{v}}(\bar{\mathbf{m}}_c), \\ d_{u2} &= \frac{1}{2} \frac{\partial^2 d}{\partial \bar{u}^2}(\bar{\mathbf{m}}_c), \quad d_{uv} = \frac{1}{2} \frac{\partial^2 d}{\partial \bar{u} \partial \bar{v}}(\bar{\mathbf{m}}_c), \quad d_{v2} = \frac{1}{2} \frac{\partial^2 d}{\partial \bar{v}^2}(\bar{\mathbf{m}}_c) \end{aligned}$$

Nous pouvons donc définir une expression  $\zeta_c(\bar{\mathbf{m}}, \bar{\mathbf{m}}_c, \mathbf{q})$  de l'approximation à l'ordre deux de la transformation locale stéréoscopique  $\zeta(\bar{\mathbf{m}})$  au voisinage du point  $\bar{\mathbf{m}}_c$  :

$$\zeta_c(\bar{\mathbf{m}}, \bar{\mathbf{m}}_c, \mathbf{q}) = \left( \bar{v} + d + d_u \delta_u + d_v \delta_v + d_{u2} \delta_u^2 + d_{uv} \delta_u \delta_v + d_{v2} \delta_v^2 \right) \quad (3.5)$$

où  $\mathbf{q} = (d \ d_u \ d_v \ d_{u2} \ d_{uv} \ d_{v2})^\top$  est le vecteur des paramètres de la transformation stéréoscopique locale. DEVERNAY et FAUGERAS ont proposé dans [DF94] d'estimer les paramètres  $d$ ,  $d_u$ ,  $d_v$ ,  $d_{u2}$ ,  $d_{uv}$  et  $d_{v2}$  (ou un sous-ensemble de ceux-ci suivant l'ordre choisi de l'approximation) par un processus d'optimisation tel que ceux présentés section 2.2.3, page 51, en utilisant la transformation stéréoscopique  $\zeta$  en lieu et place de  $\Phi$ . Ils ont alors nommé cette technique **stéréo-corrélation fine** et l'ont utilisée non seulement pour améliorer la précision des appariements obtenus, mais aussi pour déterminer des propriétés différentielles des surfaces. Géométriquement, cette approche consiste à adapter localement la géométrie des domaines de corrélation à la courbure et à l'orientation locale de la surface observée, *cf.* figure 3.11.



**FIG. 3.11** : Approximation au premier et au second ordre de la déformation d'un domaine rectangulaire de l'image gauche dans l'image droite, pour une paire d'images rectifiée (d'après [Dev97]).

Lorsque la géométrie épipolaire estimée via le processus de calibrage du capteur stéréoscopique est peu précise, il est possible d'autoriser la recherche des appariements hors de leur droite épipolaire respective. La transformation stéréoscopique locale peut alors être approximée au premier ordre par une transformation affine à 6 paramètres [Bra95, Lan97]. Nous

n'avons pas retenu cette solution du fait de la grande précision que nous avons toujours obtenue pour la géométrie épipolaire estimée via la méthode **GLOB** de calibrage du capteur de stéréovision présentée au chapitre 1.

### 3.1.5 Méthode développée : amélioration de la stéréo-corrélation précise

DEVERNAY souligne dans sa thèse [Dev97] qu'il doit être possible d'obtenir une meilleure précision de la stéréo-corrélation fine en opérant directement sur les images stéréoscopiques initiales, *i.e.* non rectifiées. Cette remarque est motivée par le fait que le processus de rectification génère deux images dont les niveaux de gris ont été calculés par interpolation de ceux des images initiales, conduisant alors à une dégradation du contenu des images rectifiées. La méthode que nous avons développée opère donc sur la paire d'images stéréoscopiques initiales et représente *in fine* une amélioration de la stéréo-corrélation fine de DEVERNAY comme nous pourrions l'observer dans la section 3.3.

Lorsque les images sont rectifiées, les critères de corrélation de la stéréoscopie calculent la similarité entre un domaine de pixels voisin d'un point  $\bar{\mathbf{m}}_c$  dans l'image de référence, et un autre domaine de pixels dans l'autre image, tel qu'à un point  $\bar{\mathbf{m}}$  voisin de  $\bar{\mathbf{m}}_c$  est associé le point  $\bar{\mathbf{m}}' = \zeta_c(\bar{\mathbf{m}}, \bar{\mathbf{m}}_c, \mathbf{p})$  dans l'autre image, *cf.* section précédente. Comment peut-on relier cette relation du cas d'une paire d'images stéréoscopiques rectifiée à celle d'une paire initiale ?

Soient  $\mathbf{m}_c$  et  $\mathbf{m}$  deux points de l'image initiale de référence, *i.e.* non rectifiée et distordue (au sens de présentant de la distorsion liée à l'optique du capteur). En ces points correspondent respectivement les points corrigés et rectifiés  $\bar{\mathbf{m}}_c$  et  $\bar{\mathbf{m}}$ , calculables par correction de la distorsion (*cf.* section 1.2.3, page 18) puis rectification (*cf.* section 3.1.2). Nous appellerons cette transformation  $\Gamma^{-1}$  (*cf.* figure 3.12) :

$$\bar{\mathbf{m}}_c = \Gamma^{-1}(\mathbf{m}_c) \quad \text{et} \quad \bar{\mathbf{m}} = \Gamma^{-1}(\mathbf{m})$$

Remarquons que l'application  $\Gamma^{-1}$  établit une correspondance entre les coordonnées initiales (distordues et non rectifiées) et les coordonnées corrigées-rectifiées, et peut être efficacement implantée par un tableau précalculé de ces correspondances.

Les coordonnées  $\bar{\mathbf{m}}_c$  et  $\bar{\mathbf{m}}$  étant corrigées et rectifiées, la transformation stéréoscopique locale  $\zeta_c$  de la section précédente est applicable et permet d'exprimer les coordonnées de l'homologue  $\bar{\mathbf{m}}'$  de  $\bar{\mathbf{m}}$  (*cf.* figure 3.12) :

$$\begin{aligned} \bar{\mathbf{m}}' &= \zeta_c(\bar{\mathbf{m}}, \bar{\mathbf{m}}_c, \mathbf{q}) \\ &= \zeta_c(\Gamma^{-1}(\mathbf{m}), \Gamma^{-1}(\mathbf{m}_c), \mathbf{q}) \end{aligned}$$

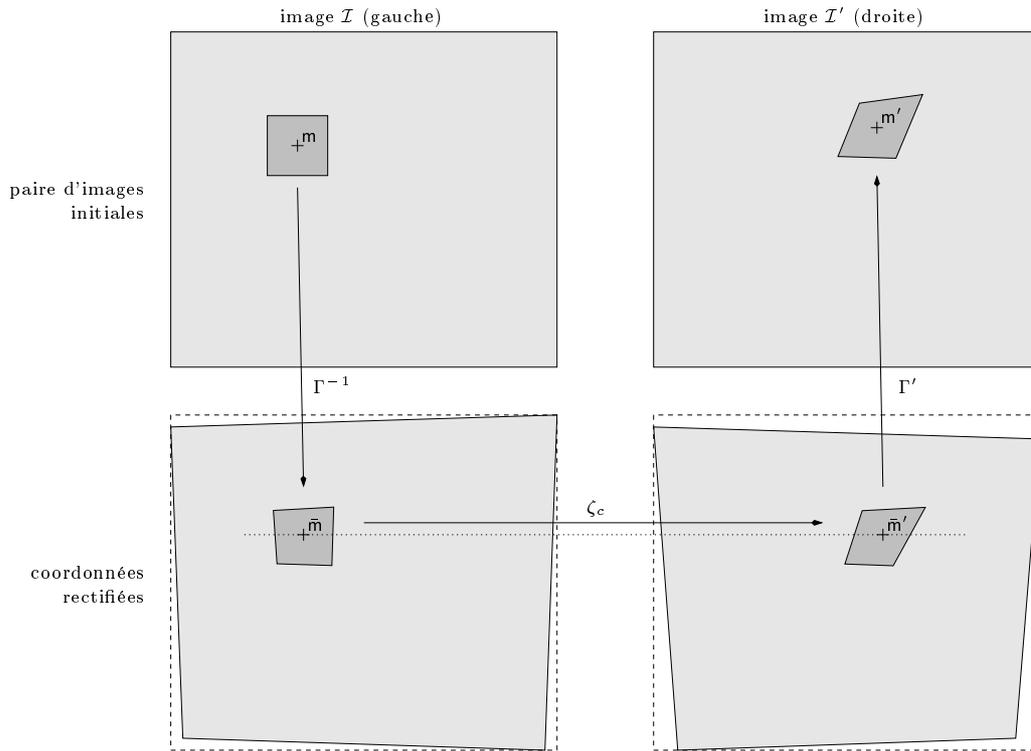
Les coordonnées  $\bar{\mathbf{m}}'$  sont corrigées et rectifiées et l'on peut déterminer son équivalent  $\mathbf{m}'$  dans l'image initiale de référence et son homologue  $\mathbf{m}'$  dans l'autre image initiale, en fonction des mêmes paramètres de la transformation stéréoscopique locale autour de  $\bar{\mathbf{m}}_c$  de la méthode de stéréo-corrélation fine :

$$\mathbf{m}' = \Gamma'(\bar{\mathbf{m}}')$$

Nous pouvons à présent établir la relation entre un point  $\mathbf{m}$ , voisin du point  $\mathbf{m}_c$ , de l'image initiale de référence et son homologue  $\mathbf{m}'$  dans l'autre image initiale, en fonction des mêmes paramètres de la transformation stéréoscopique locale autour de  $\bar{\mathbf{m}}_c$  de la méthode de stéréo-corrélation fine :

$$\mathbf{m}' = \Gamma'(\zeta_c(\Gamma^{-1}(\mathbf{m}), \Gamma^{-1}(\mathbf{m}_c), \mathbf{q}))$$

En utilisant cette transformation en lieu et place de la transformation  $\Phi_c$  dans les critères



**FIG. 3.12 :** Principe de la méthode de stéréo-corrélation fine améliorée : les images initiales fournies par le capteur sont directement exploitées. Les paramètres des transformations stéréoscopiques locales sont déterminés dans l'espace des images rectifiées, et une relation stéréoscopique entre coordonnées initiales est établie par composition des diverses transformations ( $\Gamma^{-1}$ ,  $\zeta_c$  et  $\Gamma'$ ).

de corrélation de la section 2.2.2, page 49, les mêmes techniques d'estimation des paramètres de  $\mathbf{q}$  peuvent alors être utilisées. Toutefois, l'optimisation est opérée dans l'espace des images rectifiées, et nécessite l'évaluation du gradient  $\nabla \mathcal{I}(\bar{\mathbf{m}})$ , c'est-à-dire le gradient des niveaux de gris dans l'image corrigée-rectifiée. Or, il serait souhaitable de le déterminer directement à partir du gradient des niveaux de gris dans l'image initiale  $\hat{\mathcal{I}}'$ , toujours dans le but de ne pas effectuer de calcul sur des images dégradées par la correction de la distorsion et la rectification. L'annexe D donne les détails de ce calcul en fonction de l'homographie de rectification et des paramètres de calibrage de la caméra droite. Notons qu'ici encore, notre implémentation utilise une interpolation par surface B-spline des niveaux de gris dans les images pour évaluer la luminance en des coordonnées non entières, et pour calculer leur gradient.

Enfin, l'estimée initiale de la carte des disparités pour les points initiaux de l'image de référence est obtenue en deux phases :

- la méthode de stéréo-corrélation classique est utilisée sur la paire d'images rectifiées conduisant à l'obtention d'une carte de disparités grossières ;
- cette carte est ensuite dé-rectifiée et distordue, et est utilisée comme estimée initiale pour les processus d'estimation des paramètres des transformations stéréoscopiques locales.

La reconstruction tridimensionnelle des points appariés selon la technique de stéréo-corrélation classique ou bien par sa version améliorée est présentée dans la section suivante.

## 3.2 Triangulation

Lorsque la correspondance entre deux points image d'une paire stéréoscopique est établie, le calcul du point tridimensionnel correspondant consiste à déterminer l'intersection dans l'espace des rayons optiques associés à ces deux points image. Cette méthode est appelée **triangulation**.

Quand le processus d'appariement stéréoscopique n'intègre pas explicitement la contrainte épipolaire, ou bien lorsque celle-ci est « lâche » (par exemple lorsque l'estimation de la géométrie épipolaire est peu précise, certains algorithmes d'appariement appliquent leur recherche dans une bande épipolaire), les rayons optiques ne s'intersectent pas dans l'espace. Le point tridimensionnel reconstruit est alors généralement choisi comme le point médian du segment le plus court reliant les deux rayons optiques, ou par minimisation d'un certain critère comme la méthode décrite section 1.3.3, page 25 (voir [HS97] pour un état de l'art sur la triangulation).

En revanche, dès lors que la stéréoscopie est fortement contrainte par la géométrie épipolaire, les deux rayons optiques concourent exactement en un seul et même point de l'espace. C'est le cas notamment de tous les algorithmes d'appariement respectant rigoureusement la contrainte épipolaire. Nous nous situons dans ce cas de figure.

Pour reconstruire un point de l'espace, il est nécessaire de doter ce dernier d'un référentiel. Un choix couramment effectué consiste à le confondre avec le référentiel de la caméra de gauche; la transformation  $\mathbf{T}$  reliant le repère du monde à celui de la caméra gauche est réduite à une simple identité. D'après les équations de projection (1.21) et (1.22), page 22, et celles de rectification (3.3) et (3.4), les équations de projection pour un capteur rectifié sont :

$$\tilde{\mathbf{m}} \cong \mathbf{H}\mathbf{K}\tilde{\mathbf{M}} \quad (3.6)$$

$$\tilde{\mathbf{m}}' \cong \mathbf{H}'\mathbf{K}'\mathbf{T}_s\tilde{\mathbf{M}} \quad (3.7)$$

Posons  $\mathbf{P} = \mathbf{H}\mathbf{K}$  et  $\mathbf{P}' = \mathbf{H}'\mathbf{K}'\mathbf{T}_s$ , et  $d(\bar{u}, \bar{v})$  la disparité au point rectifié  $\tilde{\mathbf{m}} = (\bar{u} \ \bar{v})^\top$ . La rectification impose que deux points homologues aient la même ordonnée  $\bar{u}$ , les matrices  $\mathbf{P}$  et  $\mathbf{P}'$  ont donc leur première et dernière lignes égales à un facteur multiplicatif  $\lambda$  non nul près :  $\mathbf{p}_1^\top = \lambda\mathbf{p}_1'^\top$  et  $\mathbf{p}_3^\top = \lambda\mathbf{p}_3'^\top$ . Les matrices  $\mathbf{P}$  et  $\mathbf{P}'$  peuvent donc s'écrire sous la forme :

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \mathbf{p}_3^\top \end{pmatrix} \quad \text{et} \quad \mathbf{P}' = \begin{pmatrix} \mathbf{p}_1^\top \\ \lambda\mathbf{p}_2'^\top \\ \mathbf{p}_3^\top \end{pmatrix} \quad \text{avec} \quad \lambda = \frac{\|\mathbf{p}_1^\top\|}{\|\mathbf{p}_1'^\top\|} = \frac{\|\mathbf{p}_3^\top\|}{\|\mathbf{p}_3'^\top\|}$$

Les équations (3.6) et (3.7) deviennent :

$$\tilde{\mathbf{m}} \cong \begin{pmatrix} \bar{u} \\ \bar{v} \\ 1 \end{pmatrix} \cong \begin{pmatrix} \mathbf{p}_1^\top \tilde{\mathbf{M}} \\ \mathbf{p}_2^\top \tilde{\mathbf{M}} \\ \mathbf{p}_3^\top \tilde{\mathbf{M}} \end{pmatrix} \quad \text{et} \quad \tilde{\mathbf{m}} \cong \begin{pmatrix} \bar{u} \\ \bar{v} + d(\bar{u}, \bar{v}) \\ 1 \end{pmatrix} \cong \begin{pmatrix} \mathbf{p}_1^\top \tilde{\mathbf{M}} \\ \lambda\mathbf{p}_2'^\top \tilde{\mathbf{M}} \\ \mathbf{p}_3^\top \tilde{\mathbf{M}} \end{pmatrix}$$

De ces deux équations, il vient l'expression du point de la carte de disparité  $(\bar{u} \ \bar{v} \ d(\bar{u}, \bar{v}) \ 1)^\top$  en fonction des coordonnées homogènes  $(x \ y \ z \ 1)^\top$  du point de la scène, et des matrices de projection rectifiée  $\mathbf{P}$  et  $\mathbf{P}'$  :

$$\underbrace{\begin{pmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \lambda\mathbf{p}_2'^\top - \mathbf{p}_2^\top \\ \mathbf{p}_3^\top \end{pmatrix}}_{\mathbf{Q}} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \cong \begin{pmatrix} \bar{u} \\ \bar{v} \\ d(\bar{u}, \bar{v}) \\ 1 \end{pmatrix}$$

Lorsque les coordonnées appariées sont corrigées de leur distorsion, et rectifiées, la triangulation consiste donc à effectuer une seule opération :

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \cong \mathbf{Q}^{-1} \begin{pmatrix} \bar{u} \\ \bar{v} \\ d(\bar{u}, \bar{v}) \\ 1 \end{pmatrix}$$

où  $\mathbf{Q}^{-1}$  est appelée **matrice de reconstruction** [Dev97]. Si, en revanche, les points appariés le sont dans l'espace des images initiales, *i.e.* distordues et non rectifiées, comme c'est le cas pour notre méthode de stéréo-corrélation fine améliorée, il est nécessaire de se ramener au cas précédent en corrigeant la distorsion et en rectifiant les coordonnées appariées, avant d'appliquer la triangulation telle qu'exposée.

### 3.3 Évaluation des performances

Nous avons évalué qualitativement les performances de notre algorithme de stéréo-corrélation fine améliorée sur deux expériences de reconstruction tridimensionnelle de la surface d'une tôle mince emboutie. Le procédé d'emboutissage est présenté dans le chapitre « applications », section 5.1, page 125.

La section 3.3.1 présente quelques résultats de la reconstruction tridimensionnelle d'une pièce emboutie, en comparant les performances de notre méthode de stéréo-corrélation fine améliorée à la méthode classique de DEVERNAY. La section 3.3.2 montre la reconstruction tridimensionnelle d'une deuxième pièce emboutie ainsi que les résultats de l'estimation de deux des paramètres de la transformation stéréoscopique locale : les dérivées de la disparité selon les lignes et les colonnes de l'image de référence (celle issue de la caméra gauche dans notre cas), exprimées dans l'espace rectifié.

#### 3.3.1 Pièce emboutie numéro 1

La figure 3.13 montre l'embouti étudié qui présente des défauts de formage : la forme obtenue n'est pas hémisphérique malgré l'utilisation d'un poinçon de cette géométrie. Nous avons choisi cet embouti comme premier exemple car il correspond à un cas particulièrement critique pour la stéréo-corrélation : la surface présente de fortes pentes qui conduisent à de fortes distorsions perspectives des domaines de corrélation dans les images.

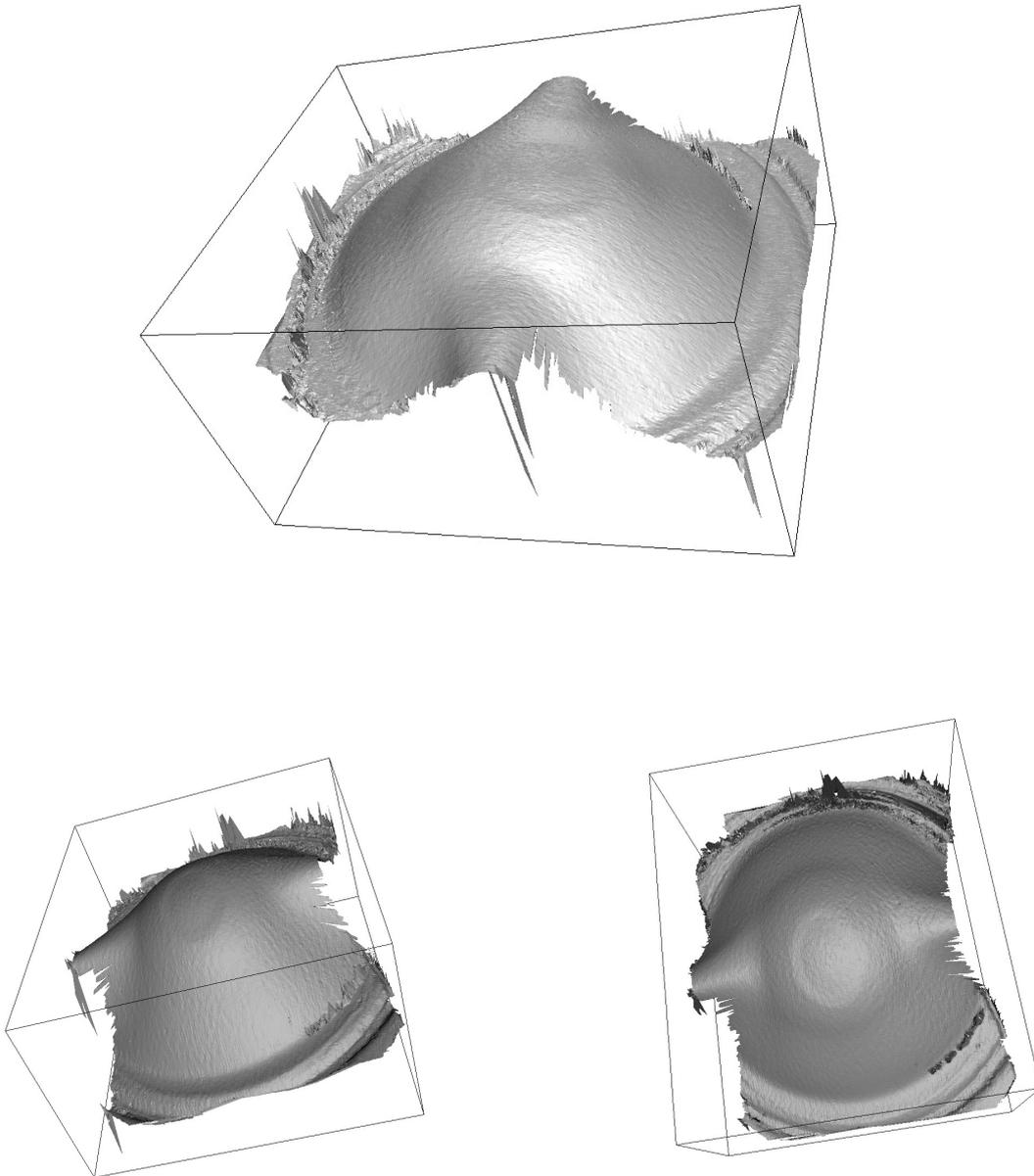
La figure 3.14 montre le résultat de la reconstruction tridimensionnelle à partir de la carte de disparités obtenue par notre méthode de stéréo-corrélation fine améliorée avec un modèle de transformation stéréoscopique locale d'ordre 1 et des domaines de corrélation de  $11 \times 11$  pixels. Cet objet tridimensionnel comporte plusieurs points faux liés aux fausses disparités estimées dans les régions où la peinture a été supprimée par frottement lors de l'emboutissage. Le métal, de l'aluminium, se trouvant découvert dans ces zones, les images acquises comportaient des reflets spéculaires mettant en échec la mise en correspondance. Par ailleurs, la reconstruction tridimensionnelle révèle de nombreux détails dont notamment la trace du contre-flan en acier percé en son centre qui se trouvait sous la tôle d'aluminium pendant l'emboutissage.

Afin d'évaluer qualitativement les performances de notre méthode par rapport à la stéréo-corrélation fine originale, nous avons choisi d'observer un détail de l'embouti — *cf.* figure 3.15. Cette zone a été choisie sur une partie de l'objet où l'orientation locale de sa surface est

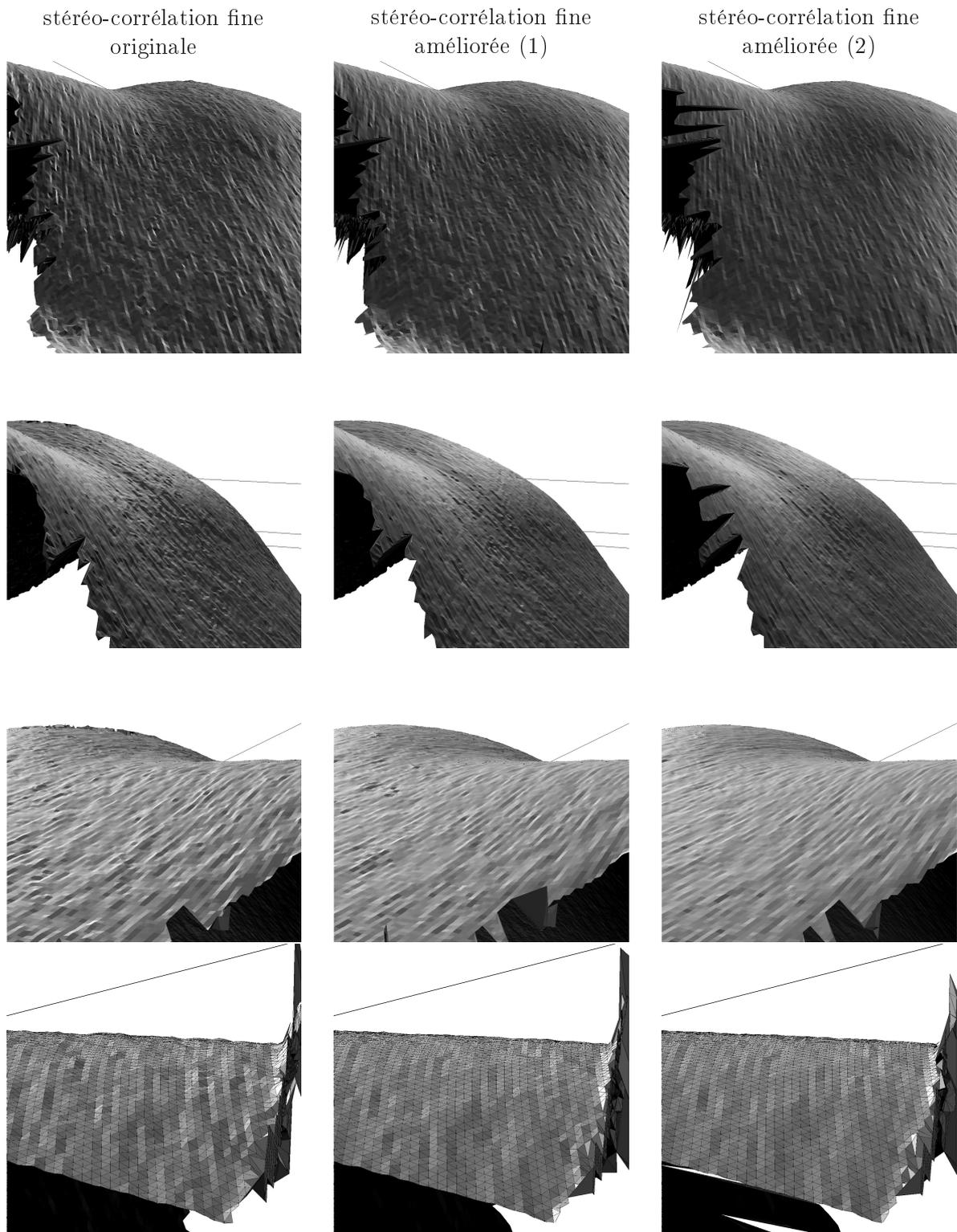


**FIG. 3.13 :** *Pièce emboutie à numériser.*

la plus inclinée par rapport aux plans images (contrainte fronto-parallèle non respectée), par conséquent mettant en difficulté les algorithmes d'appariements stéréoscopiques classiques. Ces résultats montrent clairement l'amélioration apportée par notre méthode de stéréo-corrélation fine améliorée, notamment en utilisant un modèle de luminance par spline biquintique. En effet, la surface de l'objet reconstruit ne comporte pas d'irrégularité dans la zone détaillée et la stéréo-corrélation fine améliorée conduit à une forte diminution des erreurs (bosselettes) visible dans les résultats utilisant l'algorithme original.



**FIG. 3.14 :** *Trois vues du résultat de la numérisation 3D en utilisant notre méthode de stéréo-corrélation fine améliorée. Les erreurs grossières de reconstruction sont dues à des faux appariements causés par les reflets spéculaires dus à la disparition de la peinture dans les zones de frottements du procédé d'emboutissage.*



**FIG. 3.15 :** Comparaison de la reconstruction 3D utilisant la méthode de stéréo-corrélation fine originale et notre méthode. Les différentes vues du détail observé correspondent au « repli » gauche de l'embouti montré figure 3.14. Les résultats de la méthode originale sont montrés colonnes de gauche (interpolation bilinéaire des niveaux de gris), ceux de notre méthode avec le même type d'interpolation sont dans la colonne du milieu, et enfin, colonne de droite, sont montrés les résultats de notre méthode en utilisant une interpolation par spline biquintique des niveaux de gris.

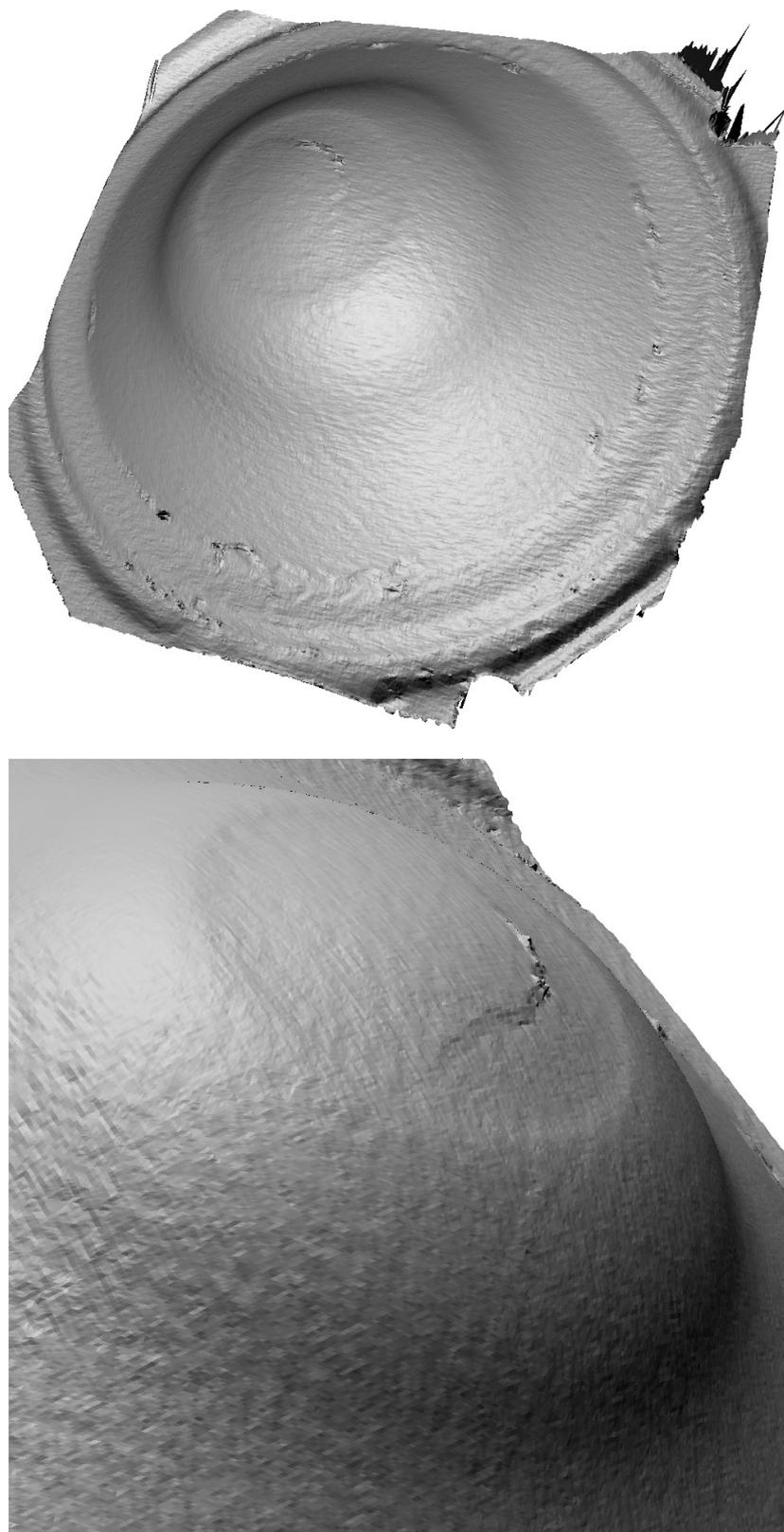
### 3.3.2 Pièce emboutie numéro 2

Nous avons procédé à un deuxième essai d'emboutissage similaire au précédent, mais ne présentant pas de défaut majeur de formage hormis la rupture de la tôle en son centre. La figure 3.16 montre le résultat de la reconstruction tridimensionnelle de cet objet, avec notamment un grossissement de son sommet pour apprécier la finesse avec laquelle la trace du contre-flan en acier est reconstruite, ainsi que la zone de la rupture de la tôle.

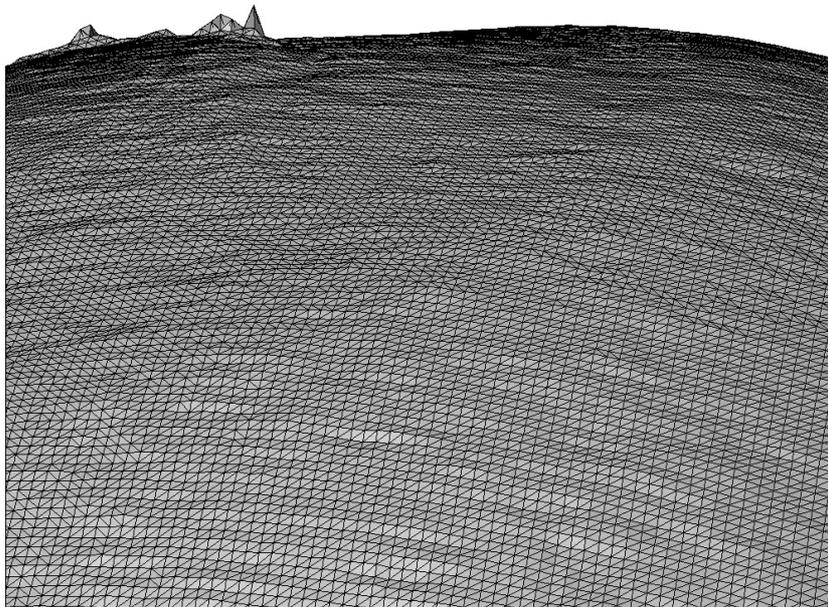
La figure 3.17 est une vue rasante du sommet de l'embouti montrant le maillage de la reconstruction tridimensionnelle. On peut y observer une certaine régularité de la surface reconstruite, ce qui est en accord avec la vraie pièce emboutie qui ne présente pas d'aspérité en surface. Cette illustration permet aussi de relativiser l'importance des défauts de type « bossette » apparaissant sur la figure 3.16 et amplifiés artificiellement par notre logiciel de rendu tridimensionnel.

Enfin, la figure 3.18 montre les deux cartes des dérivées partielles au premier ordre estimées de la disparité, respectivement selon l'axe vertical et horizontal des points exprimés dans l'espace rectifié. On peut y observer une bonne régularité de ces estimations qui correspondent à des variations douces de l'orientation de la surface par rapport aux caméras sur la majeure partie de l'objet. La trace du contre-flan est aussi visible près du centre de ces deux cartes au travers d'une variation locale de l'orientation de la surface, se traduisant par une variation des deux dérivées partielles au premier ordre de la disparité. Cette finesse de détail a pu être obtenue notamment grâce à l'utilisation de domaines de corrélation de faible dimension :  $11 \times 11$  pixels. À l'intérieur de la marque circulaire du contre-flan, on peut aussi observer l'effet de la discontinuité de la disparité liée à la rupture de la tôle qui produit dans ces deux cartes une discontinuité des dérivées partielles.

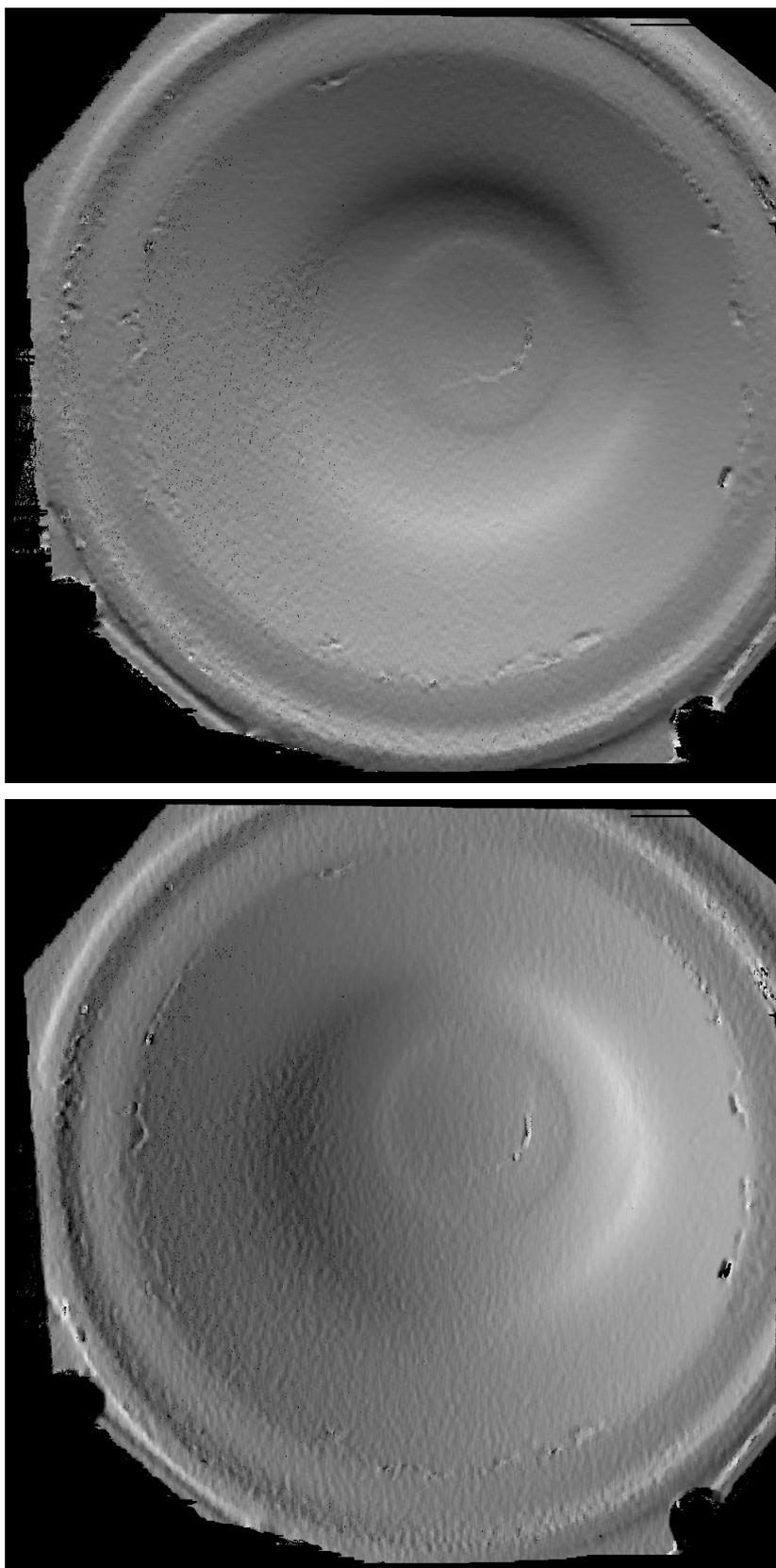
Cette reconstruction tridimensionnelle, utilisant notre méthode de stéréo-corrélation fine améliorée, est donc très satisfaisante sur le plan du bruit apparent dans les coordonnées tridimensionnelles reconstruites et sur la finesse des détails détectés en surface de la pièce observée.



**FIG. 3.16** : Reconstruction tridimensionnelle du second embouti en utilisant notre méthode de stéréocorrélation fine améliorée et un modèle de luminance par spline biquintique. On peut observer la marque circulaire du contre-flan en acier et la rupture de la tôle en son centre, notamment sur le grossissement de la figure du bas.



**FIG. 3.17 :** *Vue détaillée du sommet de la reconstruction tridimensionnelle du second embouti (au voisinage de la marque du contre-flan). La régularité de la surface maillée donne une indication qualitative sur la finesse de la reconstruction obtenue.*



**FIG. 3.18 :** Carte des dérivées partielles estimées de la disparité selon les lignes (en haut) et les colonnes (en bas), pour un modèle de transformation stéréoscopique locale d'ordre 1 en utilisant notre méthode de stéréo-corrélation fine améliorée et un modèle de luminance par spline biquintique.

### 3.3.3 Rapidité

Nous n'avons pas cherché spécialement à mettre au point des implantations rapides de nos algorithmes de stéréoscopie par corrélation. Le but principal recherché a été avant tout l'obtention de mesures les plus précises possible. Toutefois, nous avons mis en œuvre quelques techniques qui nous ont permis de réduire les temps de calcul pour obtenir des résultats dans un délai raisonnable.

La première de ces techniques concerne l'approche pyramidale pour l'appariement de paires d'images présentée dans la section 3.1.3. Tout en rendant plus robuste la mise en correspondance de points, elle permet de réduire le temps d'obtention de la carte de disparités au pixel près par rapport à une méthode directe traitant directement la paire d'images au niveau le plus fin (*i.e.* celles de niveau 0). Ainsi, le temps de calcul a été réduit de plusieurs minutes à quelques dizaines de secondes. Nous aurions pu encore améliorer significativement ce délai en mettant en œuvre une technique de *box-filtering* pour le calcul des scores de corrélation qui permet de réduire sa complexité algorithmique en un temps indépendant de la dimension des domaines de corrélation [McD81]. Malheureusement, ce genre d'approche rend très complexe la compréhension des algorithmes codés et nous avons préféré assurer une bonne lisibilité de notre code au dépend de ce gain potentiel en temps de calcul.

Concernant la phase de stéréo-corrélation fine améliorée présentée dans la section 3.1.5, nous avons pré-calculé les coordonnées corrigées de leur distorsion et rectifiées de chacun des points de l'image gauche. Cette transformation, que nous avons noté  $\Gamma^{-1}$ , n'introduit donc aucune pénalité en terme de temps de calcul. Nous avons en revanche effectué littéralement les calculs de la transformation de dé-rectification/distorsion  $\Gamma'$  pour chacune des coordonnées idéales/rectifiées de l'image droite. Ce choix provient du fait que ces coordonnées calculées ne sont pas entières : un tableau de correspondances directes pré-calculées n'est pas possible. Il est toutefois envisageable de supprimer ce coût en temps de calcul en pré-calculant un tableau de correspondances contenant les évaluations de la transformation  $\Gamma'$  pour chacune des coordonnées entières, puis en procédant par interpolation (par surface B-spline par exemple). La perte en précision devrait être minimale étant donné la quasi-linéarité de la transformation  $\Gamma'$  sur un voisinage de quelques pixels. Cette amélioration technique permettrait de se ramener à des temps de calcul similaires à la méthode de stéréo-corrélation fine utilisant une paire d'images rectifiées. Nous avons aussi modifié l'implantation de l'algorithme de LEVENBERG-MARQUARDT proposée dans [PTVF92] pour la minimisation des critères de type SSD, afin de prendre en compte l'aspect symétrique de la matrice à inverser dans l'étape 3 de l'algorithme décrit dans l'annexe F. Nous avons donc remplacé avantageusement la méthode de GAUSS-JORDAN utilisée initialement par celle proposée par RUTISHAUSER et décrite dans [Rut63], ce qui nous a permis d'accélérer notablement la minimisation.

Ainsi, ces quelques améliorations apportées nous ont permis d'obtenir chaque carte dense d'environ 1 million de disparités précises en environ 1 heure<sup>2</sup> sur une station de travail de type PC Pentium III cadencé à 850 MHz pour des domaines de corrélation de  $11 \times 11$  pixels. Nous pensons que les autres améliorations possibles, et notamment celle concernant le pré-calculs des coordonnées par la transformation  $\Gamma'$  devraient permettre de réduire encore davantage ce délai. De plus, à l'heure où nous écrivons cette section, des machines de type PC trois fois plus rapides (voire plus !) sont déjà sur le marché pour un coût relativement faible. Pour ces raisons, nous pensons que la méthode de stéréo-corrélation fine améliorée est d'ores et déjà

---

<sup>2</sup>Ce temps de calcul peut évidemment être réduit à quelques secondes si un faible nombre de points, tels que les nœuds d'un maillage de DELAUNAY, est considéré (*cf.* par exemple fig. 2.5, fig. 4.10 et fig. 4.11).

---

exploitable dans un cadre d'application de métrologie.

### 3.4 Conclusion

Nous souhaitons développer une technique de stéréoscopie par corrélation permettant la reconstruction tridimensionnelle précise de la surface d'un objet. Pour cela, nous nous sommes basé sur la même approche que celle utilisée dans le chapitre 2 en modélisant la transformation géométrique des domaines de pixels dans les critères de corrélation, pour le cas d'une paire d'images rectifiée. Cette approche nous a conduit à la méthode de stéréocorrélation fine de DEVERNAY. Nous nous sommes ensuite intéressé à son amélioration en exprimant la transformation géométrique des domaines de corrélation directement dans les images initiales plutôt que celle rectifiées et donc dégradées. L'intérêt de cette approche a été montré.



## Bibliographie du chapitre

- [Aya89] Nicolas AYACHE. *Vision stéréoscopique et perception multisensorielle*. Collection Science Informatique. InterEditions, 1989.
- [BF82] S. T. BARNARD et M. A. FISCHLER. Computational Stereo. *Computing Surveys*, 14(4):553–572, décembre 1982.
- [Bra95] Pascal BRAND. *Reconstruction tridimensionnelle à partir d'une caméra en mouvement : de l'influence de la précision*. Thèse de doctorat, Université Claude Bernard – Lyon I, octobre 1995.
- [DA89] U. R. DHOND et J. K. AGGARWAL. Structure from Stereo — A Review. *IEEE Transactions on Systems, Man and Cybernetics*, 19(16):1489–1510, novembre 1989.
- [Dev97] Frédéric DEVERNAY. *Vision stéréoscopique et propriétés différentielles des surfaces*. Thèse de doctorat, École Polytechnique (France), février 1997.
- [DF94] F. DEVERNAY et O. FAUGERAS. Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models. Rapport de recherche 2304, INRIA, juillet 1994.
- [Fau93] Olivier FAUGERAS. *Three-Dimensional Computer Vision : A Geometric Viewpoint*. The MIT Press, 1993. ISBN 0-262-06158-9.
- [Fua91] Pascal FUA. Combining Stereo and Monocular Information to Compute Dense Depth Maps that Preserve Depth Discontinuities. Dans *International Joint Conference on Artificial Intelligence (IJCAI'91)*, Sydney (Australie), août 1991.
- [HM95] R. HORAUD et O. MONGA. *Vision par ordinateur : outils fondamentaux*. Hermès, 1995. Deuxième édition.
- [HS97] R. I. HARTLEY et P. STURM. Triangulation. Dans *Computer Vision and Image Understanding (CVIU'97)*, volume 68 :2, pp. 146–157, novembre 1997.
- [Lan97] Zhong-Dan LAN. *Méthodes robustes en vision : application aux appariements visuels*. Thèse de doctorat, Institut National Polytechnique de Grenoble (France), mai 1997.
- [LG95] P. LASSERRE et P. GRANDJEAN. Stereovision improvements. Dans *7th International Conference on Advanced Robotics (ICAR'95)*, St Feliu de Guixols (Espagne), septembre 1995.
- [LZ99] C. LOOP et Z. ZHANG. Computing Rectifying Homographies for Stereo Vision. Rapport technique MSR-TR-99-21, Microsoft Research, avril 1999.
- [McD81] M. J. McDONNELL. Box-Filtering Techniques. *Computer Graphics and Image Processing*, 17:65–70, 1981.
- [PTVF92] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, et B. P. FLANNERY. *Numerical Recipes in C — The Art of Scientific Computing*. Cambridge University Press, édition 2nd, 1992.
- [Rut63] H. RUTISHAUSER. Algorithm 150 : syminv2. Dans *Communications of the ACM*, volume 6, pp. 67–68. ACM Press, New York, NY (USA), février 1963.
- [Zel96] Cyril ZELLER. *Calibration projective, affine et euclidienne en vision par ordinateur et application à la perception tridimensionnelle*. Thèse de doctorat, École Polytechnique (France), février 1996.

- [Zha93a] Zhengyou ZHANG. Le problème de la mise en correspondance : l'état de l'art. rapport de recherche 2146, INRIA, décembre 1993.
- [Zha93b] ChangSheng ZHAO. *Reconstruction de surfaces tridimensionnelles en vision par ordinateur*. Thèse de doctorat, Institut National Polytechnique de Grenoble (France), décembre 1993.

## Chapitre 4

# Mesure de champs de déplacements tridimensionnels par stéréo-corrélation et appariement temporel d'images

Nous souhaitons mesurer le champ de déplacements d'un objet par l'analyse d'une séquence de paires d'images stéréoscopiques de celui-ci, relatives à différents états de sa déformation. Généralement, le but poursuivi est le calcul des déformations de surface, qui peuvent être obtenus en suivant au cours du temps le déplacement de certains points répartis sur la surface de l'objet : on parlera de champ de déplacements.

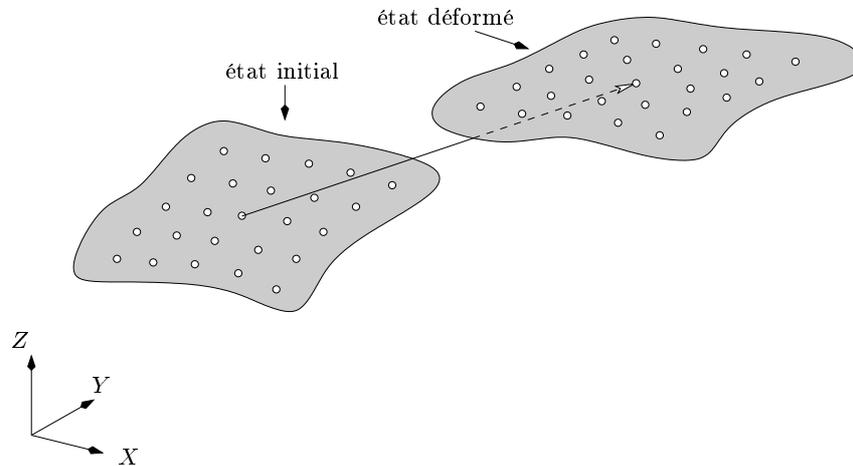
L'extensométrie bidimensionnelle discutée au chapitre 2 permet d'accéder seulement aux champs de déplacements bidimensionnels à la surface d'un objet plan. En couplant la stéréoscopie par corrélation (chapitre 3) et le suivi de pixels par corrélation (chapitre 2) nous pouvons accéder aux champs de déplacements tridimensionnels par la mise en correspondance de paires d'images acquises à différents instants. Une fois ces champs de déplacements déterminés, il est possible de calculer les déformations à chaque instant. Certains auteurs se sont intéressés à la mesure de champ de déplacements par appariement stéréoscopique de marqueurs [VL89], puis sont apparues les techniques de stéréo-corrélation [KJC90, LCSP93, OGD97, SS99, GH99, LAE00, SMHS00]. Nous allons commencer par étudier la technique classique utilisée par la plupart des chercheurs dans la section 4.1, puis nous proposerons une nouvelle méthode plus précise dans la section 4.2.

### 4.1 Méthode classique

L'approche classiquement utilisée pour la mesure de déplacements tridimensionnels par stéréo-corrélation d'images [SS99, GO99, CMB99, SMHC00] se décline en trois phases successives illustrées par la figure 4.2 :

- pour un point donné dans une image de la paire stéréoscopique à l'instant  $t$ , son homologue dans l'autre image est recherché par stéréo-corrélation en utilisant une technique telle que celle proposée au chapitre 3. Cette phase d'appariement conduit, par triangulation, à la connaissance des coordonnées tridimensionnelles du point considéré ;
- ce même point est suivi dans l'image acquise par la même caméra à l'instant  $t + \Delta_t$  par un algorithme tel que celui étudié au chapitre 2, afin de déterminer son homologue temporel ;

- une méthode similaire à celle de la première phase est mise en œuvre pour connaître les coordonnées tridimensionnelles du point à l'instant  $t + \Delta_t$ . Par simple différence des coordonnées des deux points tridimensionnels reconstruits, le vecteur déplacement tridimensionnel est alors déterminé (cf. figure 4.1).



**FIG. 4.1 :** Calcul du déplacement tridimensionnel d'un point à partir de ses deux reconstructions tridimensionnelles correspondantes aux deux états de déformation de l'objet.

Cette méthode présente l'inconvénient majeur d'apparier les points successivement, contribuant ainsi à combiner les incertitudes d'appariement sur la mesure finale. De plus, l'appariement temporel n'est effectué que sur les images d'une seule caméra (la caméra gauche dans l'illustration de la figure 4.2). Une amélioration possible consisterait à exploiter l'information redondante qu'apporterait l'appariement temporel dans les images de l'autre caméra, ou bien entre l'image initiale gauche et l'image déformée droite. Cependant, cela nécessiterait de définir une stratégie afin d'exploiter au mieux ces appariements supplémentaires pour améliorer la précision de la mesure finale. Plusieurs solutions sont possibles :

- outre les trois étapes de la méthode classique, l'appariement temporel dans les deux images de droite peut être effectué. Sur la figure 4.3, cela représente les étapes 1, 2, 3 et 2'. Deux appariements sont alors déterminés dans l'image droite après déformation (points a et b dans l'illustration). Le point médian sur le segment [ab] peut alors être calculé et utilisé pour reconstruire par triangulation les coordonnées tridimensionnelles du point correspondant de l'objet déformé. Toutefois, il est intéressant de noter que les équations de triangulation présentées dans la section 3.2, page 93, ne peuvent plus s'appliquer pour le cas présent étant donné que l'appariement dans l'image droite, après déformation, ainsi déterminé n'est plus nécessairement sur la droite épipolaire associée au point de l'image gauche après déformation. Cela signifie que les rayons optiques associés ne s'intersectent plus dans l'espace et que donc une méthode de triangulation appropriée doit être utilisée, par exemple celle décrite dans la section 1.3.3, page 25, ou bien l'une des méthodes proposées par HARTLEY et STURM dans [HS97] ;
- une autre solution consiste à effectuer les étapes classiques présentées en début de section (numérotées 1, 2 et 3 sur la figure 4.3) ainsi que l'étape repérée 3' dans l'illustration. Cela conduit au même problème évoqué dans la solution précédente, à savoir que deux appariements, notés a et c, sont alors candidats à la triangulation avec le point de

l'image gauche après déformation. Ici encore, le point médian sur le segment  $[ac]$  peut être utilisé dans les équations de triangulation, en sachant qu'ici aussi ce point n'est pas nécessairement sur la droite épipolaire associée au point de l'image gauche après déformation. Il est intéressant de noter que cette solution nécessite, pour l'obtention de l'appariement  $c$ , la prise en compte d'un modèle de transformation géométrique des domaines de corrélation incluant à la fois la transformation stéréoscopique locale et celle liée à la déformation locale de l'objet ;

- enfin, une solution basée sur l'utilisation des trois appariements  $a$ ,  $b$  et  $c$  est aussi possible. Le point barycentre du triangle formé par ceux-ci peut alors être utilisé dans les équations de triangulation (*idem* la remarque pour les deux solutions précédentes : ce barycentre n'est pas nécessairement sur la droite épipolaire associée au point de l'image gauche après déformation).

La prise en compte d'un deuxième, voire troisième, appariement dans l'image droite après déformation est aussi intéressante pour valider cette mise en correspondance. Il est en effet possible de se donner une distance maximale entre ces points au delà de laquelle la mise en correspondance sera considérée comme un échec.

Malgré cette amélioration potentielle de la méthode classique, le processus reste séquentiel et donc cumule les erreurs d'appariement entre chaque étape. La prise en compte simultanée des quatre images dans la mise en correspondance doit certainement permettre d'accroître la précision des mesures effectuées : c'est l'idée maîtresse qui nous a conduit à développer une nouvelle méthode, combinant simultanément les deux phases de stéréo-corrélation et le suivi de pixels, exposée dans la section suivante.

## 4.2 Suivi de pixels et stéréo-corrélation simultanés

Dans le but d'éviter un cumul des incertitudes d'appariement de la méthode classique, il semble intéressant d'estimer tous les paramètres des différentes transformations géométriques locales simultanément (*cf.* figure 4.4). Cette approche nécessite de définir un critère de corrélation basé sur les deux paires d'images : c'est l'objet de l'étude de la section 4.2.1. Nous utiliserons alors une technique d'optimisation similaire à celles utilisées aux chapitres 2 et 3 pour estimer simultanément tous les paramètres des diverses transformations géométriques. Dans ce but, nous déterminerons une paramétrisation minimale de ces transformations dans la section 4.2.2.

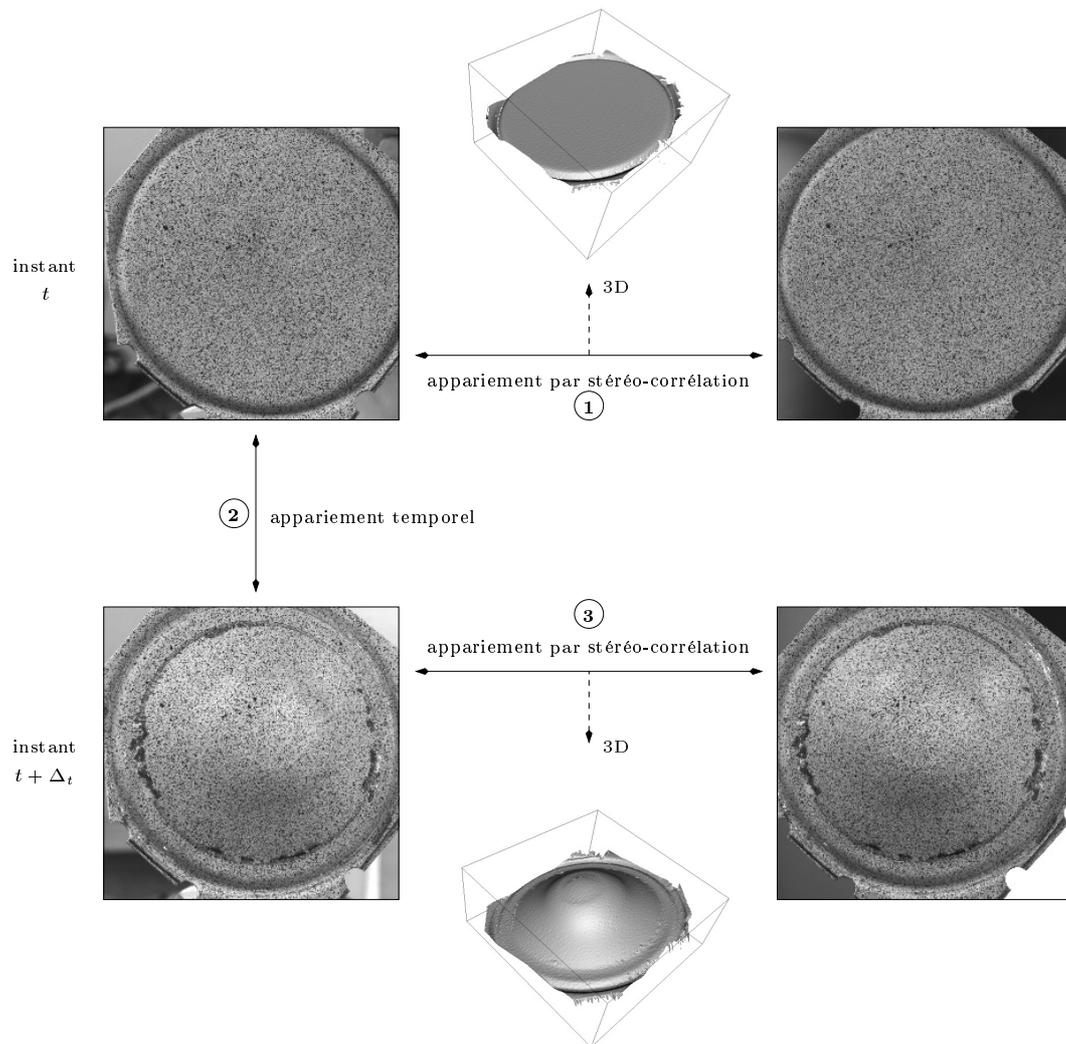
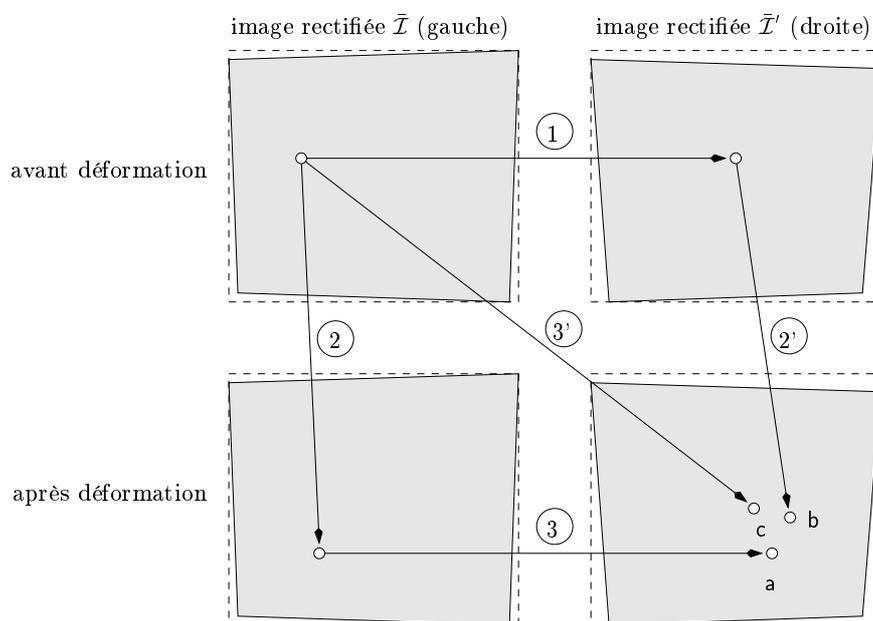


FIG. 4.2 : Méthode classique d'obtention du champ de déplacements tridimensionnels par stéréo-corrélation et suivi de pixels par corrélation.



**FIG. 4.3 :** L'appariement d'un point de l'image gauche avant déformation dans l'image droite après déformation peut être déterminé de trois manières (étapes numérotées 2', 3 et 3').

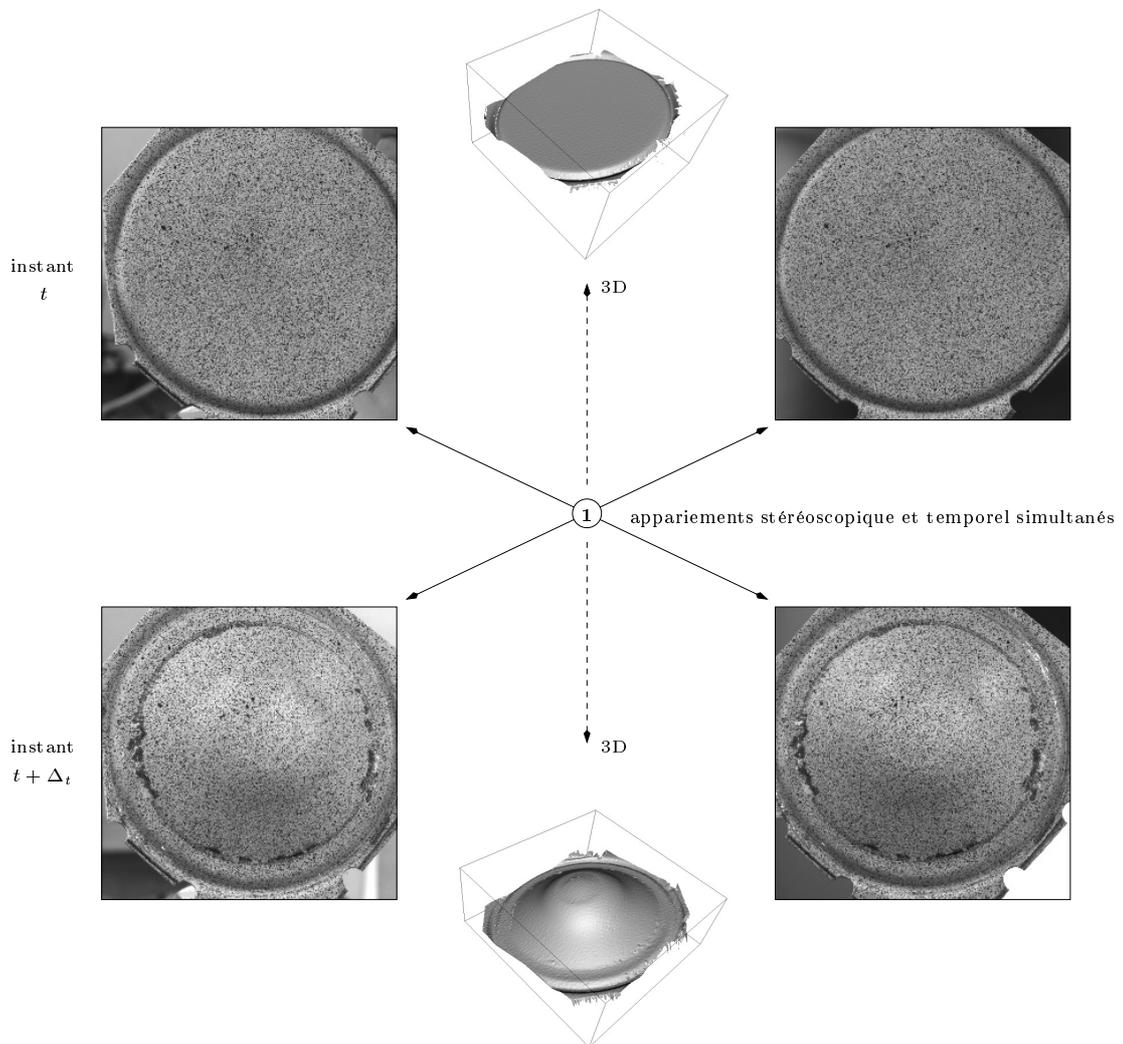
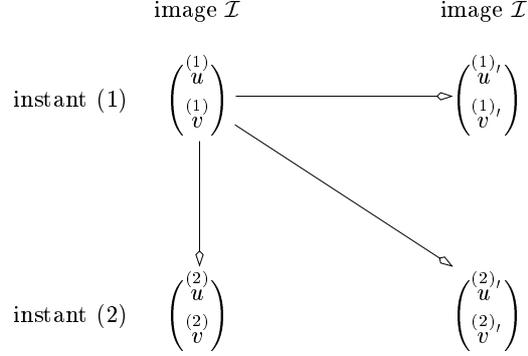


FIG. 4.4 : Obtention du champ de déplacements tridimensionnels par suivi de pixels et stéréoscopie par corrélation simultanés.

### 4.2.1 Critère de corrélation global



**FIG. 4.5 :** Notations utilisées pour désigner quatre points en correspondance stéréoscopique ou temporelle. Les flèches représentent les trois relations que nous souhaitons exprimer pour définir un critère de corrélation global.

Si nous considérons deux instants de déformation d'un objet, nous obtenons deux paires d'images stéréoscopiques de celui-ci à ces deux instants. Il est alors possible d'exprimer un critère de corrélation basé sur les quatre images en choisissant l'une de ces images comme référence. Sans perte de généralité, considérons l'image gauche au premier instant comme référence. Un critère de corrélation de type SSD mesurant simultanément la similarité du domaine de référence avec les trois autres domaines est aisé à définir : il consiste en la somme des différences au carré des niveaux de gris entre le domaine de pixels de l'image de référence et chacun des trois domaines des autres images. Nous notons  $x^{(t)}$  toute quantité  $x$  considérée à l'instant  $t$ . D'après la figure 4.5, il vient :

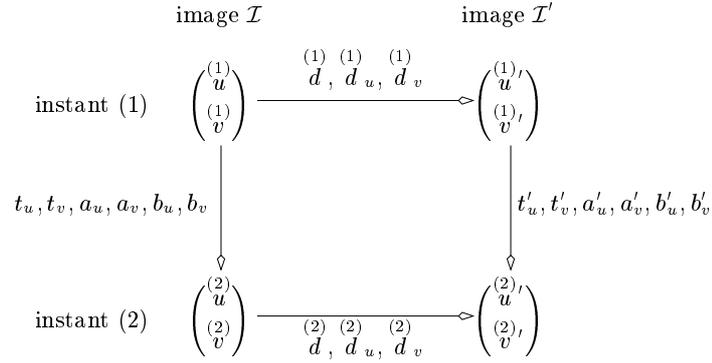
$$C(\mathbf{m}_c, \mathbf{v}) = \sum_{\substack{(1) \\ \mathbf{m}_i \in V(\mathbf{m}_c)}} \left[ \left( \mathcal{I}'(\mathbf{m}'_i) - \mathcal{I}(\mathbf{m}_i) \right)^2 + \left( \mathcal{I}(\mathbf{m}_i) - \mathcal{I}(\mathbf{m}_i) \right)^2 + \left( \mathcal{I}'(\mathbf{m}'_i) - \mathcal{I}(\mathbf{m}_i) \right)^2 \right] \quad (4.1)$$

Le vecteur  $\mathbf{v}$  représente l'ensemble des paramètres des différentes transformations géométriques locales entre le domaine de pixels de référence et les trois autres domaines. Ces paramètres sont fonction de l'ordre utilisé des approximations et sont étudiés dans la section suivante.

### 4.2.2 Paramétrisation minimale des transformations géométriques locales

Afin de simplifier l'exposé, nous limiterons notre étude aux transformations géométriques locales d'ordre 1, cf. figure 4.6. Une paramétrisation minimale possible des trois transformations montrées figure 4.5 comporte seulement 12 paramètres au lieu de 18 (cf. annexe E pour une démonstration) :

$$\mathbf{v} = \left( d^{(1)} \ d_u^{(1)} \ d_v^{(1)} \ d^{(2)} \ d_u^{(2)} \ d_v^{(2)} \ t_u \ t_v \ a'_u \ a'_v \ b'_u \ b'_v \right)^\top$$



**FIG. 4.6 :** Les différents paramètres des transformations géométriques locales à l'ordre 1 tels que définis dans les chapitres 2 et 3. Ces 18 paramètres ne sont pas tous indépendants les uns des autres.

Les six autres paramètres se déduisent de ceux-ci par les relations de dépendance suivantes :

$$\begin{aligned}
 t'_u &= t_u \\
 t'_v &= \overset{(2)}{d} - \overset{(1)}{d} + t_v \\
 a_u &= a'_u + a'_v \overset{(1)}{d}_u \\
 a_v &= a'_v \left(1 + \overset{(1)}{d}_v\right) \\
 b_u &= \frac{b'_u + \overset{(1)}{d}_u \left(1 + b'_v\right) - \overset{(2)}{d}_u \left(1 + a'_u + a'_v \overset{(1)}{d}_u\right)}{1 + \overset{(2)}{d}_v} \\
 b_v &= \frac{b'_v + \overset{(1)}{d}_v \left(1 + b'_v\right) - \overset{(2)}{d}_v - a'_v \overset{(2)}{d}_u \left(1 + \overset{(1)}{d}_v\right)}{1 + \overset{(2)}{d}_v}
 \end{aligned}$$

En posant  $\delta_u = \overset{(1)}{u} - \overset{(1)}{u}_c$  et  $\delta_v = \overset{(1)}{v} - \overset{(1)}{v}_c$ , les trois transformations géométriques locales qui nous intéressent dans l'expression du critère de corrélation (4.1) s'écrivent alors sous les formes

paramétrées minimales suivantes :

$$\begin{aligned}
{}^{(1)}u &= {}^{(1)}u \\
{}^{(1)}v &= {}^{(1)}v + d + \delta_u {}^{(1)}d_u + \delta_v {}^{(1)}d_v \\
{}^{(2)}u &= {}^{(1)}u + t_u + \delta_u (a'_u + a'_v {}^{(1)}d_u) + \delta_v a'_v (1 + {}^{(1)}d_v) \\
{}^{(2)}v &= {}^{(1)}v + t_v + \delta_u \frac{b'_u + {}^{(1)}d_u (1 + b'_v) - {}^{(2)}d_u (1 + a'_u + a'_v {}^{(1)}d_u)}{1 + {}^{(2)}d_v} + \\
&\quad \delta_v \frac{b'_v + {}^{(1)}d_v (1 + b'_v) - {}^{(2)}d_v - a'_v {}^{(2)}d_u (1 + {}^{(1)}d_v)}{1 + {}^{(2)}d_v} \\
{}^{(2)}u' &= {}^{(1)}u + t_u + \delta_u (a'_u + a'_v {}^{(1)}d_u) + \delta_v (a'_v + a'_v {}^{(1)}d_v) \\
{}^{(2)}v' &= {}^{(1)}v + d + t_v + \delta_u (d_u + b'_u + b'_v d_u) + \delta_v (d_v + b'_v + b'_v d_v)
\end{aligned}$$

### 4.2.3 Estimation des paramètres

Nous venons d'étudier une paramétrisation minimale des trois transformations géométriques locales intervenant dans l'expression du critère de corrélation global donné par l'équation (4.1). Cette étude préalable est nécessaire pour les processus de minimisation afin d'éviter l'emploi de techniques d'optimisation sous contraintes qui sont généralement plus lentes et moins robustes. Le critère de corrélation donné par l'équation (4.1) est exprimé sous la forme d'une somme de résidus quadratiques : l'emploi de la méthode de LEVENBERG-MARQUARDT est donc possible si l'on sait exprimer les dérivées partielles de chaque résidu par rapport à chacun des paramètres à estimer, et si l'on peut obtenir une estimée initiale du vecteur de paramètres.

Afin de bénéficier des mêmes performances que permet d'atteindre notre méthode de stéréocorrélation fine améliorée présentée au chapitre 3, nous estimons les paramètres en utilisant seulement les deux paires d'images initiales, *i.e* non rectifiées. L'annexe G.3, page 175, détaille le calcul des dérivées partielles du critère de corrélation par rapport à chacun des paramètres à estimer.

Une estimée initiale de chacun des 12 paramètres peut être obtenue facilement en procédant selon l'approche classique par corrélation décrite dans la section 4.1. En pratique, nous avons pu observer que seules les disparités  $d^{(1)}$  et  $d^{(2)}$ , et les composantes  $t_u$  et  $t_v$  du vecteur de translation ont besoin d'être fournies assez précisément (au pixel près) pour faire converger la méthode d'optimisation ; les autres paramètres pouvant tous être initialisés à 0. En général, la convergence est atteinte en moins de 10 itérations.

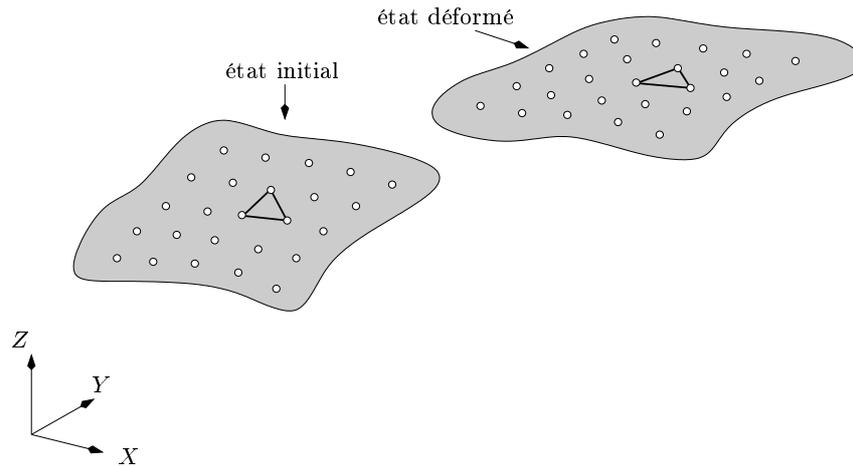
Les résultats d'application de cette technique sont présentés au chapitre 5, section 5.1, page 125.

## 4.3 Calcul du champ de déformations

Nous venons de présenter, dans les sections précédentes, deux techniques basées sur la corrélation d'images permettant de mesurer des déplacements de points en surface d'un objet déformé. De nombreuses applications en mécanique expérimentale nécessitent de connaître

le champ de déformations associé au champ de déplacements ainsi mesuré. Ce calcul des déformations est alors effectué par analyse de la transformation reliant les deux formes tridimensionnelles respectivement avant et après la sollicitation de l'objet. Cette analyse est généralement effectuée sur les éléments d'un maillage triangulaire des points tridimensionnels reconstruits, et se décline selon deux approches.

La première méthode consiste à analyser la transformation subie par chacun des triangles entre la géométrie à l'état initial et celle à l'état déformé [SDC86, MD92]. Les déformations ainsi calculées sont alors associées à chacun de leur triangle respectif (*cf.* figure 4.7).



**FIG. 4.7 :** *Calcul de la déformation par analyse de la transformation reliant un triangle entre son état initial et son état déformé. La déformation ainsi calculée est associée au triangle, et non pas à un nœud comme pour les déplacements.*

Une deuxième méthode s'attache à calculer les déformations non plus pour chaque triangle, mais en chaque nœud du maillage en prenant en compte un ensemble de triangles voisins de chaque nœud [Dou00] (*cf.* figure 4.8). Cette approche permet de réduire l'amplitude du bruit des déformations calculées mais conduit à une perte du caractère local de la mesure d'autant plus grande que le voisinage utilisé est étendu [Dou00]. Cela peut alors être gênant si le champ de déformation réel est hétérogène ou présente de forts gradients.

Nous avons développé une interface entre notre logiciel et le logiciel de calcul par éléments finis ABAQUS™ pour exporter les résultats de mesure et calculer les déformations. La surface initiale est le résultat de la reconstruction tridimensionnelle de l'objet avant déformation ; le maillage triangulaire des points de cette surface correspond à celui utilisé lors de l'appariement dans les images. Nous utilisons dans ce logiciel des éléments de type « coque mince » qui nous permettent d'attribuer à chacun des nœuds le déplacement tridimensionnel mesuré. ABAQUS™ calcule ensuite les déformations pour chacun des éléments (similaire à la première méthode exposée).

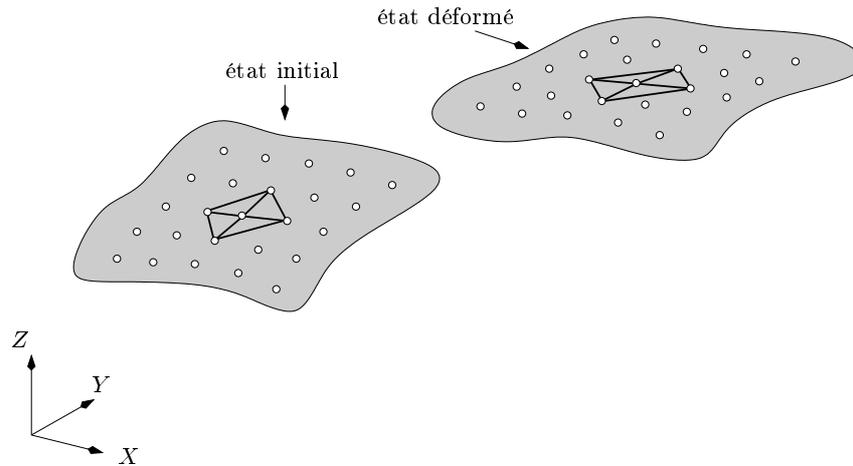


FIG. 4.8 : Calcul de la déformation sur un ensemble de triangles voisins d'un nœud. La déformation ainsi calculée est associée au nœud considéré.

#### 4.4 Extension à la mesure de champ de déplacements en dynamique

Certaines applications nécessitent la connaissance des champs de déplacements/déformations à plusieurs instants de sollicitation de l'objet. C'est par exemple le cas pour l'étude du comportement de membranes en élastomère soufflées (cf. chapitre 5). La méthode consiste alors à analyser hors-ligne (pas de temps réel) une séquence de paires d'images, acquises à différents instants, pour mesurer les champs de déplacements.

Nous avons présenté, dans les sections précédentes, deux méthodes de détermination d'un champ de déplacements à partir de deux paires d'images stéréoscopiques. Nous pouvons alors reprendre le principe évoqué dans la section 2.3, figure 2.6, 54, pour étendre ces méthodes au cas d'une séquence de plus de deux paires d'images. Nous en rappelons brièvement le principe général. Il consiste à appairer les points entre la première paire d'images et la paire d'images d'indice  $i$  dans la séquence. Les points de la paire d'images d'indice  $i$  sont appariés dans la paire d'images d'indice  $i + 1$  afin de fournir des estimées pour l'appariement des points entre la première paire d'images et la paire d'images d'indice  $i + 1$ . Cette approche évite de cumuler les erreurs d'appariement en se ramenant toujours à un problème d'appariement entre la première paire d'images et chacune des autres paires.

Des résultats d'application de cette méthode sont donnés dans le chapitre 5, section 5.2, page 132.

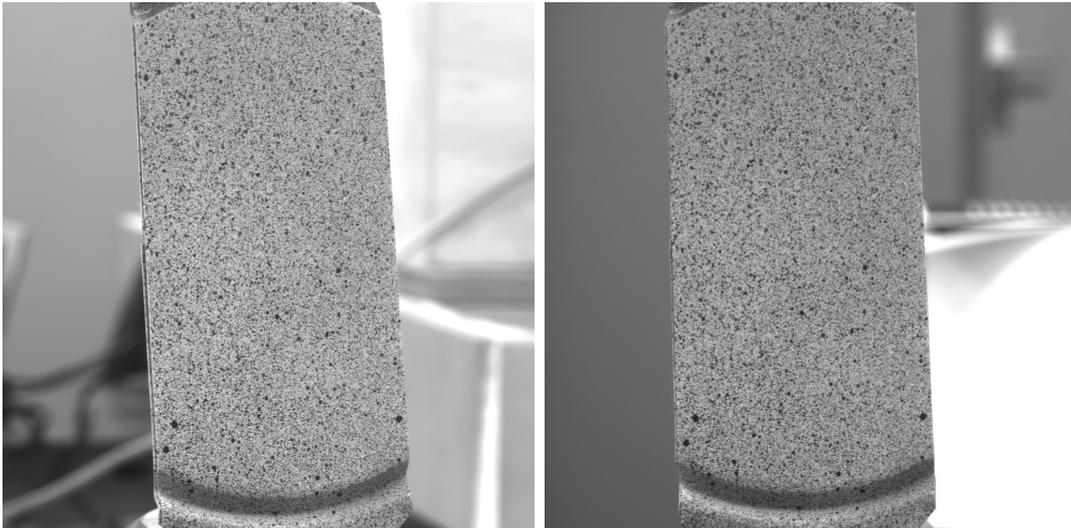
#### 4.5 Évaluation des performances

Pour évaluer les performances de notre méthode, il faut connaître la « vérité terrain ». Or, dans le cas de la mesure de déplacements en surface d'un objet déformé, il est très difficile de l'obtenir. En effet, les jauges de déformations ne sont pas adaptées du fait de leur dimension trop importante par rapport au champ de vue analysé (ou bien cela nécessiterait la mise en place d'une expérience de déformation d'un objet de grande dimension, ce qui est expérimentalement difficile à réaliser). De plus, ces capteurs ne donnent qu'une information ponctuelle,

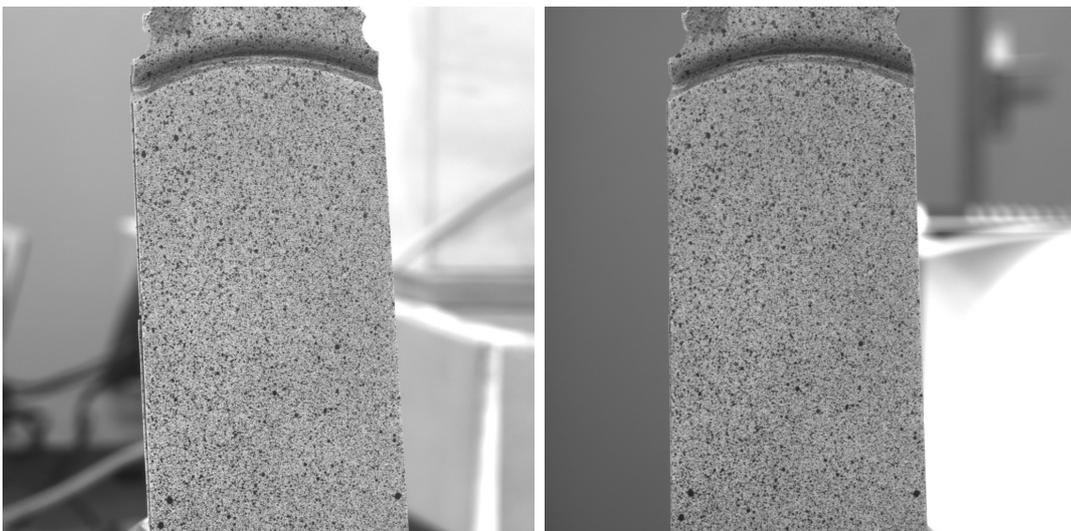
par opposition aux méthodes optiques que nous souhaitons comparer et qui fournissent une information de champ. Nous avons alors opté pour la seule expérience dont la « vérité terrain » est connue *a priori* : l'expérience « **zéro déformation** ».

L'expérience « zéro déformation » consiste à acquérir deux paires d'images stéréoscopiques d'un objet ayant subi une simple transformation de corps rigide. La mesure d'un champ de déplacements entre les deux positions de cet objet est alors effectuée, et le calcul du champ de déformations associé doit conduire à des déformations nulles en tout point.

La figure 4.9 montre les deux paires d'images stéréoscopiques de cette expérience : un objet plan texturé artificiellement par projection de fines gouttelettes de peinture blanche et noire, et translaté entre les deux instants d'acquisition selon l'axe vertical. Les figures 4.10 et 4.11



paire d'images stéréoscopiques de l'objet dans sa première position



paire d'images stéréoscopiques de l'objet dans sa deuxième position

**FIG. 4.9 :** Deux paires d'images stéréoscopiques d'un objet plan dans deux positions différentes. Le champ de déformation en surface de cet objet est théoriquement nul.

montrent respectivement le résultat de l'appariement stéréoscopique et temporel d'environ 150 nœuds, respectivement 3000 nœuds, qui composent le maillage défini dans l'image gauche de la paire stéréoscopique initiale. Le calcul des déformations en surface de cette pièce a été effectué à partir des mesures de déplacements obtenues selon notre méthode avec des domaines de corrélation de  $21 \times 21$  pixels, et en utilisant le logiciel ABAQUS<sup>TM</sup>. Nous avons alors obtenu deux champs de déformations correspondant aux deux maillages utilisés, avec un calcul de déformation par triangle. Puisque théoriquement nous devrions obtenir des déformations nulles en tout point, nous avons caractérisé la dispersion de ces mesures pour chacun des deux maillages. Le tableau 4.1 montre les résultats obtenus. On peut y observer que l'utilisation d'un maillage

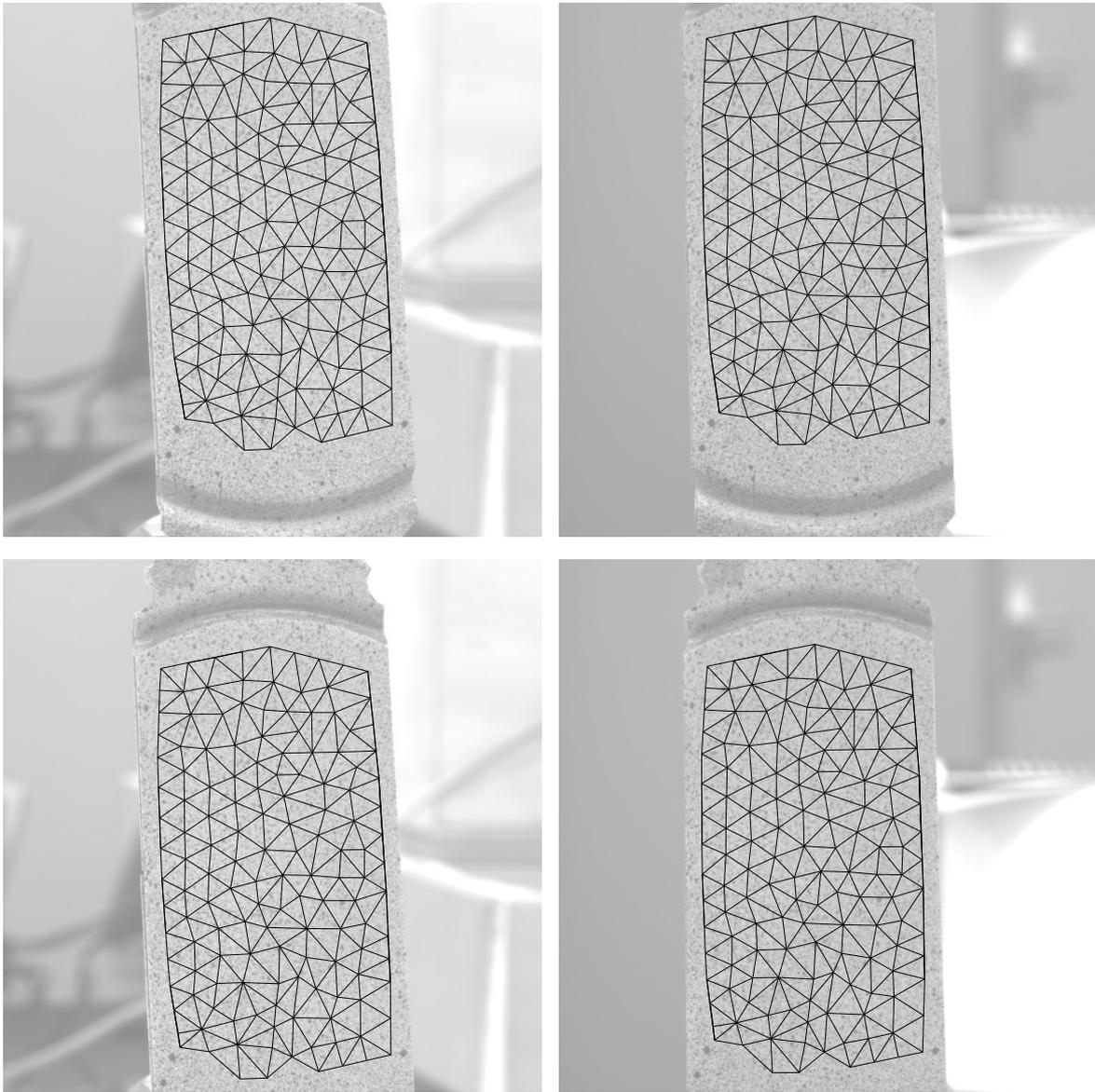
maillage	moyenne		écart-type	
	déf. max.	déf. min.	déf. max.	déf. min.
fin	-0,1150%	0,1923%	0,1359%	0,1380%
grossier	-0,0539%	0,0693%	0,0488%	0,0704%

**TAB. 4.1 :** *Dispersion des résultats de calcul des déformations maximales (déf. max.) et minimales (déf. min.) pour l'expérience « zéro déformation », pour les deux maillages utilisés.*

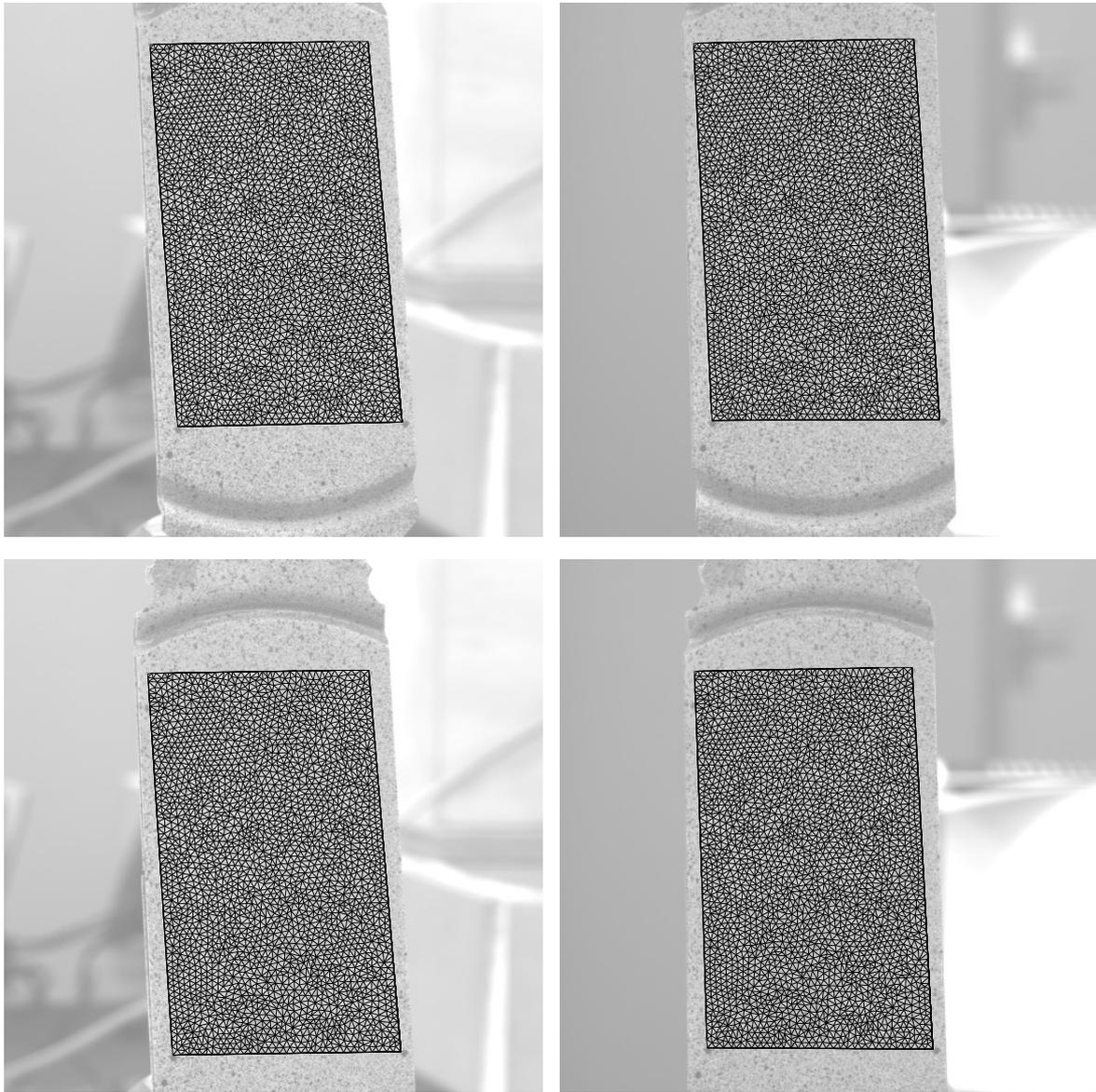
composé de triangles plus gros permet de réduire l'incertitude sur le calcul des déformations par rapport à un maillage plus fin. Cela s'explique facilement par le fait que les incertitudes des appariements (identiques dans les deux cas) rapportées aux dimensions des triangles est plus faible dans le cas d'un maillage grossier que pour le cas d'un maillage fin. Néanmoins, l'utilisation d'un maillage plus grossier qui fournit une meilleure précision de mesure conduit à une perte du caractère local de la mesure ce qui peut être gênant si le champ de déformation réel est hétérogène ou présente de forts gradients. Ce tableau met en évidence une précision de calcul des déformations de l'ordre de 0,1% pour des petites mailles de taille  $12 \times 12$  pixels. Il est à noter cependant que cette valeur a été obtenue sur un objet plan et non distordu. Il est probable que lors du calcul des déformations d'une surface gauche, les approximations que nous avons faites pour la modélisation des transformations géométriques locales des domaines de corrélation, à la fois pour le suivi de pixel et la stéréoscopie, dégradent ces performances. Une étude complémentaire pour quantifier la précision de calcul des déformations devra être menée.

## 4.6 Conclusion

Ce chapitre répond à la deuxième problématique majeure de cette thèse : proposer un outil de mesure pour la mécanique expérimentale permettant la mesure de champs de déplacements tridimensionnels et le calcul de champs de déformations de surface. Pour cela, nous proposons une nouvelle méthode combinant simultanément la stéréoscopie et le suivi de pixel par corrélation. Contrairement à l'approche classique qui consiste à préalablement rectifier et corriger la distorsion des images, l'appariement de points est effectué directement dans les images acquises par le capteur de stéréovision. Cela évite donc de traiter des images dégradées par ces opérations, et donc d'accroître la précision des mesures effectuées. Une évaluation des performances de notre approche nous donne une indication sur la précision des déformations calculées de l'ordre de 0,1% pour des petites mailles de taille  $12 \times 12$  pixels.



**FIG. 4.10 :** *Suivi des nœuds d'un maillage triangulaire par stéréo-corrélation et suivi de pixels simultanés. Le maillage a été initialement défini dans l'image gauche de la paire stéréoscopique initiale (images du haut). Ce maillage grossier comporte environ 150 nœuds et 250 triangles d'environ 50 pixels de côté.*



**FIG. 4.11 :** *Suivi des nœuds d'un maillage triangulaire par stéréo-corrélation et suivi de pixels simultanés. Le maillage a été initialement défini dans l'image gauche de la paire stéréoscopique initiale (images du haut). Ce maillage fin comporte environ 3000 nœuds et 5000 triangles d'environ 12 pixels de côté.*



## Bibliographie du chapitre

- [CMB99] P. CLERC, F. MORESTIN, et M. BRUNET. Measurement of Deformation Fields Using Numerical Images Correlation. Dans *International Conference on Advances in Materials and Processing Technologies (AMPT'99)*, pp. 1889–1897, Dublin (Irlande), août 1999.
- [Dou00] Pascal DOUMALIN. *Microextensométrie locale par corrélation d'images numériques. Application aux études micromécaniques par microscopie électronique à balayage*. Thèse de doctorat, École Polytechnique (France), juin 2000.
- [GH99] K. GALANULIS et A. HOFMANN. Determination of Forming Limit Diagrams Using an Optical Measurement System. Dans *7th International Conference on Sheet Metal (SheMet'99)*, pp. 245–252, Erlangen (Allemagne), septembre 1999.
- [GO99] D. GARCIA et J.-J. ORTEU. 3D Deformation Measurement using Stereo-Correlation applied to the Forming of Metal or Elastomer Sheets. Dans *International Workshop on Video-Controlled Materials Testing and In-Situ Microstructural Characterization*, Nancy (France), novembre 1999.
- [HS97] R. I. HARTLEY et P. STURM. Triangulation. Dans *Computer Vision and Image Understanding (CVIU'97)*, volume 68 :2, pp. 146–157, novembre 1997.
- [KJC90] Z. L. KHAN-JETTER et T. C. CHU. Three-dimensional displacement measurements using digital image correlation and photogrammic analysis. *Experimental Mechanics*, 30(1):10–16, 1990.
- [LAE00] N. LANGERAK, E. ATZEMA, et J. ESSING. Strain measurements with the PHAST™ system. Dans *International Deep Drawing Research Group (IDDRG'2000) — Meeting of the Working Groups*, Ann Arbor, Michigan (USA), 2000.
- [LCSP93] P. F. LUO, Y. J. CHAO, M. A. SUTTON, et W. H. PETERS. Accurate measurement of three-dimensional deformable and rigid bodies using computer vision. *Experimental Mechanics*, 33(2):123–132, 1993.
- [MD92] Z. MARCINIAK et J. L. DUNCAN. *The Mechanics of Sheet Metal Forming*. Edward Arnold, 1992.
- [OGD97] J.J. ORTEU, V. GARRIC, et M. DEVY. Camera calibration for 3D reconstruction : application to the measure of 3D deformations on sheet metal parts. Dans *European Symposium on Lasers, Optics and Vision in Manufacturing*, Munich (Allemagne), juin 1997.
- [SDC86] R. SOWERBY, J. L. DUNCAN, et E. CHU. The Modelling of Sheet Metal Stamping. *International Journal of Mechanical Science*, 28(7):415–430, 1986.
- [SMHC00] M. A. SUTTON, S. R. MCNEILL, J. D. HELM, et Y. J. CHAO. Advances in Two-Dimensional and Three-Dimensional Computer Vision. *Photomechanics, Topics Applied Physics*, 77:323–372, 2000.
- [SMHS00] M. A. SUTTON, S. R. MCNEILL, J. D. HELM, et H. W. SCHREIER. Computer vision applied to shape and deformation measurement. Dans *International Conference on Trends in Optical Nondestructive Testing and Inspection, Elsevier Science*, pp. 571–589, Lugano (Suisse), mai 2000.
- [SS99] P. SYNNERGREN et M. SJÖDAHL. A stereoscopic digital speckle photography system for 3-D displacement field measurements. *Optics and Lasers in Engineering*, 31:425–443, 1999.

- [VL89] J. H. VOGEL et D. LEE. An Automated Two-View Method For Determining Strain Distributions on Deformed Surfaces. Dans *Journal of Material Shaping Technology*, volume 6, 1989.

# Chapitre 5

## Applications

### 5.1 Tôles minces embouties

Dans le domaine de l'emboutissage, il est important de connaître la limite supérieure de formage admissible pour une tôle avant l'apparition de la striction localisée. Pour obtenir cette limite, on procède généralement à un essai expérimental et on mesure le champ de déformations sur la tôle emboutie [VL89, MBB95, GH99]. Nous illustrons dans cette section l'application de la technique de stéréo-corrélation et suivi de pixels simultanés pour l'obtention du champ de déplacements tridimensionnels sur une tôle mince emboutie. La figure 5.1 illustre le principe de cet essai d'emboutissage, le schéma de la figure 5.2 montre une coupe du flan à emboutir. La figure 5.3 montre la configuration de cette expérience<sup>1</sup>, avec notamment les deux caméras numériques Hamamatsu de résolution  $1280 \times 1024$  pixels, 8 bits par pixel, deux objectifs de focale 25 mm, et une vue de l'embouti à numériser. Le calibrage de ce capteur stéréoscopique a été très satisfaisant car l'évaluation du critère épipolaire sur la projection des  $8 \times 8$  points de la mire réestimées nous a conduit à une erreur meilleure que  $\frac{1}{80}$  pixel. Cet objet à une hauteur d'environ 14 cm, une largeur de 9,5 cm et une profondeur de 3 cm. Il est placé à une distance de 110 cm du capteur de stéréovision dont la base est d'environ 30 cm. Sa surface a été recouverte d'une fine couche de peinture blanche et de gouttelettes de peinture noire pour lui conférer un aspect convenablement texturé et améliorer les performances des méthodes d'appariement par corrélation. Il est à noter que cette opération ne nécessite qu'à peine quelques minutes, ce qui est un avantage important par rapport aux méthodes classiques de suivi de marqueurs dont le dépôt est beaucoup plus difficile à réaliser. Cette préparation a été effectuée préalablement à l'opération d'emboutissage ce qui explique que des parties de l'objet ayant subi des frottements aient perdu leur texture, révélant de ce fait l'aspect métallique de la tôle d'aluminium.

---

<sup>1</sup>Nous remercions messieurs Fabrice MORESTIN et Patrice CLERC du laboratoire de mécanique des solides (LMS) de l'Institut National des Sciences Appliquées de Lyon (INSAL) pour avoir accepté de faire avec nous à Lyon, avec leur presse d'emboutissage, une expérience de mesure de champs de déplacements 3D par stéréo-corrélation sur pièces embouties.

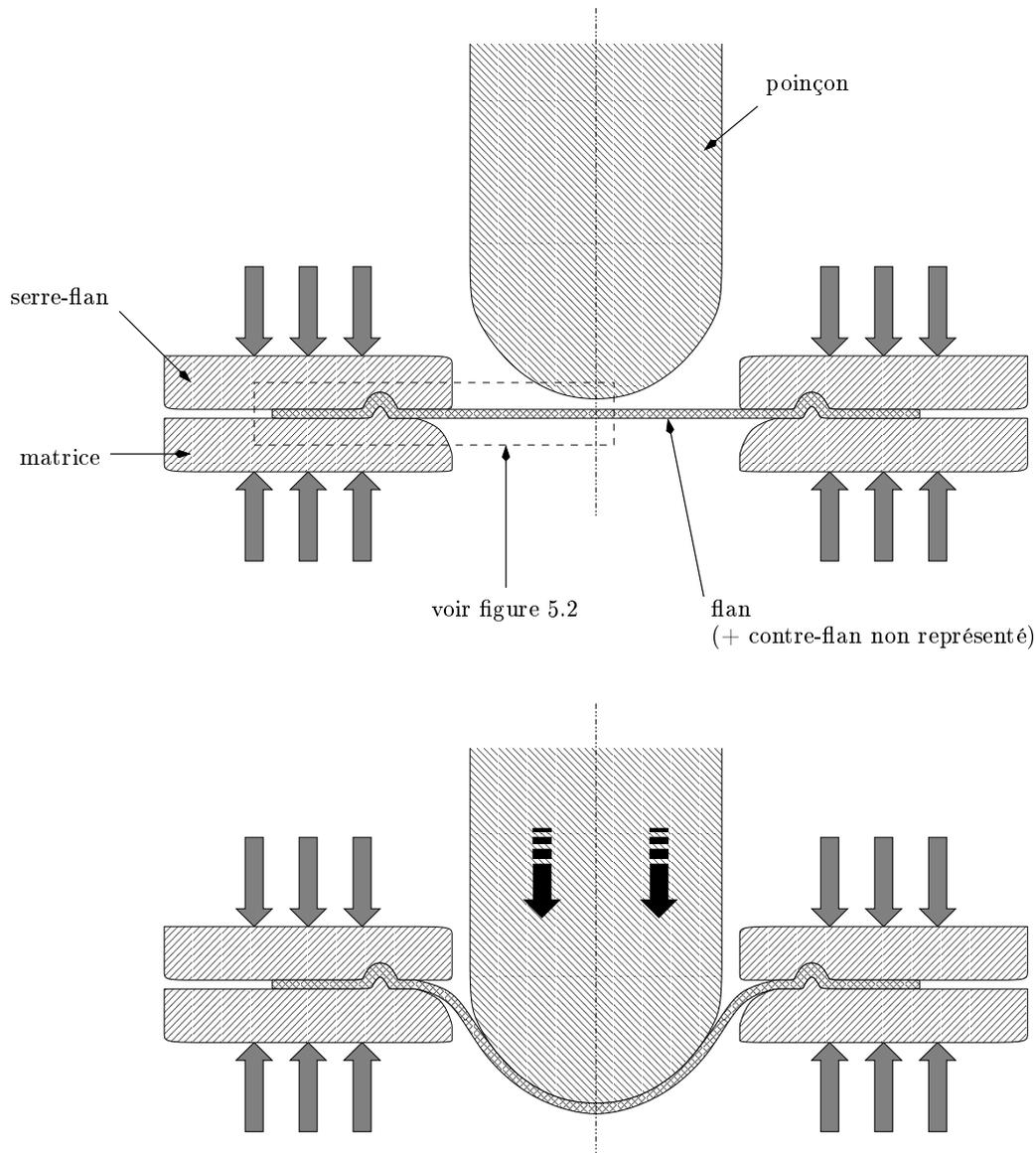


FIG. 5.1 : Schéma de principe de l'essai de coupelle à fond hémisphérique.

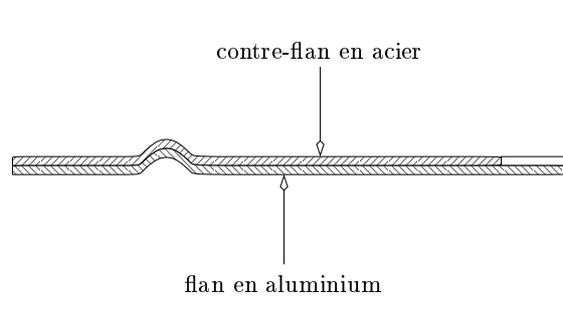
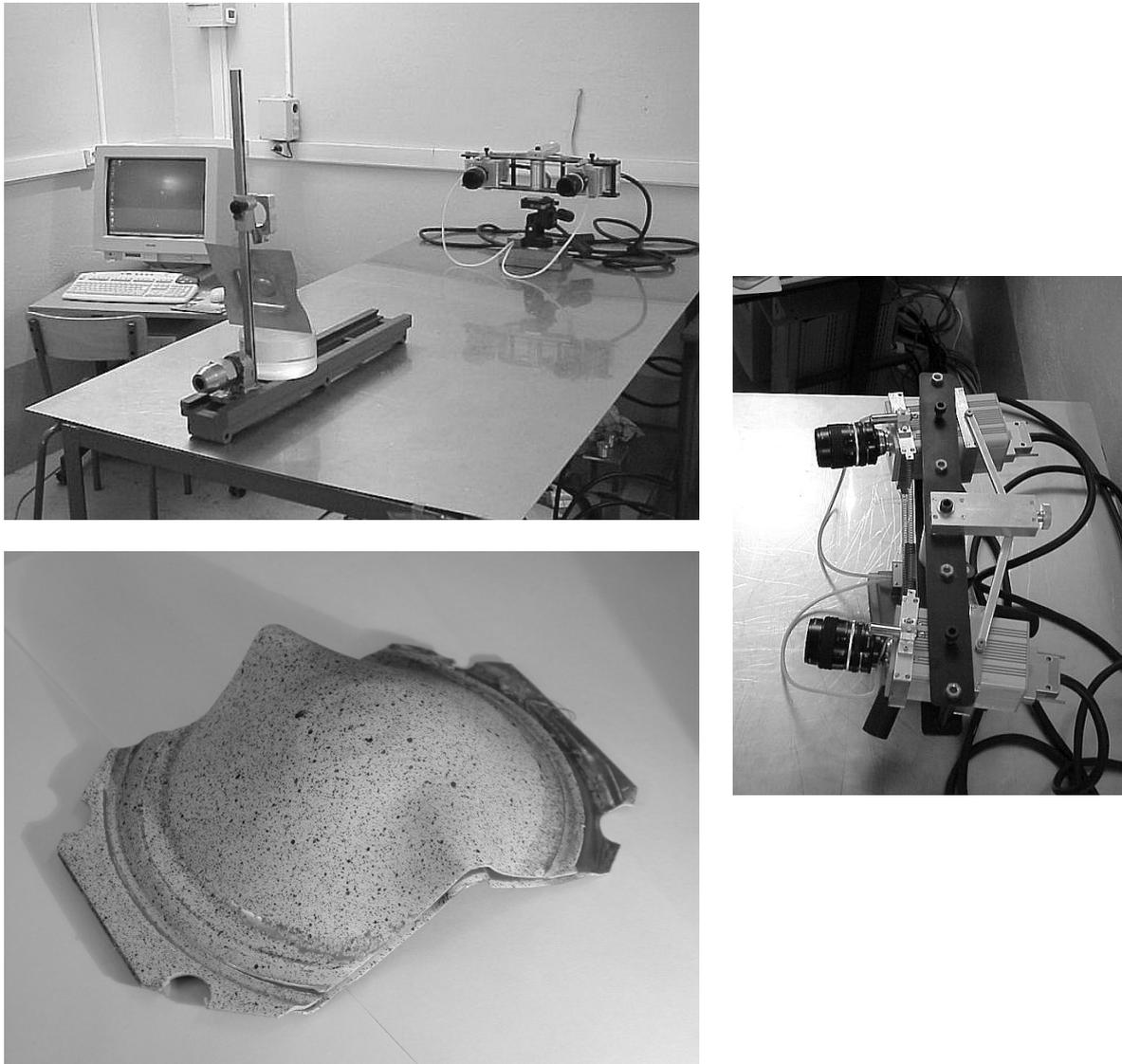
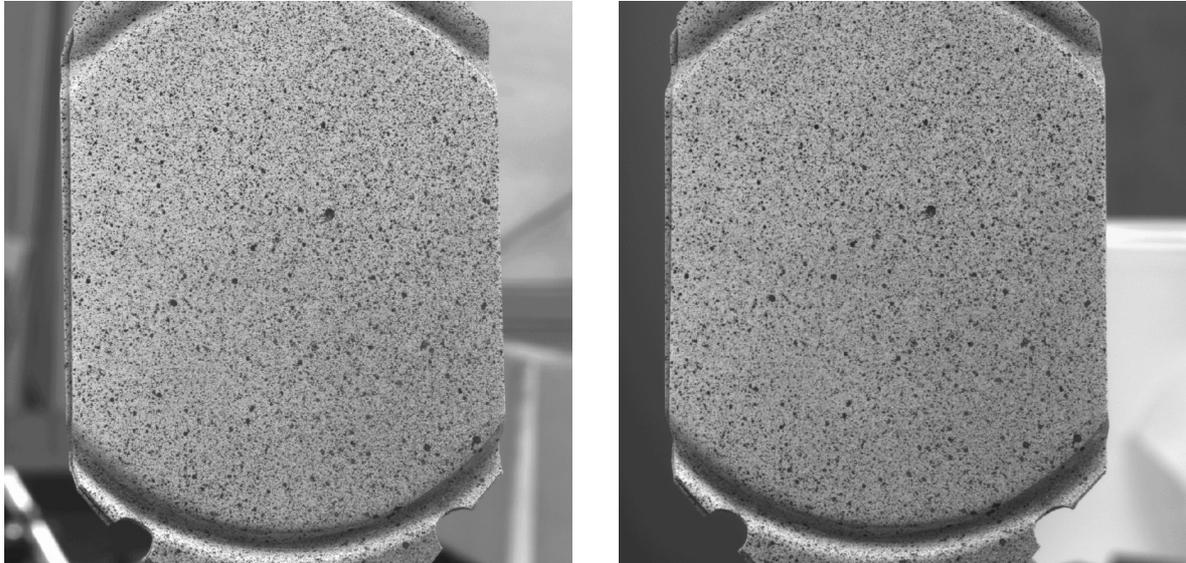


FIG. 5.2 : Vue en coupe du flan en aluminium et du contre-flan en acier percé en son centre. Le contre-flan permet de localiser la striction dans la zone percée lors de l'emboutissage.

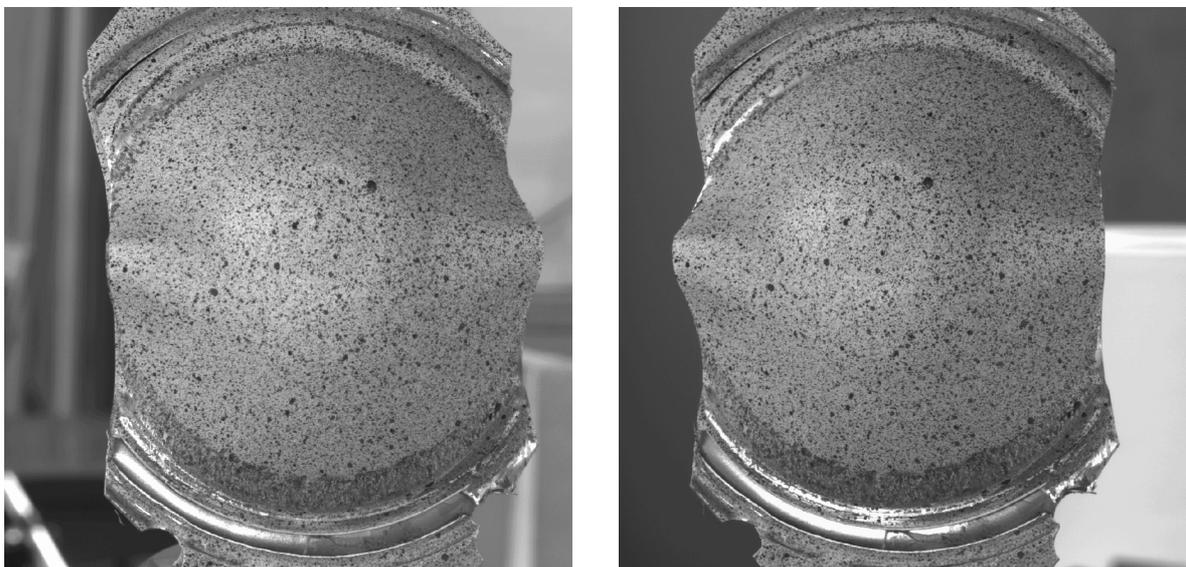


**FIG. 5.3 :** Configuration de l'expérience de reconstruction tridimensionnelle d'une tôle mince emboutie. La photo en haut à gauche montre le capteur de stéréovision de l'INSA de Lyon observant un objet placé à environ un mètre, ainsi que le poste d'acquisition des paires d'images. L'objet à numériser est montré en bas à gauche. À droite est montré un gros plan du capteur de stéréovision.

Une paire d'images stéréoscopiques du flan avant emboutissage est acquise (*cf.* figure 5.4), la pièce est emboutie et une deuxième paire d'images stéréoscopiques de la pièce emboutie est acquise (*cf.* figure 5.5). En appliquant la technique de stéréo-corrélation fine améliorée du



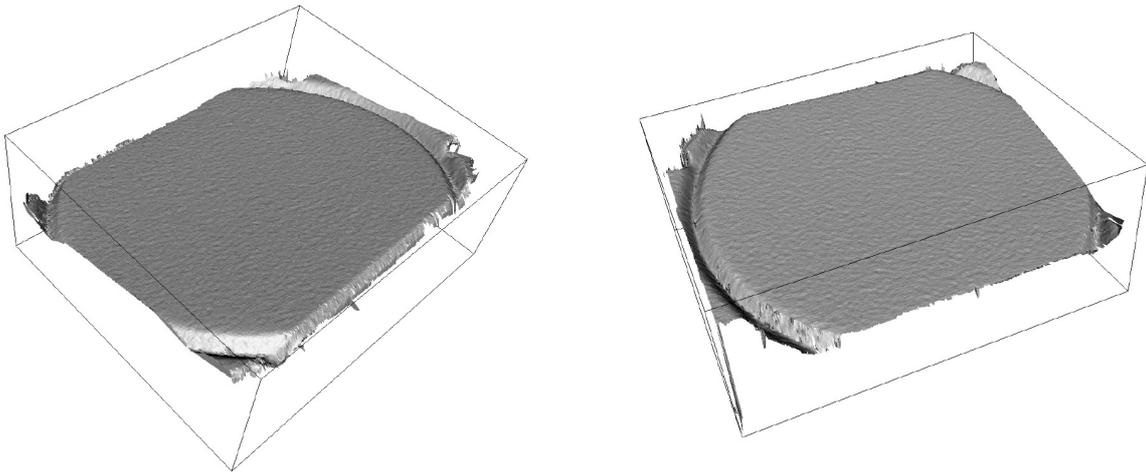
**FIG. 5.4 :** *Paire d'images stéréoscopique du flan avant emboutissage.*



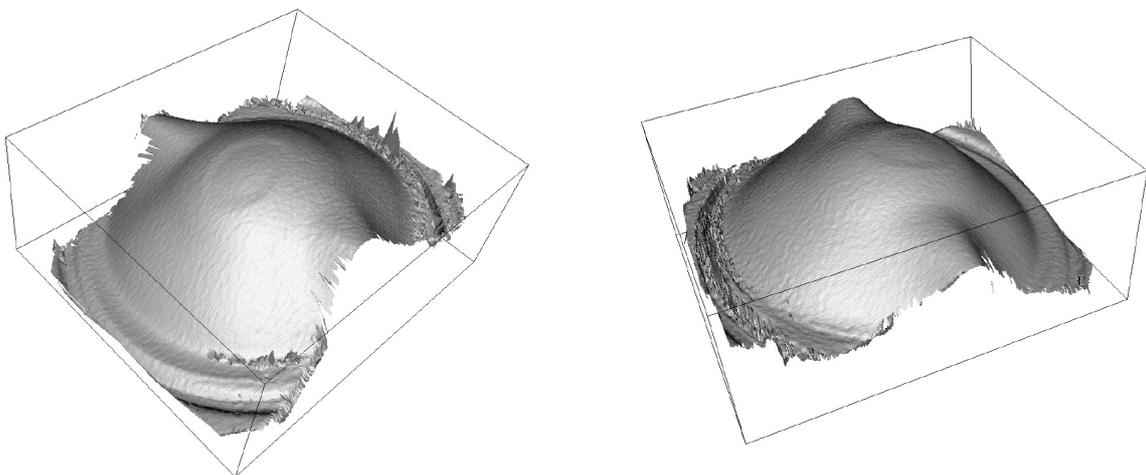
**FIG. 5.5 :** *Paire d'images stéréoscopique de la pièce emboutie.*

chapitre 3, sur ces deux paires d'images stéréoscopiques, on obtient la forme tridimensionnelle de la pièce avant et après emboutissage (*cf.* figures 5.6 et 5.7). Cette expérience correspond à l'emboutissage d'une tôle en aluminium en utilisant un contre-flan en acier percé en son centre, et un poinçon hémisphérique. On peut constater sur la reconstruction tridimensionnelle de la figure 5.7 que l'embouti étudié présente des défauts de formage car la forme obtenue n'est pas hémisphérique. Nous avons choisi cet embouti comme exemple car il correspond à un cas

particulièrement critique pour la stéréo-corrélation : la surface présente de fortes pentes qui conduisent à de fortes distorsions perspectives des domaines de corrélation dans les images. On peut aussi remarquer sur cette figure une trace au sommet de l'embouti qui correspond au marquage du contre-flan, ce qui donne une certaine indication sur la précision de numérisation tridimensionnelle de la méthode.



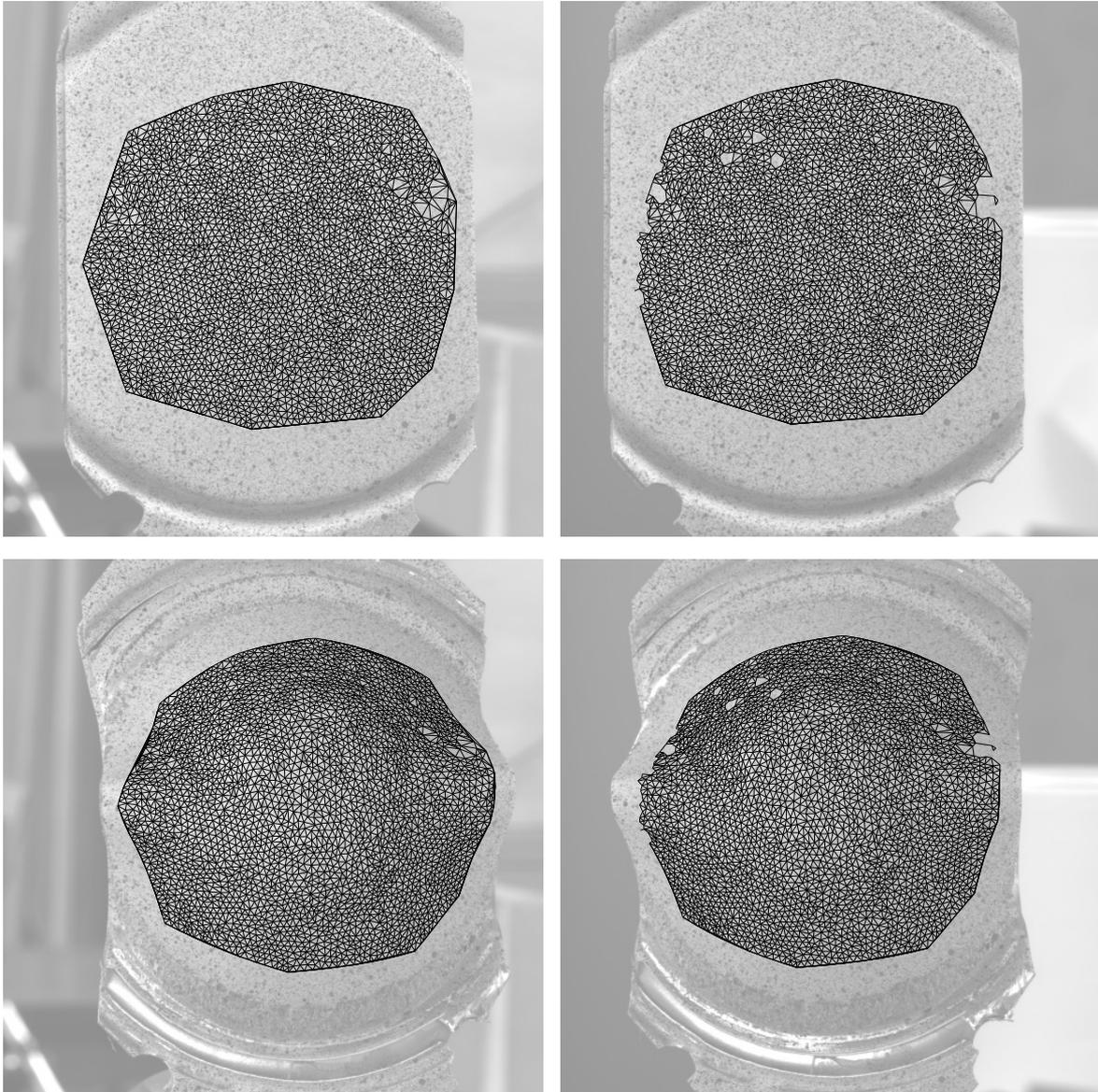
**FIG. 5.6 :** Deux vues de la forme tridimensionnelle du flan avant emboutissage calculée à partir de la paire d'images de la figure 5.4.



**FIG. 5.7 :** Deux vues de la forme tridimensionnelle de la pièce emboutie calculée à partir de la paire d'images de la figure 5.5.

Nous avons ensuite calculé les déformations en surface de cet embouti en utilisant la méthode que nous proposons dans le chapitre 4. La figure 5.8 montre le maillage triangulaire dont les nœuds ont été appariés dans les deux paires d'images stéréoscopiques. Ce maillage comporte 3000 points et 6000 triangles dont les côtés ont une longueur d'environ 10 pixels. La figure 5.9 montre les valeurs des déformations principales minimales et maximales calculées. On peut y observer une déformation maximale d'environ 25%, localisée au sommet de l'embouti,

ce qui était prévisible grâce de l'utilisation d'un contre-flan percé en son centre.



**FIG. 5.8 :** *Suivi des nœuds d'un maillage triangulaire par stéréo-corrélation et suivi de pixels simultanés. La maillage a été initialement défini dans l'image gauche de la paire stéréoscopique initiale (images du haut). Ce maillage comporte 3000 nœuds, 6000 triangles d'environ 10 pixels de côté.*

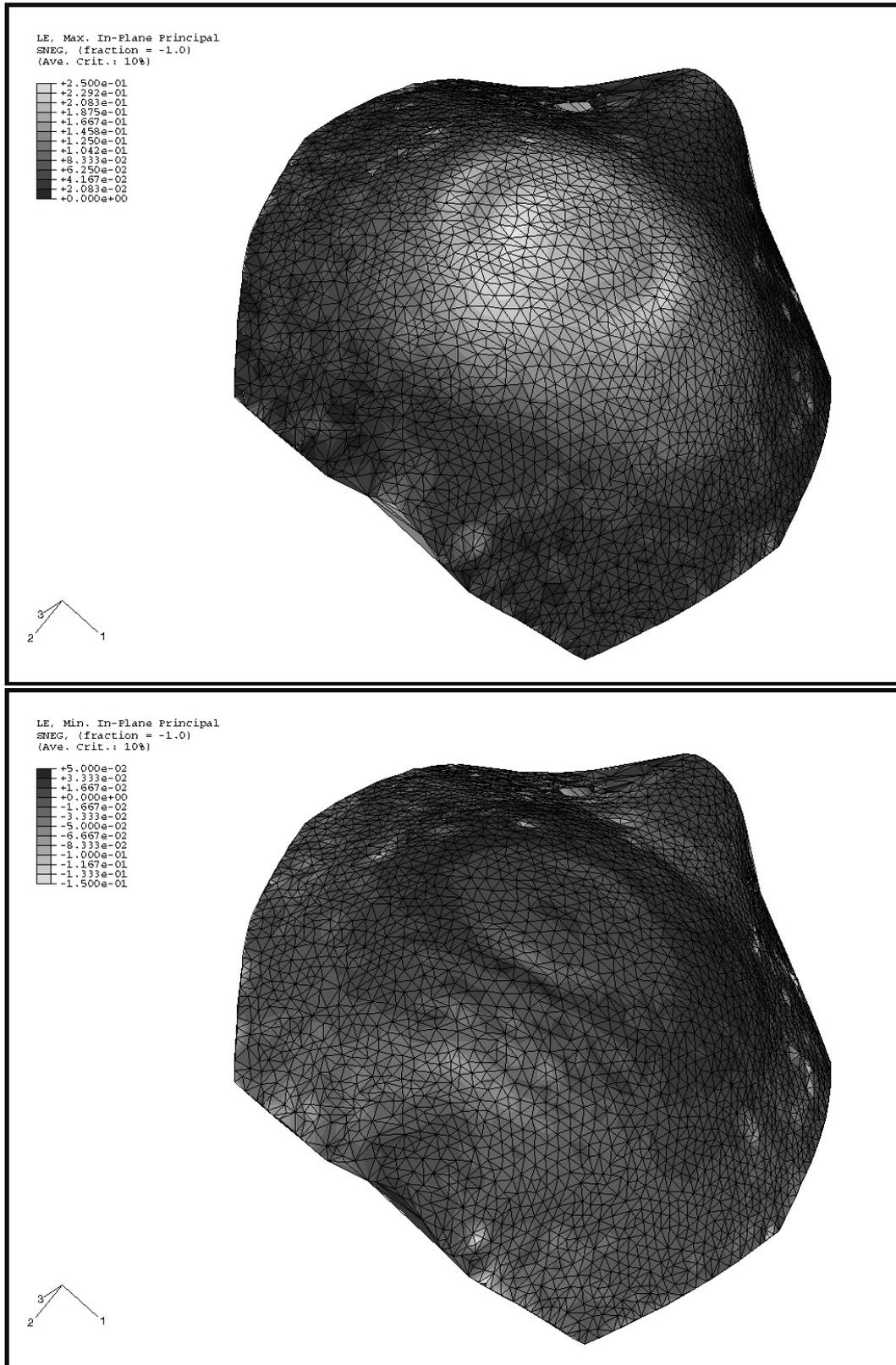
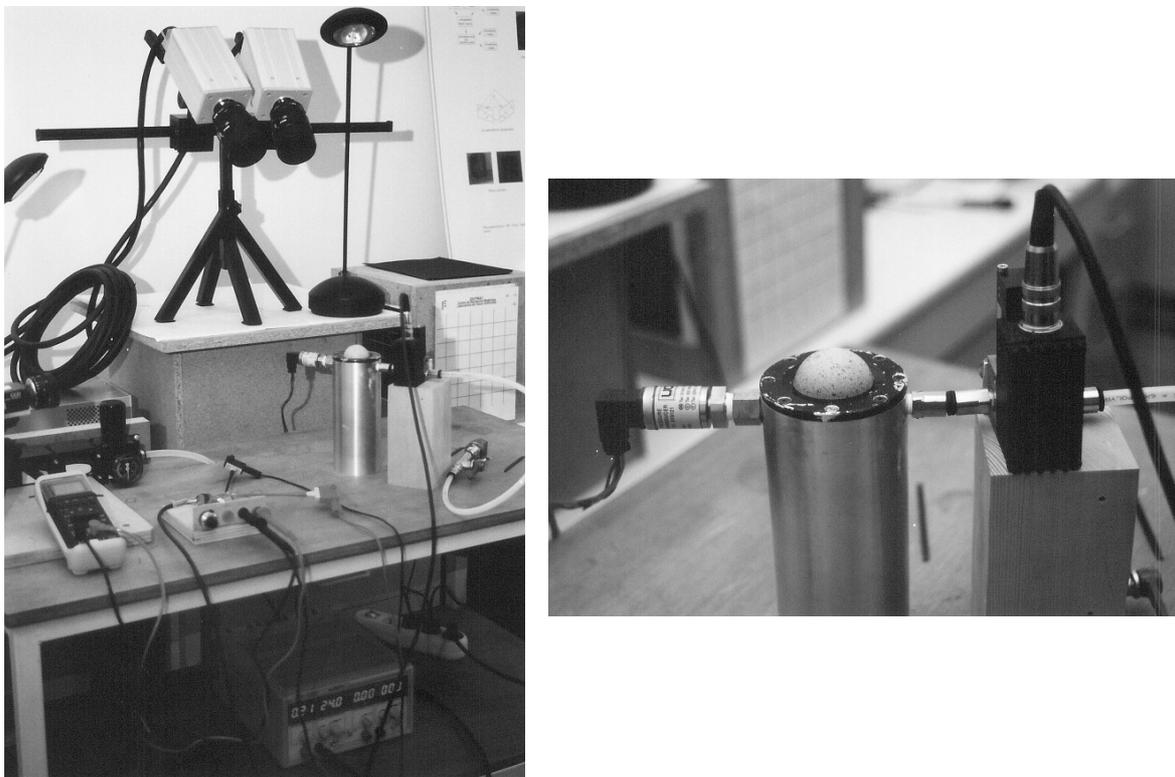


FIG. 5.9 : Champs des déformations principales maximales (en haut) et minimales (en bas) de la tôle mince emboutie.

## 5.2 Membranes en élastomère soufflées

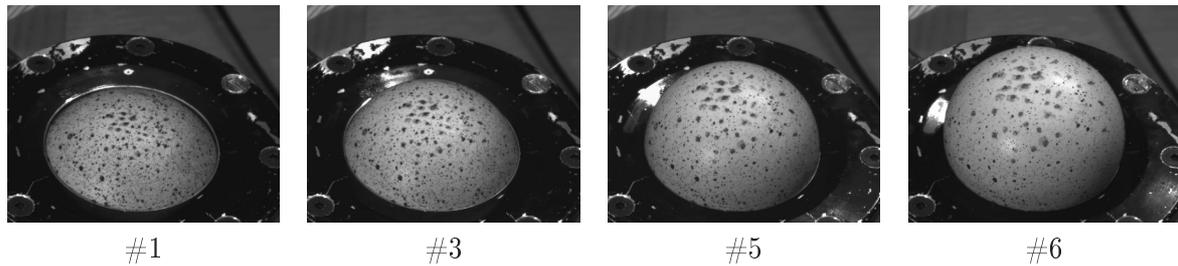
Dans le domaine de la mise en forme des polymères, il est fondamental de pouvoir accéder au calcul des déformations pour une meilleure compréhension du comportement des polymères durant la phase de mise en forme. SCHMIDT *et al.* [SRLM<sup>+</sup>99, RSLM<sup>+</sup>01] ont développé à l'École des Mines d'Albi un montage expérimental pour étudier le comportement rhéologique d'un élastomère. Le principe de ce rhéomètre consiste à imposer une pression régulée sous un flan (disque en élastomère) pour le déformer. Dans leurs travaux, le soufflage de la membrane est enregistré par une seule caméra ce qui permet de déterminer les propriétés rhéologiques biaxiales de l'élastomère. Afin d'accéder au champ de déplacements tridimensionnels de la membrane soufflée, nous avons proposé d'utiliser notre méthode de stéréovision (*cf.* figure 5.10 qui représente le dispositif de bullage et le capteur de stéréovision en position de mesure).



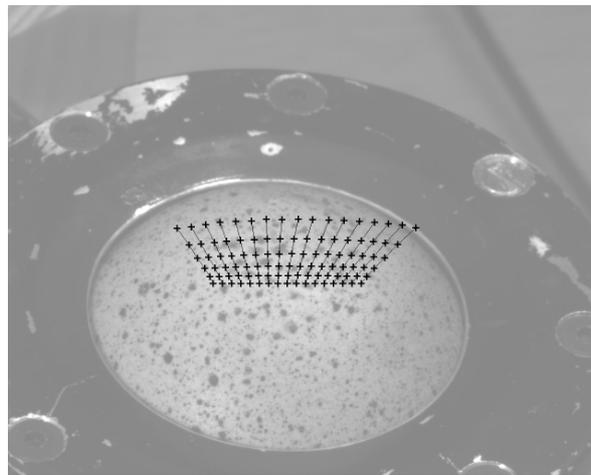
**FIG. 5.10 :** *Dispositif de bullage et capteur de stéréovision en position de mesure (image gauche). La membrane en élastomère est maintenue par un serre-flan et soumise à une pression de gonflage régulée (l'image droite montre une vue détaillée de la membrane gonflée).*

Les figures 5.11 à 5.14 illustrent une expérience de mesure d'un champ de déplacements tridimensionnels d'une membrane en élastomère en train de gonfler, ainsi que la forme tridimensionnelle de la membrane à différents instants d'acquisition. La surface de la membrane a été préalablement texturée par un dépôt de mouchetis de peinture.

Il est à noter que dans le cas de l'emboutissage les déformations à mesurer sont de l'ordre de quelques % (on parle de petites déformations), alors que dans le cas des polymères, les déformations peuvent atteindre plusieurs centaines de %. Dans ce dernier cas, on enregistre



**FIG. 5.11** : Quatre images extraites d'une séquence de six images acquises par le capteur de stéréovision (on ne représente ici que les images acquises par la caméra gauche).



**FIG. 5.12** : Suivi d'une ligne de points à travers six images de la séquence. Le résultat est affiché en surimpression de la première image.

une séquence d'images avec une fréquence garantissant que la déformation entre chaque prise de vue reste dans la limite des déformations mesurables par corrélation. Cette limite dépend de nombreux paramètres (taux de déformation locale de la surface, son orientation par rapport aux deux plans images du capteur, *etc.*). Toutefois, nous avons pu constater que l'on arrive à apparier des points sans difficulté pour des taux de déformation allant jusqu'à 30% environ.

La figure 5.11 montre quatre images (numéros 1, 3, 5 et 6) extraites d'une séquence de six images acquises par la caméra gauche du capteur de vision stéréoscopique.

Nous avons appliqué la méthode de stéréo-corrélation et suivi de pixel simultanés discutée au chapitre 4 pour suivre dans le temps un ensemble de points répartis sur une grille virtuelle. Sur la figure 5.12, on montre le chemin suivi par chacun des points de la première ligne de la grille. Sur la figure 5.13, on montre la grille initiale sur la première image de la séquence et la grille obtenue par cette méthode dans la dernière (sixième image). Par triangulation des appariements stéréoscopiques obtenus, on peut calculer la reconstruction tridimensionnelle correspondant à chacune des vues de la membrane en élastomère, puis le champ de déplacements tridimensionnels (cf. figure 5.14).

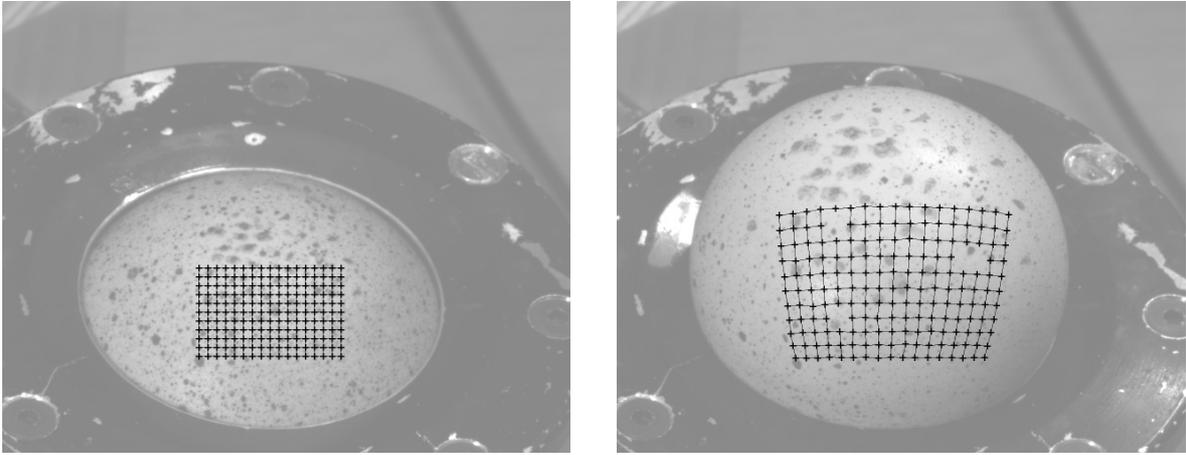


FIG. 5.13 : Appariement temporel d'une grille virtuelle entre la première et la dernière image de la séquence.

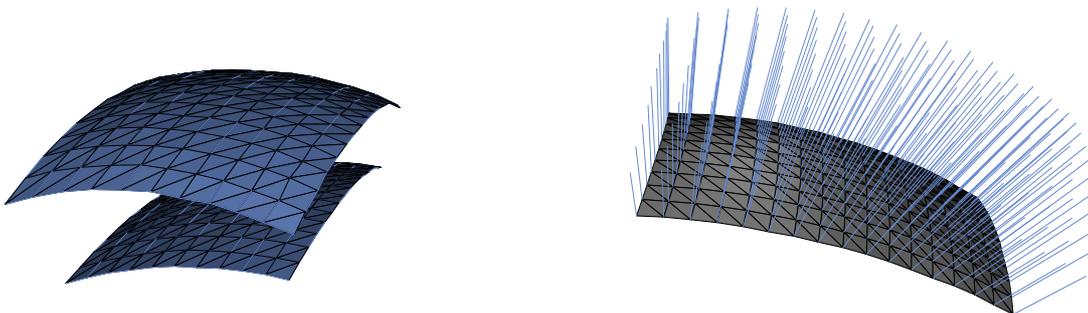
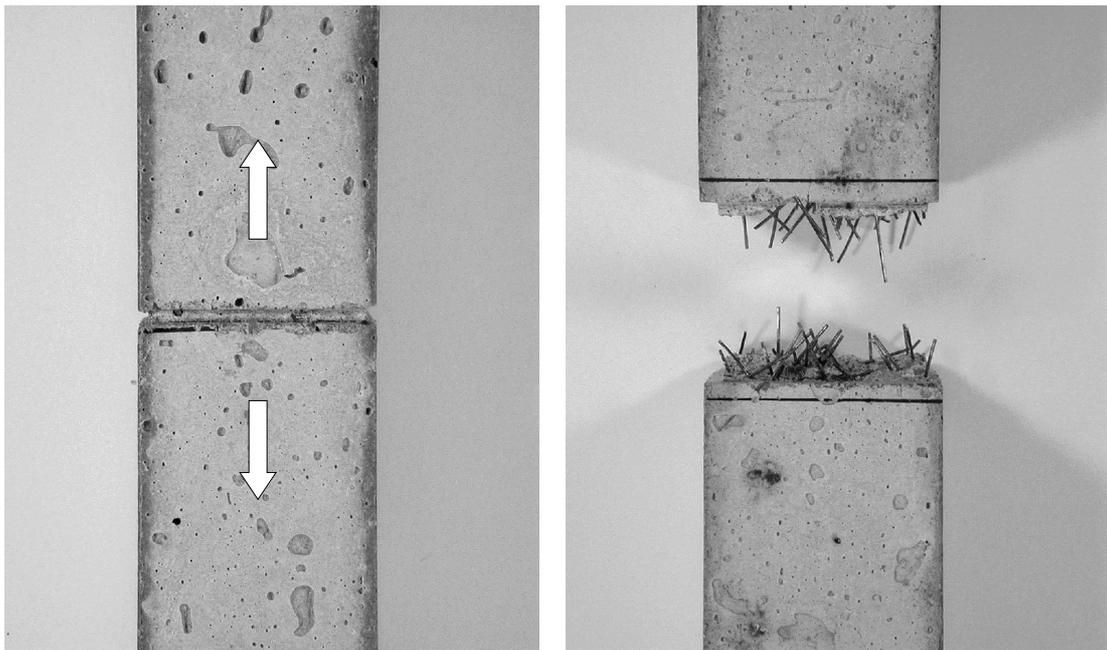


FIG. 5.14 : Reconstruction tridimensionnelle de la membrane en élastomère (instants 1 et 4) et champ de déplacements tridimensionnels correspondant.

### 5.3 Bétons réfractaires renforcés de fibres métalliques

L'introduction de fibres métalliques dans les bétons réfractaires a pour objectif de diminuer la fragilité qui caractérise leur comportement mécanique. Les essais de traction sur éprouvettes pré-entaillées (*cf.* figure 5.15) permettent de caractériser et de quantifier le gain obtenu sur les performances mécaniques d'un matériau ainsi renforcé [CCLB99, Cai01]. En cours d'essai, le suivi permanent de champs de déplacements et/ou de déformations à proximité de l'entaille et celui de l'effort appliqué permettent de tracer les courbes charge/déplacement et/ou charge/déformation.



**FIG. 5.15 :** *Éprouvette en béton renforcée de fibres métalliques avant et après rupture suite à un essai de traction uniaxial.*

Nous avons réalisé un essai de traction sur une éprouvette parallélépipédique pré-entaillée (*cf.* figure 5.15) équipée de deux extensomètres mécaniques placés sur deux faces opposées de l'éprouvette (*cf.* figures 5.16 et 5.17). Nous avons acquis une séquence de paires d'images stéréoscopiques pendant l'essai, en mémorisant pour chaque paire la charge appliquée et les valeurs des déplacements mesurés par les extensomètres mécaniques. Par stéréo-corrélation, nous avons mesuré le déplacement tridimensionnel d'un certain nombre de points (une centaine de couples) répartis de part et d'autre de la fissure (extensomètres virtuels) [GDO<sup>+</sup>01]. La figure 5.16 montre les couples de points dans la première image gauche et la figure 5.17 les montre dans la dernière image gauche de la séquence.

Sur la figure 5.18, nous représentons le déplacement mesuré par chacun des extensomètres virtuels à différents instants d'acquisition des images. Sur cette figure, il faut distinguer l'ensemble des points représentatifs de la face avant de l'éprouvette et ceux de la face latérale. Concernant la face avant, la zone de coïncidence avec les tiges de l'extensomètre mécanique est située entre les couples 10 et 20. La figure 5.19 montre la bonne correspondance entre les déplacements mesurés par l'extensomètre mécanique et ceux calculés à partir de la moyenne arithmétique des valeurs de déplacements fournies par les extensomètres virtuels situés dans

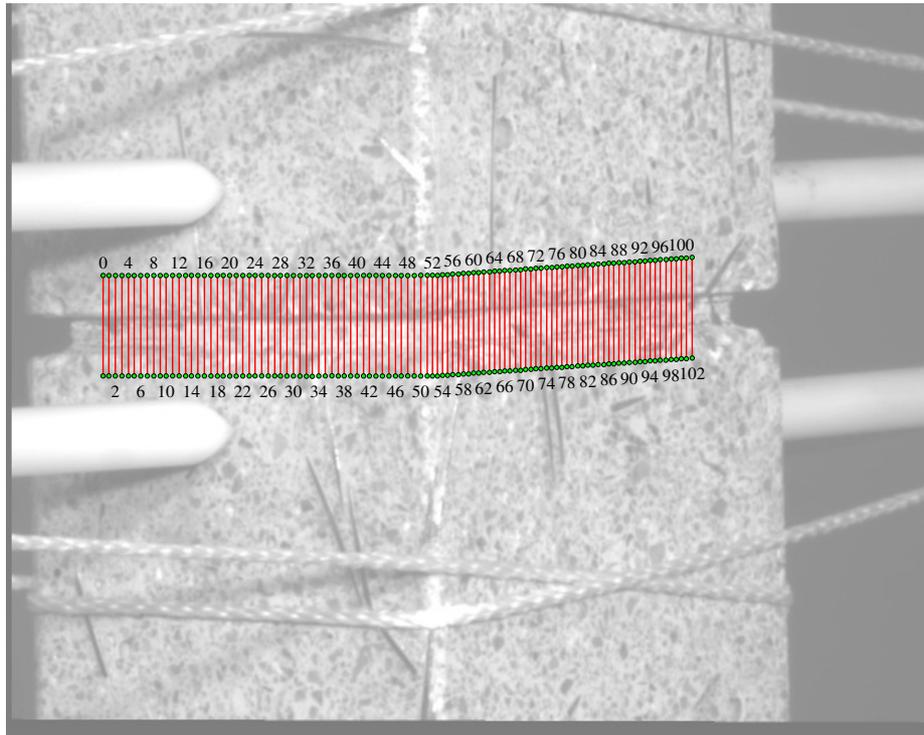


FIG. 5.16 : Extensomètres virtuels et extensomètres mécaniques (tiges blanches), en début d'essai.

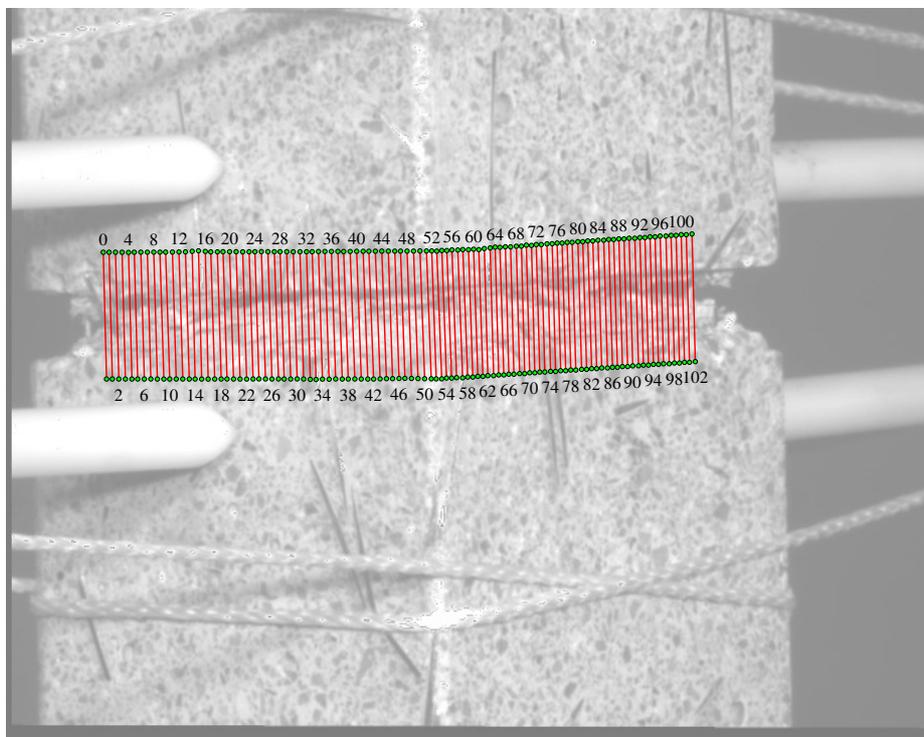


FIG. 5.17 : Extensomètres virtuels et extensomètres mécaniques (tiges blanches), en fin d'essai.

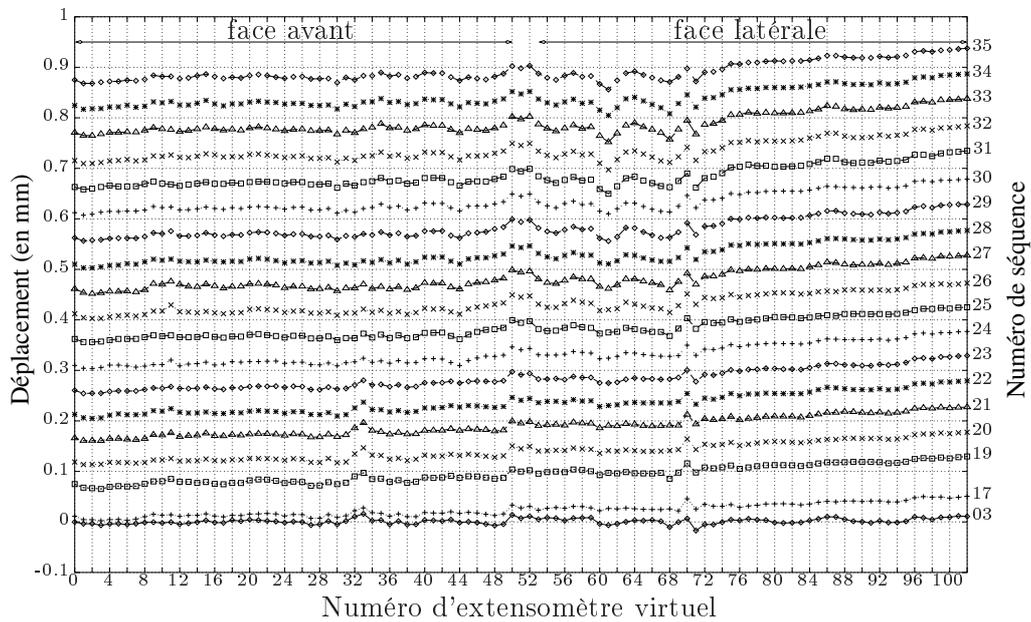


FIG. 5.18 : Déplacements mesurés par chacun des extensomètres virtuels à différents instants d'acquisition des images.

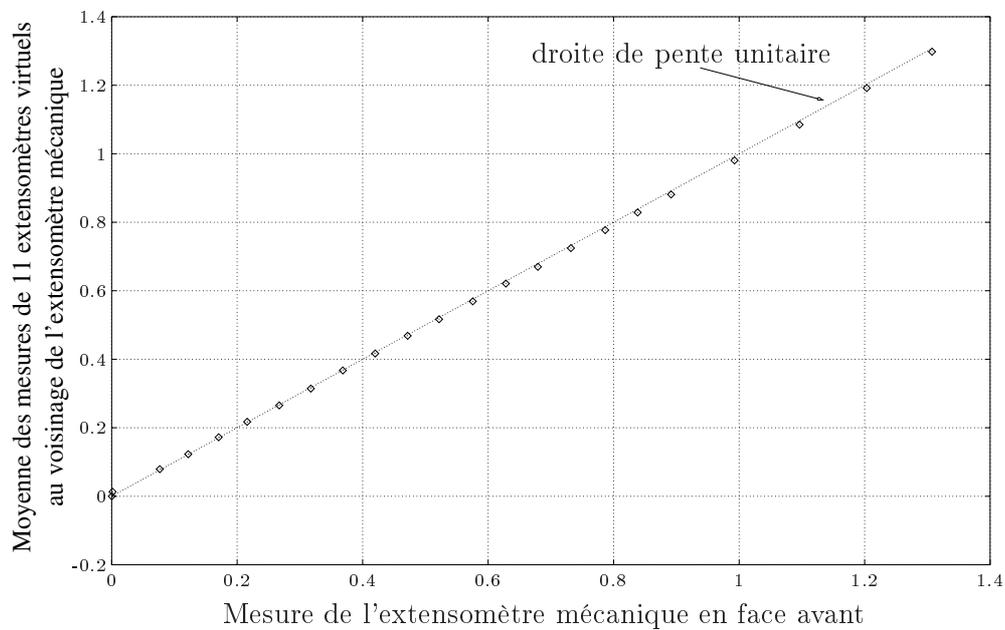


FIG. 5.19 : Comparaison des déplacements mesurés par l'extensomètre mécanique et par stéréovision.

cette zone.

Les courbes de la figure 5.18 montrent des valeurs accidentées aux alentours des extensomètres numérotés 50 à 70 : nous avons vérifié que cela correspondait à une zone de l'éprouvette où une partie du béton était en train de s'arracher. Le grand nombre d'extensomètres virtuels permet d'estimer si les deux côtés de l'entaille réalisée dans l'éprouvette s'ouvrent de manière parallèle. C'est effectivement le cas sur la face avant puisque pour un déplacement imposé donné, l'ensemble des valeurs des extensomètres virtuels peut être représenté par une droite horizontale. Il n'en est pas de même pour la face latérale pour laquelle les extensomètres virtuels indiquent une tendance à l'augmentation des valeurs de déplacement lorsque les points considérés se rapprochent de la face arrière de l'éprouvette. Une fois encore, une très bonne correspondance est obtenue entre les valeurs données par les extensomètres mécaniques et les extensomètres virtuels. Par exemple, pour la courbe numéro 29, les déplacements mesurés par les extensomètres mécaniques placés en face avant et en face arrière sont respectivement de 0,575 et de 0,642. Les valeurs données par les extensomètres virtuels varient entre 0,57 et 0,62 (le dernier extensomètre virtuel n'est pas vraiment en face arrière de l'éprouvette). Ces résultats indiquent que le mode de sollicitation mécanique en traction de l'éprouvette n'introduit pas de flexion parasite dans les plans parallèles à la face avant, ce qui n'est pas le cas dans les plans parallèles à la face latérale. La mesure par stéréovision permet ainsi de quantifier cette part de flexion dans chacun de ces plans sans avoir recours à quatre extensomètres mécaniques.

## Bibliographie du chapitre

- [Cai01] Emmanuel CAILLEUX. *Microstructure et comportement thermomécanique d'un béton réfractaire renforcé par des fibres métalliques*. Thèse de doctorat, École des Mines de Paris (France), mai 2001.
- [CCLB99] T. CUTARD, E. CAILLEUX, P. LOURS, et G. BERNHART. Structural and Mechanical Properties of a Refractory Concrete for Superplastic Forming Tools. *Industrial Ceramics*, 19(2):100–102, 1999.
- [GDO<sup>+</sup>01] D. GARCIA, B. DAVID, J.-J. ORTEU, T. CUTARD, et E. CAILLEUX. Application de la stéréovision à l'étude du comportement de bétons renforcés de fibres métalliques. Dans *Photomécanique'2001*, pp. 24–26, Futuroscope, Poitiers (France), avril 2001.
- [GH99] K. GALANULIS et A. HOFMANN. Determination of Forming Limit Diagrams Using an Optical Measurement System. Dans *7th International Conference on Sheet Metal (SheMet'99)*, pp. 245–252, Erlangen (Allemagne), septembre 1999.
- [MBB95] F. MORESTIN, P. BOGAERT, et M. BOIVIN. Mesure de déformations par imagerie : détermination des courbes limites de formage. Dans *Photomécanique'95*, ENS Cachan (France), mars 1995.
- [RSLM<sup>+</sup>01] N. REUGE, F. M. SCHMIDT, Y. LE MAOULT, M. RACHIK, et F. ABBÉ. Elastomer Biaxial Characterization Using Bubble Inflation Technique. I : Experimental Investigations. *Polymer Engineering and Science*, 41(3):522, mars 2001.
- [SRLM<sup>+</sup>99] F. SCHMIDT, N. REUGE, Y. LE MAOULT, J.P. ARCENS, B. DAVID, et F. ABBÉ. Inflation technique optimisation for an elastomer biaxial characterisation. Dans *Polymer Processing Society - 15th annual meeting*, 's-Hertogenbosch (Pays-Bas), juin 1999.
- [VL89] J. H. VOGEL et D. LEE. An Automated Two-View Method For Determining Strain Distributions on Deformed Surfaces. Dans *Journal of Material Shaping Technology*, volume 6, 1989.



# Conclusion générale et perspectives

L'objectif de cette thèse était de développer un système permettant de mesurer, sans contact et de manière précise, à la fois la forme tridimensionnelle d'un objet statique ou le champ de déplacements/déformations subi par un objet sous sollicitation mécanique.

Pour la mesure de formes tridimensionnelles, nous avons mis en œuvre une technique de stéréo-corrélation couplant la stéréovision et l'appariement de points par corrélation. La précision de la reconstruction tridimensionnelle obtenue dépend à la fois de la qualité du calibrage du capteur de stéréovision, et de la précision d'appariement des paires d'images stéréoscopiques.

Habituellement le calibrage du capteur de stéréovision est obtenu à partir du calibrage indépendant de chacune des deux caméras. En vue d'améliorer la qualité du calibrage du capteur de stéréovision, nous avons proposé une méthode originale de type photogrammétrique qui considère ce capteur comme un tout indivisible et qui détermine de façon globale l'ensemble des paramètres de ce capteur. Nous avons montré l'intérêt d'une telle approche globale en la comparant à la méthode de calibrage classique par une simulation et un calibrage réel.

En ce qui concerne l'appariement stéréoscopique des images par corrélation, nous avons proposé une méthode basée sur des domaines de corrélation déformables selon un modèle de transformation stéréoscopique locale. Cette technique, appelée stéréo-corrélation fine améliorée, calcule les scores de corrélation directement dans les images acquises par le capteur et non pas dans celles dégradées par les opérations de rectification et de correction de la distorsion, comme c'est le cas pour la méthode classique. Une expérience de reconstruction tridimensionnelle d'un objet selon la méthode classique et la nôtre a permis de montrer l'amélioration que cela apporte.

Pour la mesure de champs de déplacements/déformations tridimensionnels, nous avons mis en œuvre une technique couplant la stéréo-corrélation et l'appariement temporel d'images par corrélation. La technique développée pour l'appariement temporel d'images utilise des domaines de corrélation déformables et un modèle de luminance des images par surface B-spline. Plutôt que d'adopter la démarche classique qui consiste à traiter de façon séquentielle deux paires d'images acquises à deux états de déformation, nous avons proposé de considérer de façon globale les quatre images afin de déterminer de façon optimale les appariements stéréoscopiques et temporels. Comme pour la mesure de formes tridimensionnelles, les techniques mises en œuvre reposent sur les images brutes (non dégradées par les opérations de rectification et de correction de la distorsion).

À travers les applications que nous avons mises en œuvre (emboutissage de tôles minces, soufflage de membranes en élastomère, essais de traction sur éprouvettes en béton renforcées de fibres métalliques), nous avons montré l'intérêt de ce type de technique de mesure dans le cadre de la mécanique expérimentale. Il est à noter que peu de méthodes de mesure de déformations permettent à la fois d'obtenir une information tridimensionnelle (déformations dans le plan et hors plan) et de mesurer des taux de déformations pouvant aller de quelques centièmes de % à plusieurs centaines de %.

Il reste du travail à faire en terme de validation des performances métrologiques du système (incertitudes sur les mesures de formes, de déplacements et de déformations). En ce qui concerne les mesures de formes, notre laboratoire participe à une intercomparaison de systèmes de mesures tridimensionnelles (lasers de poursuite, vidéogrammétrie), dans le cadre du groupe Mesure Tridimensionnelles par Procédés Optiques (MTPO) du Collège de Métrologie du Mouvement Français pour la Qualité (MFQ), qui aura lieu courant 2002 (elle était initialement prévue en 2001). Cette intercomparaison devrait nous permettre d'évaluer plus finement les performances métrologiques de notre système.

En général, on s'intéresse à la mesure de champs de déformations de matériaux sollicités thermo-mécaniquement. Dans le cadre de cette thèse, nous n'avons considéré que des mesures de champs de déformations à température ambiante (sollicitations purement mécaniques). Dans le cadre de sollicitations thermiques, il conviendrait d'étudier les problèmes que l'on pourrait rencontrer en utilisant la stéréo-corrélation pour les mesures à chaud.

# Annexes



## Annexe A

# Du modèle optique d'une caméra au modèle géométrique

Nous considérerons tout d'abord que le système optique est composé d'une lentille idéale, i.e. une lentille mince, utilisée avec des angles d'incidence de rayons et des diamètres de faisceaux suffisamment faibles (approximation de GAUSS) pour que les défauts optiques, ou aberrations géométriques, puissent être négligés. Nous aborderons ensuite le cas d'un système optique composé d'une lentille épaisse. Puis nous verrons qu'en négligeant les effets de flou, ces différents modèles optiques peuvent se ramener à un modèle géométrique simple de projection perspective : le modèle sténopé. Cette section est basée sur la présentation de Patrick GROS dans [Gro93] et sur l'ouvrage [Rog96].

### A.1 Modèle lentille mince

Une lentille mince est une lentille, ou une associations de lentilles, pour laquelle on considère son épaisseur négligeable par rapport à sa focale ou à toute autre de ses dimensions. L'approximation de GAUSS, ou **approximation paraxiale**, considère que les rayons incidents à la lentille sont peu inclinés avec son axe pour pouvoir faire l'approximation au premier ordre suivante :

$$\sin \theta \approx \tan \theta \approx \theta$$

Cette approximation permet le tracé rapide de certains rayons grâce aux règles suivantes :

- tout rayon passant par le centre de la lentille n'est pas dévié ;
- deux rayons parallèles entre eux se rejoignent dans le plan focal de la lentille.

La lentille mince est placée dans le plan  $Cxy$  et centrée au point  $C$  : son **centre optique** (cf. figure A.1). L'axe  $z$  est appelé **axe optique**. Le détecteur est confondu avec un plan désigné  $\mathcal{P}_i$  orthogonal à l'axe optique, et distant du centre optique  $C$  d'une longueur  $\gamma$  (cf. figure A.2). L'axe optique intersecte  $\mathcal{P}_i$  en un point que l'on notera  $\Omega$  et que l'on nomme **point principal**. Le point  $F$  est le **foyer objet**,  $F'$  est le **foyer image**, et la distance  $f = \overline{CF} = -\overline{CF'}$  est appelée **distance focale**. Nous allons chercher à déterminer l'image  $M''$  du point objet  $M$  sur le plan  $\mathcal{P}_i$  du détecteur.

Le point  $M'$ , appelé point image de  $M$ , est l'intersection de trois rayons lumineux :

- le rayon  $(MC)$  qui n'est pas dévié ;
- le rayon parallèle à l'axe  $Cz$  qui est dévié par la lentille en un rayon passant par le foyer image  $F'$  ;

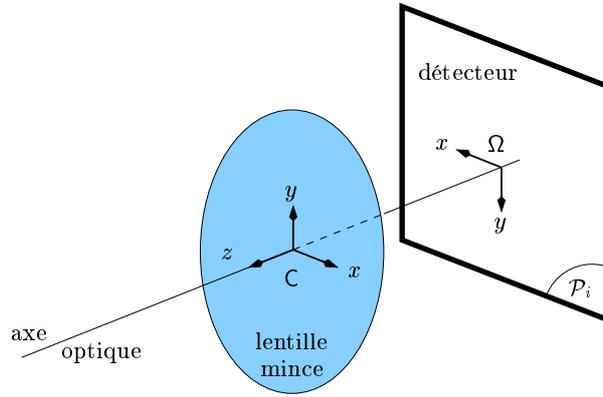


FIG. A.1 : Le détecteur et la lentille mince avec leur repère associé.

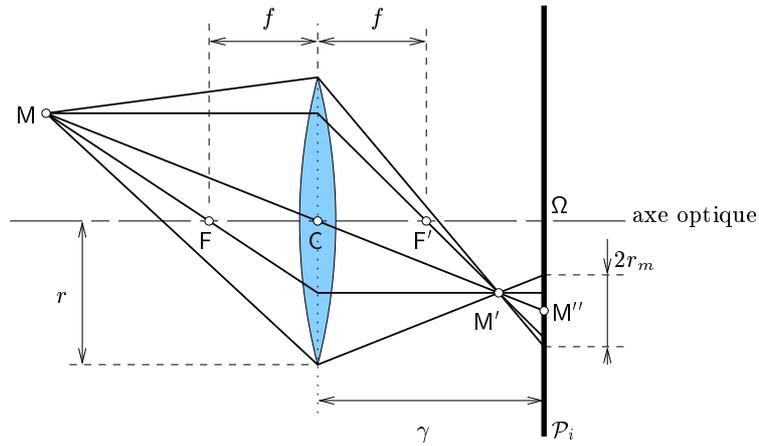


FIG. A.2 : Vue en coupe d'un système optique mince.

- le rayon passant par le foyer objet  $F$  qui est dévié par la lentille en un rayon parallèle à l'axe  $Cz$ .

En utilisant les distances algébriques  $\overline{CM}$  et  $\overline{CM'}$ , l'équation de GAUSS de la lentille est :

$$\frac{1}{\overline{CM}} - \frac{1}{\overline{CM'}} = \frac{1}{f}$$

Nous en déduisons les deux relations suivantes qui lient les points  $M$  et  $M'$  :

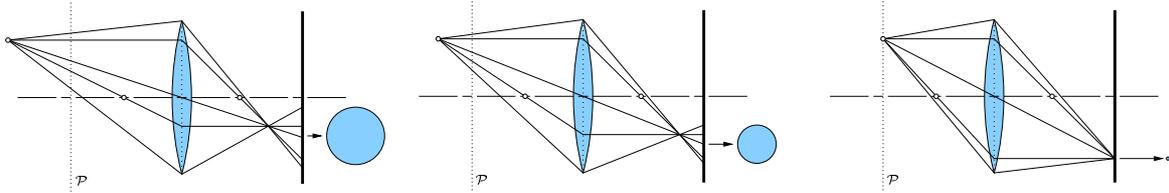
$$M = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto M' = \begin{pmatrix} \frac{fx}{f-z} \\ \frac{fy}{f-z} \\ \frac{fz}{f-z} \end{pmatrix} \quad \text{et} \quad M = \begin{pmatrix} \frac{fx'}{f+z'} \\ \frac{fy'}{f+z'} \\ \frac{fz'}{f+z'} \end{pmatrix} \mapsto M' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (\text{A.1})$$

Pour que l'image  $M''$  du point  $M$  soit nette sur le détecteur, il faut que  $M' \in \mathcal{P}_i$ . Cette condition est remplie pour tout point objet  $M$  dont l'image  $M'$  à pour coordonnée  $z' = -\gamma$ , c'est à dire

pour tout point M de coordonnée  $z$  égale à :

$$z' = -\gamma = \frac{fz}{f-z} \quad \Rightarrow \quad z = \frac{\gamma f}{\gamma - f}$$

Cela décrit un plan  $\mathcal{P}$  orthogonal à l'axe optique et distant de C d'une longueur  $\frac{\gamma f}{\gamma - f}$ . Ce plan est appelé **plan de netteté** ou **plan de mise au point**. Pour tout point objet M n'appartenant pas au plan de netteté, son image M' n'appartiendra pas au plan du détecteur  $\mathcal{P}_i$ . (cf. figure A.3). Nous allons calculer les coordonnées du point M'' sur le détecteur.



**FIG. A.3 :** L'image sur le plan du détecteur d'un point objet est d'autant moins floue que celui-ci est proche du plan de netteté  $\mathcal{P}$ .

Nous supposons que la lentille est circulaire de rayon  $r$ . Nous allons considérer tous les rayons d'origine le point  $M = (x \ y \ z)^\top$  et traversant la lentille. Nous notons  $(l \cos \theta \ l \sin \theta \ 0)^\top$  ces points d'intersection, avec  $\theta \in [0; 2\pi[$  et  $l \in [0; r]$ . D'après la figure A.2, il est aisé de déterminer la relation suivante :

$$\frac{x' - l \cos \theta}{z'} = \frac{x' - x''}{\gamma + z'}$$

En utilisant cette équation et en substituant  $z'$  et  $x'$  à partir de (A.1) nous déterminons le point  $M''_{l,\theta}$  en fonction du point objet M, des paramètres  $\theta$  et  $l$ , de la distance focale  $f$  et de  $\gamma$  :

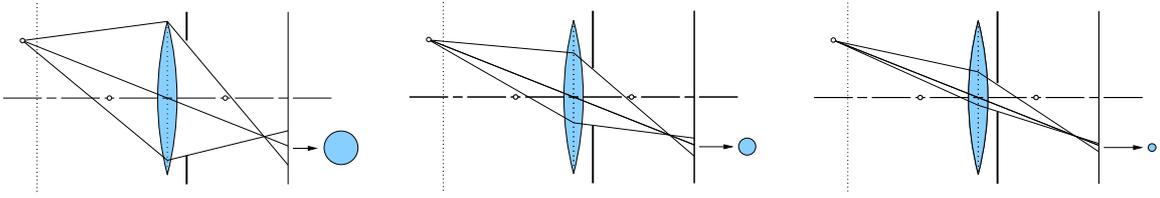
$$M''_{l,\theta} = \begin{pmatrix} l \cos \theta \left( 1 + \gamma \frac{f-z}{fz} \right) - \gamma \frac{x}{z} \\ l \sin \theta \left( 1 + \gamma \frac{f-z}{fz} \right) - \gamma \frac{y}{z} \\ -\gamma \end{pmatrix}$$

Si on fait varier les paramètres  $l$  et  $\theta$  alors l'image du point objet M sur le plan du détecteur est une tache circulaire de centre  $M''$  et de rayon  $r_m$ , avec :

$$M'' = \begin{pmatrix} -\gamma \frac{x}{z} \\ -\gamma \frac{y}{z} \\ -\gamma \end{pmatrix} \quad \text{et} \quad r_m = r \left( 1 + \gamma \frac{f-z}{fz} \right) \quad (\text{A.2})$$

Cette dernière équation explique le phénomène de flou survenant lorsque un point objet n'appartient pas au plan de mise au point : son image sur le détecteur n'est pas ponctuelle mais un cercle de rayon  $r_m$  non nul. Sans changer les cotes du système optique, le rayon de ce cercle peut être réduit en introduisant un **diaphragme d'ouverture** placé devant et/ou derrière la lentille (cf. figure A.4).

Si l'on se place dans le plan du détecteur, alors le centre de l'image du point M a pour projeté un point bidimensionnel que nous noterons  $m$  et dont les coordonnées homogènes se



**FIG. A.4 :** *L'adjonction d'un diaphragme d'ouverture permet de contrôler la dimension de la tache lumineuse projetée sur le plan du détecteur lorsque le point objet n'est pas dans le plan de netteté.*

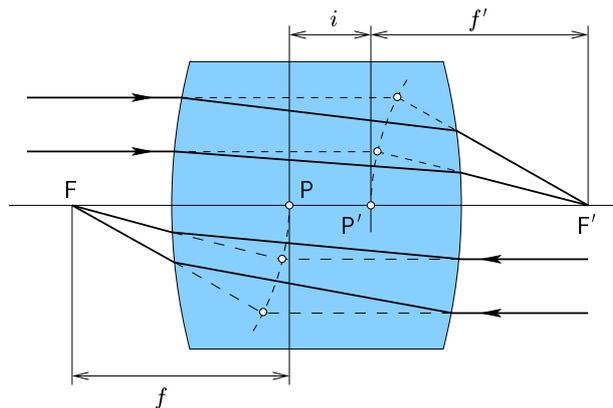
déduisent de l'équation (A.2)<sup>1</sup> :

$$\tilde{\mathbf{m}} \cong \begin{pmatrix} \gamma \frac{x}{z} \\ \gamma \frac{y}{z} \\ 1 \end{pmatrix} \cong \begin{pmatrix} \gamma & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Si l'on néglige les effets de flou, on peut donc décrire la transformation d'un système optique mince associé à un détecteur par une simple transformation projective. Le lecteur trouvera dans [Fau93] une introduction à la géométrie projective.

## A.2 Modèle lentille épaisse

Pour pouvoir être réalisées, les lentilles doivent être suffisamment rigides et donc épaisses. Généralement, l'épaisseur est faible par rapport à la focale et le modèle lentille mince est une bonne approximation. Il arrive cependant que cette approximation ne soit plus suffisante et l'on doit alors revoir les équations du modèle lentille mince pour tenir compte de cette épaisseur.



**FIG. A.5 :** *Plans principaux d'une lentille épaisse.*

La figure A.5 montre le cheminement des rayons dans une lentille épaisse. On montre que les formules des lentilles minces sont encore valables pour peu que l'on connaisse les **points**

<sup>1</sup>Le changement de signe provient de notre convention d'orientation des repères montrés figure A.1.

**principaux** P objet et P' image. On définit deux plans orthogonaux à l'axe optique et passant par ces deux points : les **plans principaux**. Ces derniers sont distants d'une longueur  $i$  nommée **interstice** de la lentille. Si l'on considère un rayon parallèle à l'axe optique entrant dans la lentille du coté gauche de la figure, et si l'on prolonge le rayon réfracté jusqu'à son intersection avec le rayon incident, le plan perpendiculaire à l'axe contenant P' est appelé **plan principal image**. Une lentille mince placée en cet endroit donnerait, de ce rayon incident, le même rayon réfracté. La distance focale image  $f'$  d'une lentille épaisse est la distance du foyer image F' au plan principal image.

Si l'on considère à présent un rayon parallèle à l'axe optique arrivant sur la lentille du coté droit de la figure, la distance focale objet  $f$  obtenue est égale à la distance du foyer objet F au **plan principal objet**, i.e. égale à  $f'$  si la lentille est dans l'air.

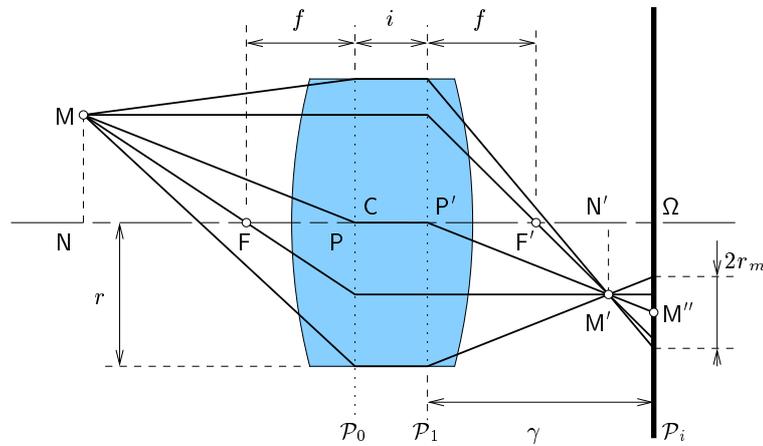


FIG. A.6 : Vue en coupe d'un système optique épais.

La figure A.6 montre un système optique épais avec un détecteur. À partir de cette figure nous pouvons écrire :

$$\frac{1}{\overline{PM}} - \frac{1}{\overline{P'M'}} = \frac{1}{f} \quad \text{et} \quad \frac{\overline{N'M'}}{\overline{NM}} = \frac{\overline{P'N'}}{\overline{PN}}$$

Nous en déduisons les deux relations suivantes :

$$M = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto M' = \begin{pmatrix} \frac{fx}{f-z} \\ \frac{fy}{f-z} \\ \frac{fz}{f-z} - i \end{pmatrix} \quad \text{et} \quad M = \begin{pmatrix} \frac{fx'}{f+i+z'} \\ \frac{fy'}{f+i+z'} \\ \frac{f(z'+i)}{f+i+z'} \end{pmatrix} \mapsto M' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

En supposant la lentille circulaire de rayon  $r$ , un rayon optique d'origine un point  $M = (x \ y \ z)^\top$  et passant par la lentille aura pour image, sur le détecteur, un cercle de rayon  $r_m$  et de centre le point  $M''$ , avec :

$$M'' = \begin{pmatrix} -\gamma \frac{x}{z} \\ -\gamma \frac{y}{z} \\ -\gamma - i \end{pmatrix} \quad \text{et} \quad r_m = r \left( 1 + \gamma \frac{f-z}{fz} \right)$$

Nous notons  $\mathbf{m}$  les coordonnées de ce point exprimées dans le plan du détecteur. Les coordonnées homogènes de  $\mathbf{m}$  se formulent directement à partir de l'équation précédente en respectant

les conventions d'orientation des repères montrés figure A.1 :

$$\tilde{\mathbf{m}} \cong \begin{pmatrix} \gamma \frac{x}{z} \\ \gamma \frac{y}{z} \\ 1 \end{pmatrix} \cong \begin{pmatrix} \gamma & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Nous pouvons remarquer que, comme pour un système optique mince, si on ignore l'effet de flou nous pouvons décrire la transformation d'un système optique épais associé à un détecteur par une simple transformation projective.

### A.3 Modèle géométrique

Nous venons de montrer que, sous les conditions de GAUSS, le modèle lentille mince et, plus généralement, le modèle lentille épaisse pouvaient être décrits par une transformation projective pour peu que l'on ne tienne pas compte de l'effet de flou.

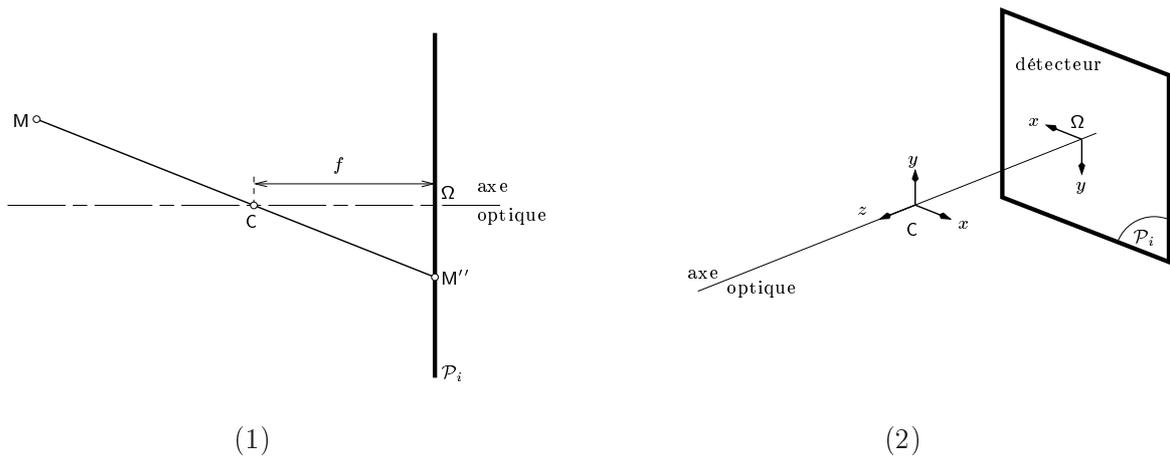


FIG. A.7 : Modèle projection perspective du couple système optique et détecteur.

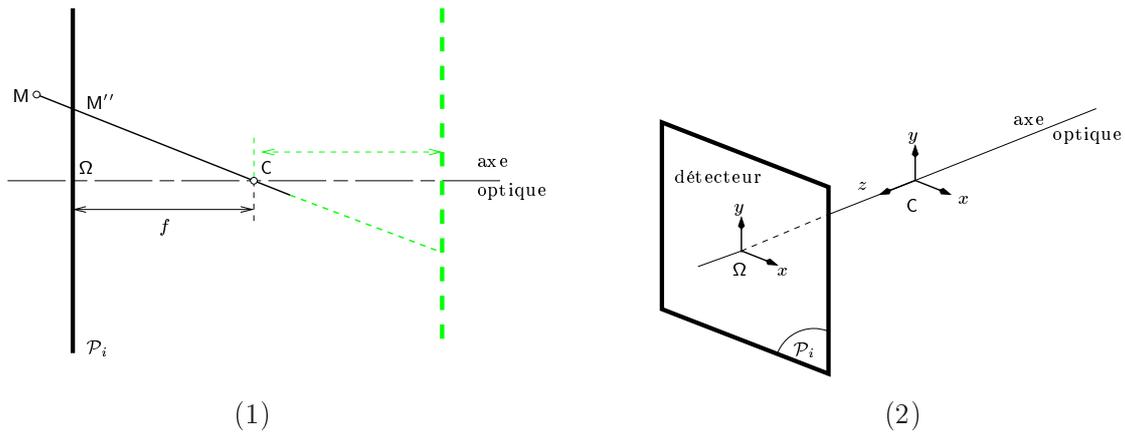
Les figures A.7-1 et A.7-2 montrent un modèle géométrique équivalent, appelé **modèle de projection perspective**. La distance  $f$  est appelée **focale** du système. Il ne faut pas confondre la focale d'une projection perspective, distance du centre optique au plan du détecteur, avec la distance focale des modèles optiques qui ne dépend pas de la position du détecteur. Il est aisé de déterminer les coordonnées du projeté  $M''$  d'un point objet  $M$  :

$$M = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto M'' = \begin{pmatrix} -f \frac{x}{z} \\ -f \frac{y}{z} \\ -f \end{pmatrix}$$

Nous notons  $\mathbf{m}$  les coordonnées de ce point exprimées dans le plan du détecteur. Les coordonnées homogènes de  $\mathbf{m}$  s'expriment directement à partir de l'équation précédente (en respectant

les orientations des repères indiqués figure A.7-2) :

$$\tilde{\mathbf{m}} \cong \begin{pmatrix} f \frac{x}{z} \\ f \frac{y}{z} \\ 1 \end{pmatrix} \cong \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (\text{A.3})$$



**FIG. A.8 :** *Modèle projection perspective équivalent : le plan image est le symétrique par rapport au centre optique C du plan image de la figure A.7.*

Il existe une autre configuration géométrique pour laquelle les coordonnées de ce point projeté sont les mêmes. Elle consiste à faire une symétrie par rapport au centre optique C du plan image  $\mathcal{P}_i$  (appelé aussi **plan rétinien**) de la figure A.7 (cf. figure A.8). Les coordonnées du projeté  $\mathbf{m}$  du point objet M sont alors déterminées via l'équation (A.3). Ce modèle géométrique est le plus couramment utilisé car l'image d'une scène formée sur le plan  $\mathcal{P}_i$  n'est pas « inversée » par rapport au modèle de la figure A.7. C'est celui que nous avons retenu pour établir le modèle de caméra permettant de déterminer les coordonnées dans l'image d'un point 3D exprimé dans un référentiel propre à la scène observée.



## Annexe B

# Calibrage d'une caméra : méthode linéaire de RAVN *et al.*

En 1993, RAVN, ANDERSEN et SORENSEN ont proposé une méthode analytique d'identification des paramètres intrinsèques et extrinsèques du modèle sténopé de caméra en utilisant plusieurs vues d'un objet plan de calibrage [RAS93]. Nous présentons en détails cette méthode après correction de quelques erreurs présentes dans leur papier, et après adaptation des équations à nos conventions de repère.

Chaque point de calibrage donne deux équations indépendantes en utilisant (1.6). Sans perte de généralité, nous pouvons faire l'hypothèse que le plan de calibrage est dans le plan  $z = 0$ . Les équations (1.6) peuvent se mettre sous la forme :

$$\frac{\alpha_u (t_x + r_{11}x + r_{12}y) + u_0 (t_z + r_{31}x + r_{32}y)}{t_z + r_{31}x + r_{32}y} = u$$
$$\frac{\alpha_v (t_y + r_{21}x + r_{22}y) + v_0 (t_z + r_{31}x + r_{32}y)}{t_z + r_{31}x + r_{32}y} = v$$

Leur idée a été de regrouper tous les paramètres à estimer dans un seul et même vecteur. Pour cela, les équations précédentes sont reformulées en :

$$\frac{\alpha_u (t_x + r_{11}x + r_{12}y) + u_0 (t_z + r_{31}x + r_{32}y)}{t_z} = u \left( 1 + \frac{r_{31}x + r_{32}y}{t_z} \right)$$
$$\frac{\alpha_v (t_y + r_{21}x + r_{22}y) + v_0 (t_z + r_{31}x + r_{32}y)}{t_z} = v \left( 1 + \frac{r_{31}x + r_{32}y}{t_z} \right)$$

qui se développe en :

$$\frac{\alpha_u t_x + u_0 t_z}{t_z} + x \frac{\alpha_u r_{11} + u_0 r_{31}}{t_z} + y \frac{\alpha_u r_{12} + u_0 r_{32}}{t_z} - ux \frac{r_{31}}{t_z} - uy \frac{r_{32}}{t_z} = u$$
$$\frac{\alpha_v t_y + v_0 t_z}{t_z} + x \frac{\alpha_v r_{21} + v_0 r_{31}}{t_z} + y \frac{\alpha_v r_{22} + v_0 r_{32}}{t_z} - vx \frac{r_{31}}{t_z} - vy \frac{r_{32}}{t_z} = v$$

Cette écriture nous amène au système linéaire suivant :

$$\begin{pmatrix} 1 & 0 & x & y & 0 & 0 & -ux & -uy \\ 0 & 1 & 0 & 0 & x & y & -vx & -vy \end{pmatrix} (\kappa_1 \ \kappa_2 \ \kappa_3 \ \kappa_4 \ \kappa_5 \ \kappa_6 \ \kappa_7 \ \kappa_8)^\top = \begin{pmatrix} u \\ v \end{pmatrix} \quad (\text{B.1})$$

avec

$$\begin{aligned} \kappa_1 &= \frac{\alpha_u t_x + u_0 t_z}{t_z} & \kappa_3 &= \frac{\alpha_u r_{11} + u_0 r_{31}}{t_z} & \kappa_5 &= \frac{\alpha_v r_{21} + v_0 r_{31}}{t_z} & \kappa_7 &= \frac{r_{31}}{t_z} \\ \kappa_2 &= \frac{\alpha_v t_y + v_0 t_z}{t_z} & \kappa_4 &= \frac{\alpha_u r_{12} + u_0 r_{32}}{t_z} & \kappa_6 &= \frac{\alpha_v r_{22} + v_0 r_{32}}{t_z} & \kappa_8 &= \frac{r_{32}}{t_z} \end{aligned}$$

Avec au moins quatre points du plan de calibrage, l'équation (B.1) donne un système d'équations pour lequel on peut déterminer  $\kappa_1, \dots, \kappa_8$ . Dans le cas de plus de quatre points, on se ramène à la résolution d'un problème surdéterminé par une méthode classique (pseudo-inverse, décomposition SVD, ...).

Chaque position de l'objet plan de calibrage donne un jeu de paramètres  $\kappa_1, \dots, \kappa_8$ . En combinant les expressions des paramètres  $\kappa_3, \dots, \kappa_8$  et en utilisant la propriété d'orthonormalité de la matrice de rotation, RAVN *et al.* établissent les équations suivantes :

$$\begin{pmatrix} \kappa_3 \kappa_8 + \kappa_4 \kappa_7 & \kappa_5 \kappa_6 & \kappa_5 \kappa_8 + \kappa_6 \kappa_7 & \kappa_7 \kappa_8 \\ 2(\kappa_3 \kappa_7 - \kappa_4 \kappa_8) & \kappa_5^2 - \kappa_6^2 & 2(\kappa_5 \kappa_7 - \kappa_6 \kappa_8) & \kappa_7^2 - \kappa_8^2 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{pmatrix} = \begin{pmatrix} -\kappa_3 \kappa_4 \\ \kappa_4^2 - \kappa_3^2 \end{pmatrix} \quad (\text{B.2})$$

avec

$$\nu_1 = -u_0 \quad \nu_2 = \frac{\alpha_u^2}{\alpha_v^2} \quad \nu_3 = -\frac{\alpha_u^2}{\alpha_v^2} v_0 \quad \nu_4 = \frac{\alpha_u^2}{\alpha_v^2} v_0^2 + u_0^2 + \alpha_u^2$$

Avec deux vues, ou plus, de l'objet plan de calibrage, le système linéaire (B.2) peut être résolu. Ils en déduisent les paramètres intrinsèques du modèle sténopé de caméra :

$$u_0 = -\nu_1 \quad v_0 = -\frac{\nu_3}{\nu_2} \quad \alpha_u = \sqrt{\nu_4 - \frac{\nu_3^2}{\nu_2} - \nu_1^2} \quad \alpha_v = -\sqrt{\frac{\nu_4 - \frac{\nu_3^2}{\nu_2} - \nu_1^2}{\nu_2}}$$

Les paramètres intrinsèques déterminés, on peut calculer les paramètres extrinsèques de chaque position de la mire plane de calibrage. En utilisant la propriété d'orthonormalité de la matrice de rotation qui nous indique que  $r_{11}^2 + r_{21}^2 + r_{31}^2 = 1$  et que  $r_{12}^2 + r_{22}^2 + r_{32}^2 = 1$ , ils calculent  $t_z$  :

$$t_z = \sqrt{\frac{2}{\frac{(\kappa_3 - \kappa_7 u_0)^2}{\alpha_u^2} + \frac{(\kappa_5 - \kappa_7 v_0)^2}{\alpha_v^2} + \frac{(\kappa_4 - \kappa_8 u_0)^2}{\alpha_u^2} + \frac{(\kappa_6 - \kappa_8 v_0)^2}{\alpha_v^2} + \kappa_7^2 + \kappa_8^2}}$$

En utilisant  $t_z$  ainsi calculé dans les expressions de  $\kappa_1$  et  $\kappa_2$ , ils déterminent  $t_x$  et  $t_y$  :

$$t_x = \frac{\kappa_1 t_z - u_0 t_z}{\alpha_u} \quad t_y = \frac{\kappa_2 t_z - v_0 t_z}{\alpha_v}$$

La translation de la transformation rigide est à présent établie. Ils s'en servent pour déterminer les coefficients de la matrice de rotation :

$$\begin{aligned} r_{11} &= t_z \frac{\kappa_3 - \kappa_7 u_0}{\alpha_u} & r_{12} &= t_z \frac{\kappa_4 - \kappa_8 u_0}{\alpha_u} & r_{21} &= t_z \frac{\kappa_5 - \kappa_7 v_0}{\alpha_v} & r_{22} &= t_z \frac{\kappa_6 - \kappa_8 v_0}{\alpha_v} \\ & & r_{31} &= \kappa_7 t_z & r_{32} &= \kappa_8 t_z & & \end{aligned}$$

Le vecteur colonne  $\mathbf{r}_3$  est déterminé en utilisant la propriété d'orthogonalité de la matrice de rotation :

$$\mathbf{r}_3 = \mathbf{r}_1 \wedge \mathbf{r}_2$$

Nous avons rencontré des problèmes avec les matrices de rotation ainsi calculées car leur orthonormalité n'est pas garantie. Nous avons alors ajouté une étape supplémentaire à la méthode de RAVN *et al.* qui permet de déterminer les matrices orthonormales  $\bar{\mathbf{R}}$  les plus proches des matrices  $\mathbf{R}$  en minimisant la quantité  $\|\mathbf{R} - \bar{\mathbf{R}}\|$ , où  $\|\cdot\|$  est la norme de FROBENIUS. On montre que ce problème de minimisation admet pour solution  $\hat{\mathbf{R}} = \mathbf{U}\mathbf{V}^\top$ , où  $\mathbf{U}$  et  $\mathbf{V}$  sont des matrices orthogonales obtenues par décomposition en valeurs singulières (SVD) de la matrice  $\mathbf{R}$  telles que :  $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\top$  (voir [Zha98a] pour une démonstration).



## Annexe C

# Calibrage par méthode photogrammétrique : unicité des solutions

Lorsque l'on calibre une caméra en réestimant sans contrainte les points tridimensionnels de la mire, les paramètres extrinsèques calculés pour chacune des positions de la mire et ses points tridimensionnels sont définis à une homographie inconnue près (nous allons montrer que c'est une similitude). En effet, l'équation (1.1) peut aussi s'écrire :

$$\begin{aligned}\tilde{M}_c &\cong \mathbf{T} \tilde{M} \\ &\cong \underbrace{\mathbf{T}\mathbf{H}}_{\hat{\mathbf{T}}} \underbrace{\mathbf{H}^{-1}\tilde{M}}_{\tilde{Q}} \\ &\cong \hat{\mathbf{T}}\tilde{Q}\end{aligned}$$

Quels sont les paramètres de la famille des homographies  $\mathbf{H}$  compatibles ? Sans perte de généralité, posons :

$$\mathbf{H} \cong \begin{pmatrix} \mathbf{A}_{3 \times 3} & \mathbf{b}_3 \\ \mathbf{c}_3^\top & \lambda \end{pmatrix}$$

Exprimons la transformation rigide  $\hat{\mathbf{T}}$  en fonction de  $\mathbf{T}$  et  $\mathbf{H}$  :

$$\begin{aligned}\hat{\mathbf{T}} &\cong \mathbf{T}\mathbf{H} \\ \begin{pmatrix} \hat{\mathbf{R}} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & 1 \end{pmatrix} &\cong \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & \lambda \end{pmatrix} \\ &\cong \begin{pmatrix} \mathbf{R}\mathbf{A} + \mathbf{t}\mathbf{c}^\top & \mathbf{R}\mathbf{b} + \lambda\mathbf{t} \\ \mathbf{c}^\top & \lambda \end{pmatrix}\end{aligned}$$

De cette équation nous tirons  $\mathbf{c}^\top = \mathbf{0}^\top$ , et elle devient :

$$\begin{pmatrix} \hat{\mathbf{R}} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & 1 \end{pmatrix} \cong \begin{pmatrix} \mathbf{R}\mathbf{A} & \mathbf{R}\mathbf{b} + \lambda\mathbf{t} \\ \mathbf{0}^\top & \lambda \end{pmatrix}$$

De cette équation nous identifions terme à terme :

$$\hat{\mathbf{R}} \cong \mathbf{R}\mathbf{A} \tag{C.1}$$

$$\hat{\mathbf{t}} \cong \mathbf{R}\mathbf{b} + \lambda\mathbf{t} \tag{C.2}$$

Or  $\widehat{\mathbf{R}}$  est une matrice de rotation. Elle possède par conséquent les propriétés d'orthonormalité suivantes :

$$\begin{aligned}\widehat{\mathbf{R}}^\top \widehat{\mathbf{R}} &= \mathbf{I} \\ |\widehat{\mathbf{R}}| &= 1\end{aligned}$$

Ces propriétés et l'équation (C.1) nous donnent :

$$\begin{aligned}\widehat{\mathbf{R}}^\top \widehat{\mathbf{R}} &= \mathbf{A}^\top \underbrace{\mathbf{R}^\top \mathbf{R}}_{=\mathbf{I}} \mathbf{A} = \mathbf{I} \\ \Downarrow \\ \mathbf{A}^\top \mathbf{A} &= \mathbf{I}\end{aligned}$$

Cette relation nous indique que  $\mathbf{A}$  est aussi une matrice orthogonale. En utilisant la propriété suivante du déterminant d'une matrice régulière :  $|k\mathbf{M}_{n \times n}| = k^n |\mathbf{M}|$ , cherchons à établir le déterminant de  $\mathbf{A}$  :

$$\begin{aligned}\widehat{\mathbf{R}} &= \alpha \mathbf{R} \mathbf{A} \quad \forall \alpha \in \mathbb{R}^* \\ \Downarrow \\ |\widehat{\mathbf{R}}| = 1 &= |\alpha \mathbf{R} \mathbf{A}| \\ &= |\mathbf{R}| \cdot |\alpha \mathbf{A}| = \alpha^3 |\mathbf{A}| \\ \Downarrow \\ |\mathbf{A}| &= \frac{1}{\alpha^3}\end{aligned}$$

Une homographie de  $\mathbb{P}^3$  étant définie à un facteur non nul près, on peut toujours définir une matrice  $\mathbf{D} = \alpha^{-1} \mathbf{A}$  telle que son déterminant soit égal à 1. L'homographie  $\mathbf{H}$  peut alors s'écrire :

$$\mathbf{H} = \alpha \begin{pmatrix} \mathbf{D} & \mathbf{e} \\ \mathbf{0}^\top & s \end{pmatrix} \cong \begin{pmatrix} \mathbf{D} & \mathbf{e} \\ \mathbf{0}^\top & s \end{pmatrix}$$

La matrice  $\mathbf{D}$  est orthogonale et de déterminant égal à 1, c'est donc une matrice de rotation. Nous pouvons conclure que la famille des homographies  $\mathbf{H}$  est composée de seulement sept paramètres :

- 3 paramètres pour la rotation  $\mathbf{D}$  ;
- 3 paramètres pour la translation  $\mathbf{e}$  ;
- 1 paramètre pour le facteur d'échelle  $s$ .

Lors de la réestimation sans contrainte des points de la mire d'étalonnage, nous venons de montrer qu'il existe une infinité de solutions toutes définies à une similitude près des points tridimensionnels et des matrices de localisation. Si nous souhaitons assurer l'unicité de ces solutions, il faut imposer au moins sept contraintes lors de l'estimation.

Si nous imposons, lors de l'estimation des paramètres, que le point origine de la mire  $(0\ 0\ 0)^\top$  soit constant, nous avons alors la relation suivante :

$$\begin{aligned}\tilde{\mathbf{M}}_c &= \underbrace{\begin{pmatrix} \hat{\mathbf{R}} & \hat{\mathbf{t}} \\ \mathbf{0}^\top & 1 \end{pmatrix}}_{\hat{\mathbf{T}}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{D} & \mathbf{e} \\ \mathbf{0}^\top & s \end{pmatrix} \underbrace{\begin{pmatrix} \mathbf{D}^\top & -\frac{1}{s}\mathbf{D}^\top\mathbf{e} \\ \mathbf{0}^\top & \frac{1}{s} \end{pmatrix}}_{=(0\ 0\ 0\ 1)^\top} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\end{aligned}$$

Sachant que  $\mathbf{D}$  est régulière et que  $s \neq 0$ , nous en déduisons que :

$$\begin{pmatrix} \mathbf{D}^\top & -\frac{1}{s}\mathbf{D}^\top\mathbf{e} \\ \mathbf{0}^\top & \frac{1}{s} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{s}\mathbf{D}^\top\mathbf{e} \\ \frac{1}{s} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \mathbf{e} = \mathbf{0}$$

Imposons qu'un deuxième point de la mire soit constant. Pour rester dans le cas le plus général, posons que ce point ait pour coordonnées non toutes nulles  $\mathbf{M} = (x\ y\ z)^\top$ . Nous avons alors la relation suivante :

$$\begin{pmatrix} \mathbf{D}^\top & \mathbf{0} \\ \mathbf{0}^\top & \frac{1}{s} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \Rightarrow \mathbf{D}^\top\mathbf{M} = s\mathbf{M}$$

L'équation précédente nous indique que  $s$  est une valeur propre de  $\mathbf{D}^\top$ , or une matrice de rotation ne peut avoir pour valeur propre réelle que la valeur 1, donc  $s = 1$ . De plus  $\mathbf{M}$  est un vecteur propre associé à la valeur propre  $s = 1$  : ceci signifie que  $(\mathbf{0}, \mathbf{M})$  est l'axe de rotation de  $\mathbf{D}^\top$ . Rendre constant un deuxième point nous à donc permis d'imposer un facteur d'échelle unitaire et de supprimer deux degrés de libertés pour la rotation. Il reste donc un seul degré de liberté correspondant à la rotation autour de l'axe définie par l'origine  $\mathbf{0}$  et le point  $\mathbf{M}$ .

Sans perte de généralité, nous pouvons toujours définir une estimation du modèle de la mire tel que les deux précédents points constants soient définis par exemple dans le plan  $z = 0$ . Éliminer le dernier degré de liberté devient alors trivial et consiste à laisser constante la coordonnée  $z$  d'un troisième point non colinéaire aux deux premiers.

Dans [LVD99], les auteurs proposent de fixer seulement deux points, ce qui laisse un degré de liberté en rotation autour de l'axe défini par ces deux points. Ce n'est donc pas suffisant pour assurer l'unicité des estimations des paramètres extrinsèques et des points tridimensionnels de la mire.



## Annexe D

# Calcul du gradient des niveaux de gris d'une image rectifiée en fonction du gradient dans l'image initiale

Cette annexe décrit en détail le calcul du gradient des niveaux de gris dans l'image rectifiée  $\nabla\bar{\mathcal{I}}'(\bar{\mathbf{m}}')$  en fonction du gradient des niveaux de gris dans l'image initiale  $\nabla\mathcal{I}'(\mathbf{m}')$  nécessaire aux méthodes d'estimation de paramètres de la stéréo-corrélation fine améliorée de la section 3.1.5, page 91. Commençons par écrire la relation  $\Gamma'$  qui relie les coordonnées corrigées-rectifiées  $\bar{\mathbf{m}}' = (\bar{u}' \ \bar{v}')^\top$  aux coordonnées dans l'image initiale  $\mathbf{m}' = (u' \ v')^\top$ . Cette fonction est la composition de plusieurs transformations élémentaires dont la première d'entre elles donne les coordonnées corrigées et non rectifiées que nous notons  $(U' \ V')^\top$  :

$$\begin{aligned} U'(\bar{u}', \bar{v}') &= \frac{k_{11}\bar{u}' + k_{12}\bar{v}' + k_{13}}{k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33}} \\ V'(\bar{u}', \bar{v}') &= \frac{k_{21}\bar{u}' + k_{22}\bar{v}' + k_{23}}{k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33}} \end{aligned}$$

où  $\mathbf{K} = (k_{ij}) = \mathbf{H}'^{-1}$ , et  $\mathbf{H}'$  est la matrice de rectification de l'image  $\mathcal{I}'$  (cf. section 3.1.2, page 80). La deuxième transformation s'applique sur les coordonnées  $(U' \ V')^\top$  pour déterminer les coordonnées rétinienne — *i.e.* non distordues, non rectifiées et normalisées —  $(x'_r \ y'_r)^\top$  (cf. section 1.1.1, page 8) :

$$\begin{aligned} x'_r(U') &= \frac{U' - u'_0}{\alpha'_u} \\ y'_r(V') &= \frac{V' - v'_0}{\alpha'_v} \end{aligned}$$

Ensuite, ces coordonnées  $(x'_r \ y'_r)^\top$  sont distordues en  $(x'_d \ y'_d)^\top$  — cf. section 1.1.3, page 13 :

$$\begin{aligned} x'_d(x'_r, y'_r) &= x'_r + x'_r \left( r_1 (x_r'^2 + y_r'^2) + r_2 (x_r'^2 + y_r'^2)^2 + r_3 (x_r'^2 + y_r'^2)^3 + \right) \\ &\quad + 2d_1 x'_r y'_r + d_2 (3x_r'^2 + y_r'^2) + p_1 (x_r'^2 + y_r'^2) \\ y'_d(x'_r, y'_r) &= y'_r + y'_r \left( r_1 (x_r'^2 + y_r'^2) + r_2 (x_r'^2 + y_r'^2)^2 + r_3 (x_r'^2 + y_r'^2)^3 + \right) \\ &\quad + 2d_2 x'_r y'_r + d_1 (x_r'^2 + 3y_r'^2) + p_2 (x_r'^2 + y_r'^2) \end{aligned}$$

Enfin, ces coordonnées sont dénormalisées pour obtenir les coordonnées image  $(u' \ v')^\top$  :

$$\begin{aligned} u'(x'_d) &= \alpha'_u x'_d + u'_0 \\ v'(y'_d) &= \alpha'_v y'_d + v'_0 \end{aligned}$$

À partir de ces transformations, nous pouvons écrire l'expression de la fonction vectorielle  $\Gamma'(\bar{\mathbf{m}}')$  :

$$\Gamma'(\bar{\mathbf{m}}') = \mathbf{m}' = \begin{pmatrix} \Gamma_u(\bar{u}', \bar{v}') \\ \Gamma_v(\bar{u}', \bar{v}') \end{pmatrix}$$

avec :

$$\begin{aligned} \Gamma_u(\bar{u}', \bar{v}') &= u' \left( x'_d \left( x'_r \left( U'(\bar{u}', \bar{v}') \right), y'_r \left( V'(\bar{u}', \bar{v}') \right) \right) \right) \\ \Gamma_v(\bar{u}', \bar{v}') &= v' \left( x'_d \left( x'_r \left( U'(\bar{u}', \bar{v}') \right), y'_r \left( V'(\bar{u}', \bar{v}') \right) \right) \right) \end{aligned}$$

Nous pouvons alors écrire l'expression recherchée de  $\nabla \bar{\mathcal{I}}'(\bar{\mathbf{m}}')$  en fonction de  $\nabla \mathcal{I}'(\mathbf{m}')$  :

$$\nabla \bar{\mathcal{I}}'(\bar{\mathbf{m}}') = \nabla \Gamma(\bar{\mathbf{m}}')^\top \cdot \nabla \mathcal{I}'(\mathbf{m}')$$

qui se développe en :

$$\begin{pmatrix} \frac{\partial \bar{\mathcal{I}}'}{\partial \bar{u}'}(\bar{\mathbf{m}}') \\ \frac{\partial \bar{\mathcal{I}}'}{\partial \bar{v}'}(\bar{\mathbf{m}}') \end{pmatrix} = \begin{pmatrix} \frac{\partial \Gamma_u}{\partial \bar{u}'}(\bar{\mathbf{m}}') & \frac{\partial \Gamma_u}{\partial \bar{v}'}(\bar{\mathbf{m}}') \\ \frac{\partial \Gamma_v}{\partial \bar{u}'}(\bar{\mathbf{m}}') & \frac{\partial \Gamma_v}{\partial \bar{v}'}(\bar{\mathbf{m}}') \end{pmatrix} \begin{pmatrix} \frac{\partial \mathcal{I}'}{\partial u'}(\mathbf{m}') \\ \frac{\partial \mathcal{I}'}{\partial v'}(\mathbf{m}') \end{pmatrix}$$

Nous allons à présent détailler l'expression du gradient  $\nabla \Gamma'(\bar{\mathbf{m}}')$ . Pour simplifier l'écriture des équations, nous allons noter toutes les dérivées partielles sous la forme réduite  $\frac{\partial f}{\partial x}$  au lieu de  $\frac{\partial f}{\partial x}(x, y)$ . Chacune des dérivées partielles suivantes sont évaluées au point  $(\bar{u}' \ \bar{v}')^\top$ .

$$\begin{aligned} \frac{\partial \Gamma_u}{\partial \bar{u}'} &= \alpha'_u \frac{\partial x'_d}{\partial u'} & \frac{\partial \Gamma_u}{\partial \bar{v}'} &= \alpha'_u \frac{\partial x'_d}{\partial v'} \\ \frac{\partial \Gamma_v}{\partial \bar{u}'} &= \alpha'_v \frac{\partial y'_d}{\partial u'} & \frac{\partial \Gamma_v}{\partial \bar{v}'} &= \alpha'_v \frac{\partial y'_d}{\partial v'} \end{aligned}$$

Ces équations nécessitent la connaissance des dérivées partielles suivantes :

$$\begin{aligned}
\frac{\partial x'_d}{\partial u'} &= \frac{\partial x'_r}{\partial u'} + \frac{\partial x'_r}{\partial u'} \left( r'_1 (x_r'^2 + y_r'^2) + r'_2 (x_r'^2 + y_r'^2)^2 + r'_3 (x_r'^2 + y_r'^2)^3 \right) \\
&\quad + 2x'_r \left( \frac{\partial x'_r}{\partial u'} x'_r + \frac{\partial y'_r}{\partial u'} y'_r \right) \left( r'_1 + 2r'_2 (x_r'^2 + y_r'^2) + 3r'_3 (x_r'^2 + y_r'^2)^2 \right) \\
&\quad + 2d'_1 \left( \frac{\partial x'_r}{\partial u'} y'_r + \frac{\partial y'_r}{\partial u'} x'_r \right) + 2d'_2 \left( 3 \frac{\partial x'_r}{\partial u'} x'_r + \frac{\partial y'_r}{\partial u'} y'_r \right) + 2p'_1 \left( \frac{\partial x'_r}{\partial u'} x'_r + \frac{\partial y'_r}{\partial u'} y'_r \right) \\
\frac{\partial x'_d}{\partial v'} &= \frac{\partial x'_r}{\partial v'} + \frac{\partial x'_r}{\partial v'} \left( r'_1 (x_r'^2 + y_r'^2) + r'_2 (x_r'^2 + y_r'^2)^2 + r'_3 (x_r'^2 + y_r'^2)^3 \right) \\
&\quad + 2x'_r \left( \frac{\partial x'_r}{\partial v'} x'_r + \frac{\partial y'_r}{\partial v'} y'_r \right) \left( r'_1 + 2r'_2 (x_r'^2 + y_r'^2) + 3r'_3 (x_r'^2 + y_r'^2)^2 \right) \\
&\quad + 2d'_1 \left( \frac{\partial x'_r}{\partial v'} y'_r + \frac{\partial y'_r}{\partial v'} x'_r \right) + 2d'_2 \left( 3 \frac{\partial x'_r}{\partial v'} x'_r + \frac{\partial y'_r}{\partial v'} y'_r \right) + 2p'_1 \left( \frac{\partial x'_r}{\partial v'} x'_r + \frac{\partial y'_r}{\partial v'} y'_r \right) \\
\frac{\partial y'_d}{\partial u'} &= \frac{\partial y'_r}{\partial u'} + \frac{\partial y'_r}{\partial u'} \left( r'_1 (x_r'^2 + y_r'^2) + r'_2 (x_r'^2 + y_r'^2)^2 + r'_3 (x_r'^2 + y_r'^2)^3 \right) \\
&\quad + 2y'_r \left( \frac{\partial x'_r}{\partial u'} x'_r + \frac{\partial y'_r}{\partial u'} y'_r \right) \left( r'_1 + 2r'_2 (x_r'^2 + y_r'^2) + 3r'_3 (x_r'^2 + y_r'^2)^2 \right) \\
&\quad + 2d'_2 \left( \frac{\partial x'_r}{\partial u'} y'_r + \frac{\partial y'_r}{\partial u'} x'_r \right) + 2d'_1 \left( \frac{\partial x'_r}{\partial u'} x'_r + 3 \frac{\partial y'_r}{\partial u'} y'_r \right) + 2p'_2 \left( \frac{\partial x'_r}{\partial u'} x'_r + \frac{\partial y'_r}{\partial u'} y'_r \right) \\
\frac{\partial y'_d}{\partial v'} &= \frac{\partial y'_r}{\partial v'} + \frac{\partial y'_r}{\partial v'} \left( r'_1 (x_r'^2 + y_r'^2) + r'_2 (x_r'^2 + y_r'^2)^2 + r'_3 (x_r'^2 + y_r'^2)^3 \right) \\
&\quad + 2y'_r \left( \frac{\partial x'_r}{\partial v'} x'_r + \frac{\partial y'_r}{\partial v'} y'_r \right) \left( r'_1 + 2r'_2 (x_r'^2 + y_r'^2) + 3r'_3 (x_r'^2 + y_r'^2)^2 \right) \\
&\quad + 2d'_2 \left( \frac{\partial x'_r}{\partial v'} y'_r + \frac{\partial y'_r}{\partial v'} x'_r \right) + 2d'_1 \left( \frac{\partial x'_r}{\partial v'} x'_r + 3 \frac{\partial y'_r}{\partial v'} y'_r \right) + 2p'_2 \left( \frac{\partial x'_r}{\partial v'} x'_r + \frac{\partial y'_r}{\partial v'} y'_r \right)
\end{aligned}$$

Ici encore, il reste à déterminer l'expression d'autres dérivées partielles :

$$\begin{aligned}
\frac{\partial x'_r}{\partial u'} &= \frac{1}{\alpha'_u} \frac{\partial U'}{\partial \bar{u}'} & \frac{\partial x'_r}{\partial v'} &= \frac{1}{\alpha'_v} \frac{\partial U'}{\partial \bar{v}'} \\
\frac{\partial y'_r}{\partial u'} &= \frac{1}{\alpha'_v} \frac{\partial V'}{\partial \bar{u}'} & \frac{\partial y'_r}{\partial v'} &= \frac{1}{\alpha'_v} \frac{\partial V'}{\partial \bar{v}'}
\end{aligned}$$

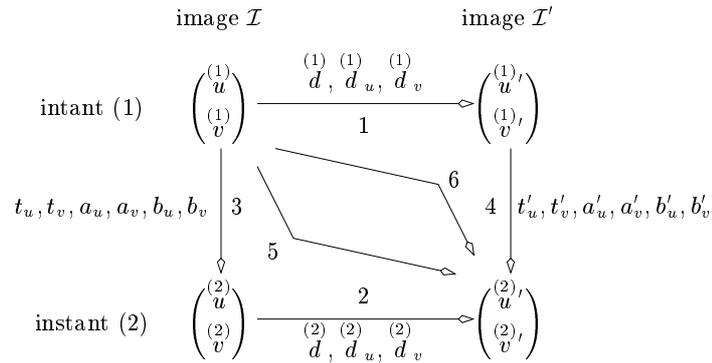
Enfin, les dérivées partielles suivantes complètent la connaissance de l'expression globale de

$\nabla\Gamma'(\bar{m}')$  :

$$\begin{aligned} \frac{\partial U'}{\partial \bar{u}'} &= \frac{k_{11}}{k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33}} - k_{31} \frac{k_{11}\bar{u}' + k_{12}\bar{v}' + k_{13}}{(k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33})^2} \\ \frac{\partial U'}{\partial \bar{v}'} &= \frac{k_{12}}{k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33}} - k_{32} \frac{k_{11}\bar{u}' + k_{12}\bar{v}' + k_{13}}{(k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33})^2} \\ \frac{\partial V'}{\partial \bar{u}'} &= \frac{k_{21}}{k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33}} - k_{31} \frac{k_{21}\bar{u}' + k_{22}\bar{v}' + k_{23}}{(k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33})^2} \\ \frac{\partial V'}{\partial \bar{v}'} &= \frac{k_{22}}{k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33}} - k_{32} \frac{k_{21}\bar{u}' + k_{22}\bar{v}' + k_{23}}{(k_{31}\bar{u}' + k_{32}\bar{v}' + k_{33})^2} \end{aligned}$$

## Annexe E

# Paramétrisation minimale des transformations géométriques locales de la méthode de stéréo-corrélation et tracking simultanés



**FIG. E.1 :** Les différents paramètres des transformations géométriques locales à l'ordre 1.

Afin de simplifier l'exposé, nous nous limiterons à l'étude de la paramétrisation minimale des transformations géométriques locales à l'ordre 1. La figure E.1 montre les différents paramètres des transformations géométriques entre les deux instants d'acquisition des paires d'images et entre les deux images de chaque paire. Exprimons chacune des transformations repérées par les chiffres un à six dans cette figure. La première d'entre elles est la relation stéréoscopique à l'instant 1 :

$$\begin{aligned} \begin{pmatrix} u' \\ v' \end{pmatrix} &= \begin{pmatrix} u \\ v \end{pmatrix} \\ \begin{pmatrix} u' \\ v' \end{pmatrix} &= \begin{pmatrix} u \\ v \end{pmatrix} + d + d_u \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} u_c \\ v_c \end{pmatrix} + d_v \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} u_c \\ v_c \end{pmatrix} \end{aligned}$$

La deuxième relation est similaire à la précédente pour l'instant 2 et s'écrit :

$$\begin{aligned} \begin{pmatrix} u' \\ v' \end{pmatrix} &= \begin{pmatrix} u \\ v \end{pmatrix} \\ \begin{pmatrix} u' \\ v' \end{pmatrix} &= \begin{pmatrix} u \\ v \end{pmatrix} + d + d_u \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} u_c \\ v_c \end{pmatrix} + d_v \begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} u_c \\ v_c \end{pmatrix} \end{aligned}$$

La troisième relation exprime la transformation temporelle entre les deux instants pour les images de la caméra gauche :

$$\begin{aligned} \begin{pmatrix} 2 \\ u \end{pmatrix} &= \begin{pmatrix} 1 \\ u \end{pmatrix} + t_u + a_u \begin{pmatrix} 1 \\ u \end{pmatrix} - \begin{pmatrix} 1 \\ u_c \end{pmatrix} + a_v \begin{pmatrix} 1 \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ v_c \end{pmatrix} \\ \begin{pmatrix} 2 \\ v \end{pmatrix} &= \begin{pmatrix} 1 \\ v \end{pmatrix} + t_v + b_u \begin{pmatrix} 1 \\ u \end{pmatrix} - \begin{pmatrix} 1 \\ u_c \end{pmatrix} + b_v \begin{pmatrix} 1 \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ v_c \end{pmatrix} \end{aligned}$$

La quatrième relation est similaire à la précédente pour la caméra droite et s'écrit :

$$\begin{aligned} \begin{pmatrix} 2 \\ u' \end{pmatrix} &= \begin{pmatrix} 1 \\ u' \end{pmatrix} + t'_u + a'_u \begin{pmatrix} 1 \\ u' \end{pmatrix} - \begin{pmatrix} 1 \\ u'_c \end{pmatrix} + a'_v \begin{pmatrix} 1 \\ v' \end{pmatrix} - \begin{pmatrix} 1 \\ v'_c \end{pmatrix} \\ \begin{pmatrix} 2 \\ v' \end{pmatrix} &= \begin{pmatrix} 1 \\ v' \end{pmatrix} + t'_v + b'_u \begin{pmatrix} 1 \\ u' \end{pmatrix} - \begin{pmatrix} 1 \\ u'_c \end{pmatrix} + b'_v \begin{pmatrix} 1 \\ v' \end{pmatrix} - \begin{pmatrix} 1 \\ v'_c \end{pmatrix} \end{aligned}$$

La cinquième relation est le résultat de la composition de la deuxième et de la troisième relation et s'écrit :

$$\begin{aligned} \begin{pmatrix} 2 \\ u' \end{pmatrix} &= \begin{pmatrix} 1 \\ u \end{pmatrix} + t_u + a_u \begin{pmatrix} 1 \\ u \end{pmatrix} - \begin{pmatrix} 1 \\ u_c \end{pmatrix} + a_v \begin{pmatrix} 1 \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ v_c \end{pmatrix} \\ \begin{pmatrix} 2 \\ v' \end{pmatrix} &= \begin{pmatrix} 1 \\ v \end{pmatrix} + \overset{(2)}{d} + t_v + \\ &\quad \left( b_u + \overset{(2)}{d}_u + a_u \overset{(2)}{d}_u + b_u \overset{(2)}{d}_v \right) \begin{pmatrix} 1 \\ u \end{pmatrix} - \begin{pmatrix} 1 \\ u_c \end{pmatrix} + \left( b_v + \overset{(2)}{d}_v + a_v \overset{(2)}{d}_u + b_v \overset{(2)}{d}_v \right) \begin{pmatrix} 1 \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ v_c \end{pmatrix} \end{aligned}$$

Enfin, la sixième relation est déterminée par composition de la première et de la quatrième relation :

$$\begin{aligned} \begin{pmatrix} 2 \\ u' \end{pmatrix} &= \begin{pmatrix} 1 \\ u \end{pmatrix} + t'_u + \left( a'_u + a'_v \overset{(1)}{d}_u \right) \begin{pmatrix} 1 \\ u \end{pmatrix} - \begin{pmatrix} 1 \\ u_c \end{pmatrix} + \left( a'_v + a'_v \overset{(1)}{d}_v \right) \begin{pmatrix} 1 \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ v_c \end{pmatrix} \\ \begin{pmatrix} 2 \\ v' \end{pmatrix} &= \begin{pmatrix} 1 \\ v \end{pmatrix} + \overset{(1)}{d} + t'_v + \left( \overset{(1)}{d}_u + b'_u + b'_v \overset{(1)}{d}_u \right) \begin{pmatrix} 1 \\ u \end{pmatrix} - \begin{pmatrix} 1 \\ u_c \end{pmatrix} + \left( \overset{(1)}{d}_v + b'_v + b'_v \overset{(1)}{d}_v \right) \begin{pmatrix} 1 \\ v \end{pmatrix} - \begin{pmatrix} 1 \\ v_c \end{pmatrix} \end{aligned}$$

En identifiant terme à terme ces deux dernières relations, nous constatons les égalités suivantes qui exhibent des dépendances des paramètres entre eux :

$$\begin{aligned} t_u &= t'_u \\ \overset{(1)}{d} + t'_v &= \overset{(2)}{d} + t_v \\ a_u &= a'_u + a'_v \overset{(1)}{d}_u \\ a_v &= a'_v + a'_v \overset{(1)}{d}_v \\ b_u + \overset{(2)}{d}_u + a_u \overset{(2)}{d}_u + b_u \overset{(2)}{d}_v &= b'_u + \overset{(1)}{d}_u + b'_v \overset{(1)}{d}_u \\ b_v + \overset{(2)}{d}_v + a_v \overset{(2)}{d}_u + b_v \overset{(2)}{d}_v &= b'_v + \overset{(1)}{d}_v + b'_v \overset{(1)}{d}_v \end{aligned}$$

Pour parvenir à une représentation minimale de ces transformations, il est nécessaire de sup-

primer ces dépendances. Pour cela nous posons :

$$\begin{aligned}
t'_u &= t_u \\
t'_v &= \overset{(2)}{d} - \overset{(1)}{d} + t_v \\
a_u &= a'_u + a'_v \overset{(1)}{d}_u \\
a_v &= a'_v \left(1 + \overset{(1)}{d}_v\right) \\
b_u &= \frac{b'_u + \overset{(1)}{d}_u \left(1 + b'_v\right) - \overset{(2)}{d}_u \left(1 + a'_u + a'_v \overset{(1)}{d}_u\right)}{1 + \overset{(2)}{d}_v} \\
b_v &= \frac{b'_v + \overset{(1)}{d}_v \left(1 + b'_v\right) - \overset{(2)}{d}_v - a'_v \overset{(2)}{d}_u \left(1 + \overset{(1)}{d}_v\right)}{1 + \overset{(2)}{d}_v}
\end{aligned}$$

La représentation minimale de l'ensemble des transformations comporte donc seulement 12 paramètres :

$$\mathbf{v} = \left( \overset{(1)}{d} \ \overset{(1)}{d}_u \ \overset{(1)}{d}_v \ \overset{(2)}{d} \ \overset{(2)}{d}_u \ \overset{(2)}{d}_v \ t_u \ t_v \ a'_u \ a'_v \ b'_u \ b'_v \right)^\top$$

En posant  $\delta_u = \overset{(1)}{u} - \overset{(1)}{u}_c$  et  $\delta_v = \overset{(1)}{v} - \overset{(1)}{v}_c$ , les trois transformations géométriques locales qui nous intéressent dans l'expression du critère de corrélation (4.1) s'écrivent alors sous les formes paramétrées minimales suivantes :

$$\begin{aligned}
\overset{(1)}{u}' &= \overset{(1)}{u} \\
\overset{(1)}{v}' &= \overset{(1)}{v} + \overset{(1)}{d} + \delta_u \overset{(1)}{d}_u + \delta_v \overset{(1)}{d}_v \\
\overset{(2)}{u} &= \overset{(1)}{u} + t_u + \delta_u \left( a'_u + a'_v \overset{(1)}{d}_u \right) + \delta_v a'_v \left( 1 + \overset{(1)}{d}_v \right) \\
\overset{(2)}{v} &= \overset{(1)}{v} + t_v + \delta_u \frac{b'_u + \overset{(1)}{d}_u \left( 1 + b'_v \right) - \overset{(2)}{d}_u \left( 1 + a'_u + a'_v \overset{(1)}{d}_u \right)}{1 + \overset{(2)}{d}_v} + \\
&\quad \delta_v \frac{b'_v + \overset{(1)}{d}_v \left( 1 + b'_v \right) - \overset{(2)}{d}_v - a'_v \overset{(2)}{d}_u \left( 1 + \overset{(1)}{d}_v \right)}{1 + \overset{(2)}{d}_v} \\
\overset{(2)}{u}' &= \overset{(1)}{u} + t_u + \delta_u \left( a'_u + a'_v \overset{(1)}{d}_u \right) + \delta_v \left( a'_v + a'_v \overset{(1)}{d}_v \right) \\
\overset{(2)}{v}' &= \overset{(1)}{v} + \overset{(2)}{d} + t_v + \delta_u \left( \overset{(1)}{d}_u + b'_u + b'_v \overset{(1)}{d}_u \right) + \delta_v \left( \overset{(1)}{d}_v + b'_v + b'_v \overset{(1)}{d}_v \right)
\end{aligned}$$



## Annexe F

# Algorithme de Levenberg-Marquardt [Mar63, PTVF92]

Soit le système de  $n$  équations non linéaires suivant :

$$\mathbf{y}(\mathbf{x}, \mathbf{p}) = \begin{pmatrix} y_1(\mathbf{x}, \mathbf{p}) \\ y_2(\mathbf{x}, \mathbf{p}) \\ \vdots \\ y_N(\mathbf{x}, \mathbf{p}) \end{pmatrix}$$

où  $\mathbf{p}$  est le vecteur des  $P$  paramètres à optimiser :  $\mathbf{p} = (p_1 \quad p_2 \quad \cdots \quad p_P)$ . Soit  $\bar{E}$  l'ensemble de  $M$  mesures (données expérimentales) suivant :

$$\bar{E} = \{(\bar{\mathbf{x}}_1, \bar{\mathbf{y}}_1), \dots, (\bar{\mathbf{x}}_M, \bar{\mathbf{y}}_M)\}$$

Nous cherchons à optimiser les paramètres  $\mathbf{p}$  des équations  $\mathbf{y}(\mathbf{x}, \mathbf{p})$  sur l'ensemble des mesures  $\bar{E}$ . Pour cela, nous introduisons les  $N$  vecteurs des erreurs d'ajustement, appelés **vecteurs résidus** :

$$\bar{\mathbf{r}}_i = \begin{pmatrix} y_i(\bar{\mathbf{x}}_1, \mathbf{p}) \\ y_i(\bar{\mathbf{x}}_2, \mathbf{p}) \\ \vdots \\ y_i(\bar{\mathbf{x}}_M, \mathbf{p}) \end{pmatrix} - \bar{\mathbf{y}}_i$$

Nous regroupons ces vecteurs dans un seul vecteur des résidus  $\mathbf{r}$  :

$$\mathbf{r} = \begin{pmatrix} \bar{\mathbf{r}}_1 \\ \bar{\mathbf{r}}_2 \\ \vdots \\ \bar{\mathbf{r}}_N \end{pmatrix}$$

L'optimisation des paramètres  $\mathbf{p}$  sur l'ensemble des mesures  $\bar{E}$  consiste à minimiser le critère suivant :

$$\chi^2(\mathbf{p}) = \sum_{i=1}^N \sum_{j=1}^M r_{ij}^2 = \mathbf{r}^\top \mathbf{r} \quad (\text{F.1})$$

La méthode de LEVENBERG-MARQUARDT est un algorithme itératif qui fait intervenir les dérivées première et seconde du critère  $\chi^2(\mathbf{p})$ . Ces dérivées sont respectivement le vecteur jacobien  $\mathbf{j}(\mathbf{p})$  et la matrice hessienne  $\mathbf{H}(\mathbf{p})$  :

$$\begin{aligned}\mathbf{j}(\mathbf{p}) &\equiv \left[ \frac{\partial \chi^2}{\partial p_i} \right]_i = 2\mathbf{A}\mathbf{r} \\ \mathbf{H}(\mathbf{p}) &\equiv \left[ \frac{\partial^2 \chi^2}{\partial p_i \partial p_j} \right]_{i,j} = 2\mathbf{A}\mathbf{A}^\top + 2 \left[ r_i \frac{\partial^2 r_i}{\partial p_j \partial p_k} \right]_{j,k}\end{aligned}\quad (\text{F.2})$$

où  $\mathbf{A}$  représente la matrice transposée de la matrice jacobienne du système  $\mathbf{y}(\mathbf{x}, \mathbf{p})$  :

$$\mathbf{A} = \left[ \frac{\partial r_j}{\partial p_i} \right]_{i,j} = \left[ \frac{\partial y_j}{\partial p_i} \right]_{i,j}$$

Au voisinage de la solution recherchée, les résidus  $r_i$  sont en principe très faibles. La matrice hessienne  $\mathbf{H}(\mathbf{p})$  peut donc être approchée par l'expression :

$$\mathbf{H}(\mathbf{p}) \approx 2\mathbf{A}\mathbf{A}^\top \quad (\text{F.3})$$

Le minimum du critère  $\chi^2(\mathbf{p})$  est itérativement recherché par une série d'approximations de la solution :

$$\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} + \delta^{(n)}$$

où nous résolvons, à chaque itération, le problème suivant :

Trouver  $\delta^{(n)}$  solution de  $\inf \left\{ q^{(n)}(\delta), \|\delta\| \leq h \right\}$ , où  $q^{(n)}(\delta)$  est une approximation de  $\chi^2(\mathbf{p}^{(n)} + \delta)$  définie par :

$$\left| \begin{array}{l} q^{(n)}(\delta) = \chi^2 + \mathbf{j}^\top \cdot \delta + \frac{1}{2} \delta^\top \cdot \mathbf{H} \cdot \delta \\ h > 0 \text{ est un paramètre à définir} \end{array} \right.$$

avec :

$$\begin{aligned}\chi^2 &\equiv \chi^2(\mathbf{p}) \\ \mathbf{j} &\equiv \mathbf{j}(\mathbf{p}) \\ \mathbf{H} &\equiv \mathbf{H}(\mathbf{p})\end{aligned}$$

LEVENBERG et MARQUARDT ont montré que si  $\delta^{(n)}$  est solution du problème posé et que si  $\|\cdot\|$  est une norme euclidienne, alors il existe un réel positif  $\lambda$  tel que :

$$\left( \mathbf{H} + \lambda \mathbf{I} \right) \delta = -\mathbf{j} \quad (\text{F.4})$$

où  $\mathbf{I}$  est la matrice identité. Leur algorithme se résume à :

1.  $\mathbf{P}^{(1)}$  et  $\lambda^{(1)}$  sont donnés ;
2. sachant  $\mathbf{P}^{(n)}$  on calcule  $\mathbf{j}^{(n)}$  et  $\mathbf{H}^{(n)}$  selon les équations (F.2) et (F.3) ;
3. on résout le système linéaire de l'équation (F.4) pour obtenir  $\delta^{(n)}$  :

$$\delta^{(n)} = - \left( \mathbf{H}^{(n)} + \lambda^{(n)} \mathbf{I} \right)^{-1} \cdot \mathbf{j}^{(n)}$$

4. on calcule  $\chi^2 \left( \mathbf{P}^{(n)} + \delta^{(n)} \right)$  par l'équation (F.1) ;
5. si  $\chi^2 \left( \mathbf{P}^{(n)} + \delta^{(n)} \right) < \chi^2 \left( \mathbf{P}^{(n)} \right)$  alors on prend :

$$\begin{aligned} \lambda^{(n+1)} &= \frac{\lambda^{(n)}}{2} \\ \mathbf{P}^{(n+1)} &= \mathbf{P}^{(n)} + \delta^{(n)} \end{aligned}$$

sinon on prend :

$$\begin{aligned} \lambda^{(n+1)} &= 2 \lambda^{(n)} \\ \mathbf{P}^{(n+1)} &= \mathbf{P}^{(n)} \end{aligned}$$

6. si le critère de convergence est satisfait on arrête le calcul, sinon  $n = n + 1$  et on reprend à l'étape 2.

Notons que initialement,  $\lambda^{(1)} > 0$  est choisi de façon arbitraire. Le processus itératif prend fin quand le nombre d'itérations  $n_{max}$  est atteint (divergence), ou quand le critère de convergence suivant est satisfait :

$$\left| p_i^{(n+1)} - p_i^{(n)} \right| \leq \varepsilon \cdot \left( \left| p_i^{(n)} \right| + \tau \right); \quad i = 1 \dots P$$

Avec  $\varepsilon$  et  $\tau$  pris généralement égaux à  $10^{-3}$ , et  $\lambda^{(1)} = 1$ .



# Annexe G

## Jacobiennes

### G.1 Jacobiennes du modèle de caméra

Les matrices jacobiennes du modèle de caméra par rapport aux paramètres  $u_0, v_0, \alpha_u$  et  $\alpha_v$ , et par rapport aux paramètres de distorsion  $r_1, r_2, r_3, p_1, p_2, d_1$  et  $d_2$ , sont :

$$\frac{\partial u, v}{\partial u_0, v_0, \alpha_u, \alpha_v} = \begin{pmatrix} 1 & 0 & x_d & 0 \\ 0 & 1 & 0 & y_d \end{pmatrix} \quad (\text{G.1})$$

$$\frac{\partial u, v}{\partial r_1, r_2, r_3, p_1, p_2, d_1, d_2} = \frac{\partial u, v}{\partial x_d, y_d} \cdot \frac{\partial x_d, y_d}{\partial r_1, r_2, r_3, p_1, p_2, d_1, d_2} \quad (\text{G.2})$$

Les jacobiennes du modèle par rapport aux paramètres extrinsèques sont :

$$\frac{\partial u, v}{\partial n_x, n_y, n_z} = \frac{\partial u, v}{\partial x_c, y_c, z_c} \cdot \frac{\partial x_c, y_c, z_c}{\partial n_x, n_y, n_z} \quad (\text{G.3})$$

$$\frac{\partial u, v}{\partial t_x, t_y, t_z} = \frac{\partial u, v}{\partial x_c, y_c, z_c} \cdot \frac{\partial x_c, y_c, z_c}{\partial t_x, t_y, t_z} \quad (\text{G.4})$$

L'équation (G.2) nous impose de connaître les dérivées partielles d'un point image en fonction de ses coordonnées normalisées distordues :

$$\frac{\partial u, v}{\partial x_d, y_d} = \begin{pmatrix} \alpha_u & 0 \\ 0 & \alpha_v \end{pmatrix}$$

Posons  $\delta = x_r^2 + y_r^2$ . La jacobienne des coordonnées normalisées distordues en fonction des paramètres de distorsion, nécessaire à l'expression de (G.2), est :

$$\frac{\partial x_d, y_d}{\partial r_1, r_2, r_3, p_1, p_2, d_1, d_2} = \begin{pmatrix} x_r \delta & x_r \delta^2 & x_r \delta^3 & \delta & 0 & 3x_r^2 + y_r^2 & 2x_r y_r \\ y_r \delta & y_r \delta^2 & y_r \delta^3 & 0 & \delta & 2x_r y_r & x_r^2 + 3y_r^2 \end{pmatrix}$$

Les jacobiennes (G.3) et (G.4) nécessitent la connaissance des jacobiennes suivantes :

$$\begin{aligned} \frac{\partial u, v}{\partial x_c, y_c, z_c} &= \frac{\partial u, v}{\partial x_d, y_d} \cdot \frac{\partial x_d, y_d}{\partial x_c, y_c, z_c} \\ \frac{\partial x_c, y_c, z_c}{\partial n_x, n_y, n_z} &= \left( \frac{\partial \mathbf{R}}{\partial n_x} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid \frac{\partial \mathbf{R}}{\partial n_y} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid \frac{\partial \mathbf{R}}{\partial n_z} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) \\ \frac{\partial x_c, y_c, z_c}{\partial t_x, t_y, t_z} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (\text{G.5})$$

où

$$\begin{aligned} \frac{\partial x_d, y_d}{\partial x_c, y_c, z_c} &= \frac{\partial x_d, y_d}{\partial x_r, y_r} \cdot \frac{\partial x_r, y_r}{\partial x_c, y_c, z_c} \\ \frac{\partial x_r, y_r}{\partial x_c, y_c, z_c} &= \begin{pmatrix} \frac{1}{z_c} & 0 & -\frac{x_c}{z_c^2} \\ 0 & \frac{1}{z_c} & -\frac{y_c}{z_c^2} \end{pmatrix} \end{aligned}$$

et

$$\frac{\partial x_d, y_d}{\partial x_r, y_r} : \begin{cases} \frac{\partial x_d}{\partial x_r} = 1 + r_1\delta + r_2\delta^2 + r_3\delta^3 + 2x_r^2 (r_1 + 2r_2\delta + 3r_3\delta^2) + 2x_r (p_1 + 3d_1) + 2d_2 y_r \\ \frac{\partial x_d}{\partial y_r} = 2x_r y_r (r_1 + 2r_2\delta + 3r_3\delta^2) + 2y_r (p_1 + d_1) + 2d_2 x_r \\ \frac{\partial y_d}{\partial x_r} = 2x_r y_r (r_1 + 2r_2\delta + 3r_3\delta^2) + 2x_r (p_2 + d_2) + 2d_1 y_r \\ \frac{\partial y_d}{\partial y_r} = 1 + r_1\delta + r_2\delta^2 + r_3\delta^3 + 2y_r^2 (r_1 + 2r_2\delta + 3r_3\delta^2) + 2y_r (p_2 + 3d_2) + 2d_1 x_r \end{cases}$$

Les matrices  $\frac{\partial \mathbf{R}}{\partial n_x}$ ,  $\frac{\partial \mathbf{R}}{\partial n_y}$  et  $\frac{\partial \mathbf{R}}{\partial n_z}$  de l'équation (G.5) sont données par l'équation (H.1). La matrice des dérivées partielles de  $u$  et  $v$  par rapport au point de mesure  $\mathbf{M}$  est donnée par :

$$\begin{aligned} \frac{\partial u, v}{\partial x, y, z} &= \frac{\partial u, v}{\partial x_c, y_c, z_c} \frac{\partial x_c, y_c, z_c}{\partial x, y, z} \\ \frac{\partial x_c, y_c, z_c}{\partial x, y, z} &= \mathbf{R} \end{aligned}$$

## G.2 Jacobiennes du modèle de capteur de stéréovision — méthode GLOB

Les jacobiennes relatives aux paramètres de la caméra gauche et aux paramètres extrinsèques sont identiques à celles développées dans la section précédente. Nous allons donc nous attacher à exprimer seulement les jacobiennes relatives aux paramètres de la caméra droite et à la transformation rigide entre les référentiels des deux caméras.

Les matrices jacobiennes de  $u'$  et  $v'$  par rapport aux paramètres  $u'_0$ ,  $v'_0$ ,  $k'_u$  et  $k'_v$ , ainsi que par rapport aux paramètres de distorsion  $r'_1$ ,  $r'_2$ ,  $r'_3$  se déduisent directement des équations (G.1) et (G.2). Les matrices des dérivées partielles de  $u'$  et  $v'$  par rapport aux paramètres de localisation de la deuxième caméra vis à vis de la première, transformation rigide exprimée sous la forme minimale  $\mathbf{T}_s(\mathbf{n}_s, \mathbf{t}_s)$ , sont :

$$\begin{aligned} \frac{\partial u', v'}{\partial n_{s_x}, n_{s_y}, n_{s_z}} &= \frac{\partial u', v'}{\partial x'_c, y'_c, z'_c} \frac{\partial x'_c, y'_c, z'_c}{\partial n_{s_x}, n_{s_y}, n_{s_z}} \\ \frac{\partial u', v'}{\partial t_{s_x}, t_{s_y}, t_{s_z}} &= \frac{\partial u', v'}{\partial x'_c, y'_c, z'_c} \frac{\partial x'_c, y'_c, z'_c}{\partial t_{s_x}, t_{s_y}, t_{s_z}} \end{aligned}$$

avec :

$$\frac{\partial x'_c, y'_c, z'_c}{\partial n_{s_x}, n_{s_y}, n_{s_z}} = \left( \frac{\partial \mathbf{R}'}{\partial n_{s_x}} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \mid \frac{\partial \mathbf{R}'}{\partial n_{s_y}} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \mid \frac{\partial \mathbf{R}'}{\partial n_{s_z}} \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \right)$$

Les deux jacobiennes de  $u'$  et  $v'$  par rapport aux paramètres extrinsèques de la première caméra sont :

$$\begin{aligned} \frac{\partial u', v'}{\partial n_x, n_y, n_z} &= \frac{\partial u', v'}{\partial x_c, y_c, z_c} \frac{\partial x_c, y_c, z_c}{\partial n_x, n_y, n_z} \\ \frac{\partial u', v'}{\partial t_x, t_y, t_z} &= \frac{\partial u', v'}{\partial x_c, y_c, z_c} \frac{\partial x_c, y_c, z_c}{\partial t_x, t_y, t_z} \end{aligned} \quad (\text{G.6})$$

La jacobienne (G.6) nécessite la connaissance de la jacobienne suivante :

$$\begin{aligned} \frac{\partial u', v'}{\partial x_c, y_c, z_c} &= \frac{\partial u', v'}{\partial x'_c, y'_c, z'_c} \frac{\partial x'_c, y'_c, z'_c}{\partial x_c, y_c, z_c} \\ \text{avec} \quad \frac{\partial x'_c, y'_c, z'_c}{\partial x_c, y_c, z_c} &= \mathbf{R}' \end{aligned}$$

La matrice des dérivées partielles de  $u'$  et  $v'$  par rapport au point de mesure  $\mathbf{M}$  est :

$$\frac{\partial u', v'}{\partial x, y, z} = \frac{\partial u', v'}{\partial x_c, y_c, z_c} \frac{\partial x_c, y_c, z_c}{\partial x, y, z}$$

### G.3 Jacobiennes du critère de corrélation de la méthode de stéréoscopie et suivi de pixels simultanés

Nous rappelons ici l'équation (4.1), expression d'un critère de corrélation de type SSD mesurant la similarité de quatre domaines simultanément (voir la section 4.2, page 109, pour une description détaillée des notations utilisées) :

$$C(\mathbf{m}_c, \mathbf{v}) = \sum_{\mathbf{m}_i \in V(\mathbf{m}_c)}^{(1)} \left[ \left( \mathcal{I}'^{(1)}(\mathbf{m}_i) - \mathcal{I}^{(1)}(\mathbf{m}_i) \right)^2 + \left( \mathcal{I}^{(2)}(\mathbf{m}_i) - \mathcal{I}^{(1)}(\mathbf{m}_i) \right)^2 + \left( \mathcal{I}'^{(2)}(\mathbf{m}_i) - \mathcal{I}^{(1)}(\mathbf{m}_i) \right)^2 \right]$$

En utilisant des modèles de transformations géométriques locales à l'ordre 1, les relations entre les points  $\mathbf{m}_i^{(1)}$  et ses correspondants  $\mathbf{m}'_i$ ,  $\mathbf{m}_i^{(2)}$  et  $\mathbf{m}'_i$  sont définies dans la section 4.2.2, page 113, et comportent 12 paramètres. Pour chacun des trois résidus exprimés, nous posons :

$$\begin{aligned} g_1(\mathbf{u}, \mathbf{v}, \mathbf{v}) &= \mathcal{I}'^{(1)}(\mathbf{u}', \mathbf{v}') - \mathcal{I}^{(1)}(\mathbf{u}, \mathbf{v}) \\ g_2(\mathbf{u}, \mathbf{v}, \mathbf{v}) &= \mathcal{I}^{(2)}(\mathbf{u}, \mathbf{v}) - \mathcal{I}^{(1)}(\mathbf{u}, \mathbf{v}) \\ g_3(\mathbf{u}, \mathbf{v}, \mathbf{v}) &= \mathcal{I}'^{(2)}(\mathbf{u}', \mathbf{v}') - \mathcal{I}^{(1)}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

La matrice jacobienne de l'expression du premier résidu  $g_1$  par rapport aux douze paramètres du vecteur  $\mathbf{v}$  est composée des éléments suivants :

$$\begin{aligned} \frac{\partial g_1}{\partial d^{(1)}} &= \frac{\partial \mathcal{I}'}{\partial v} & \frac{\partial g_1}{\partial d_u^{(1)}} &= \delta_u \frac{\partial \mathcal{I}'}{\partial v} & \frac{\partial g_1}{\partial d_v^{(1)}} &= \delta_v \frac{\partial \mathcal{I}'}{\partial v} \\ \frac{\partial g_1}{\partial d^{(2)}} &= \frac{\partial g_1}{\partial d_u^{(2)}} = \frac{\partial g_1}{\partial d_v^{(2)}} = \frac{\partial g_1}{\partial t_u} = \frac{\partial g_1}{\partial t_v} = \frac{\partial g_1}{\partial a'_u} = \frac{\partial g_1}{\partial a'_v} = \frac{\partial g_1}{\partial b'_u} = \frac{\partial g_1}{\partial b'_v} = 0 \end{aligned}$$

La matrice jacobienne de l'expression du deuxième résidu  $g_2$  par rapport aux douze paramètres du vecteur  $\mathbf{v}$  est composée des éléments suivants :

$$\frac{\partial g_2}{\partial d^{(1)}} = \frac{\partial g_2}{\partial d^{(2)}} = 0$$

$$\frac{\partial g_2}{\partial d_u^{(1)}} = \delta_u \left( a'_v \frac{\partial \mathcal{I}^{(2)}}{\partial u} + \frac{1 + b'_v - a'_v d_u^{(2)}}{1 + d_v^{(2)}} \frac{\partial \mathcal{I}^{(2)}}{\partial v} \right)$$

$$\frac{\partial g_2}{\partial d_v^{(1)}} = \delta_v \left( a'_u \frac{\partial \mathcal{I}^{(2)}}{\partial u} + \frac{1 + b'_u - a'_u d_v^{(2)}}{1 + d_u^{(2)}} \frac{\partial \mathcal{I}^{(2)}}{\partial v} \right)$$

$$\frac{\partial g_2}{\partial d_u^{(2)}} = - \frac{\delta_u \left( 1 + a'_u + a'_v d_u^{(1)} \right) + \delta_v a'_v \left( 1 + d_v^{(1)} \right) \frac{\partial \mathcal{I}^{(2)}}{\partial v}}{1 + d_v^{(2)}}$$

$$\frac{\partial g_2}{\partial d_v^{(2)}} = - \frac{\delta_u}{\left( 1 + d_v^{(2)} \right)^2} \left( b'_u + d_u^{(1)} (1 + b'_v) - d_u^{(2)} \left( 1 + a'_u + a'_v d_u^{(1)} \right) \right) \frac{\partial \mathcal{I}^{(2)}}{\partial v}$$

$$- \frac{\delta_v}{\left( 1 + d_u^{(2)} \right)^2} \left( 1 + b'_v + d_v^{(1)} (1 + b'_u) - d_v^{(2)} \left( a'_v + a'_u d_v^{(1)} \right) \right) \frac{\partial \mathcal{I}^{(2)}}{\partial v}$$

$$\frac{\partial g_2}{\partial t_u} = \frac{\partial \mathcal{I}^{(2)}}{\partial t_u^{(2)}}$$

$$\frac{\partial g_2}{\partial t_v} = \frac{\partial \mathcal{I}^{(2)}}{\partial t_v^{(2)}}$$

$$\frac{\partial g_2}{\partial a'_u} = \delta_u \frac{\partial \mathcal{I}^{(2)}}{\partial u} - \delta_u \frac{d_u^{(2)}}{1 + d_v^{(2)}} \frac{\partial \mathcal{I}^{(2)}}{\partial v}$$

$$\frac{\partial g_2}{\partial a'_v} = \left( \delta_u d_u^{(1)} + \delta_v \left( 1 + d_v^{(1)} \right) \right) \frac{\partial \mathcal{I}^{(2)}}{\partial u} - \frac{\delta_u d_u^{(1)} d_u^{(2)} + \delta_v d_u^{(2)} \left( 1 + d_v^{(1)} \right) \frac{\partial \mathcal{I}^{(2)}}{\partial v}}{1 + d_v^{(2)}}$$

$$\frac{\partial g_2}{\partial b'_u} = \frac{\delta_u}{1 + d_v^{(2)}} \frac{\partial \mathcal{I}^{(2)}}{\partial v}$$

$$\frac{\partial g_2}{\partial b'_v} = \frac{\delta_u d_u^{(1)} + \delta_v \left( 1 + d_v^{(1)} \right) \frac{\partial \mathcal{I}^{(2)}}{\partial v}}{1 + d_v^{(2)}}$$

Enfin, la matrice jacobienne de l'expression du troisième résidu  $g_3$  par rapport aux douze paramètres du vecteur  $\mathbf{v}$  est composée des éléments suivants :

$$\frac{\partial g_3}{\partial d^{(1)}} = \frac{\partial g_3}{\partial d_u^{(2)}} = \frac{\partial g_3}{\partial d_v^{(2)}} = 0$$

$$\begin{aligned}
\frac{\partial g_3}{\partial d_u^{(1)}} &= \delta_u a'_v \frac{\partial \mathcal{I}'}{\partial u} + \delta_u (1 + b'_v) \frac{\partial \mathcal{I}'}{\partial v} \\
\frac{\partial g_3}{\partial d_v^{(1)}} &= \delta_v a'_u \frac{\partial \mathcal{I}'}{\partial u} + \delta_v (1 + b'_u) \frac{\partial \mathcal{I}'}{\partial v} \\
\frac{\partial g_3}{\partial d^{(2)}} &= \frac{\partial \mathcal{I}'}{\partial v} \\
\frac{\partial g_3}{\partial t_u} &= \frac{\partial \mathcal{I}'}{\partial y} \\
\frac{\partial g_3}{\partial t_v} &= \frac{\partial \mathcal{I}'}{\partial v} \\
\frac{\partial g_3}{\partial a'_u} &= \delta_u \frac{\partial \mathcal{I}'}{\partial u} \\
\frac{\partial g_3}{\partial a'_v} &= \left( \delta_u d_u^{(1)} + \delta_v (1 + d_v^{(1)}) \right) \frac{\partial \mathcal{I}'}{\partial u} \\
\frac{\partial g_3}{\partial b'_u} &= \delta_u \frac{\partial \mathcal{I}'}{\partial v} \\
\frac{\partial g_3}{\partial b'_v} &= \left( \delta_u d_u^{(1)} + \delta_v (1 + d_v^{(1)}) \right) \frac{\partial \mathcal{I}'}{\partial v}
\end{aligned}$$

Les différents gradients des niveaux de gris exprimés dans l'espace des images rectifiées sont déterminés à partir de ceux des images initiales par la méthode exposée dans l'annexe D.



## Annexe H

# Vecteur de rotation instantanée

Une expression minimale, sous la forme d'un vecteur  $\mathbf{n}$  de rotation instantanée, d'une matrice de rotation  $\mathbf{R} = [r_{ij}]_{i,j}$ , est la suivante [Gar96] :

$$\mathbf{n} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \theta \cdot \mathbf{m}$$

avec :

$$\mathbf{m} = \begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} = \frac{1}{\sqrt{4 - (r_{11} + r_{22} + r_{33} - 1)^2}} \begin{pmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{pmatrix}$$
$$\theta = \cos^{-1} \left( \frac{r_{11} + r_{22} + r_{33} - 1}{2} \right)$$

Réciproquement, nous pouvons définir une matrice de rotation sous une forme minimale en fonction de  $\mathbf{n}$  par :

$$\mathbf{R} = \begin{pmatrix} \cos \theta + \frac{n_x^2(1-\cos \theta)}{\theta^2} & \frac{n_x n_y(1-\cos \theta)}{\theta^2} - \frac{n_z \sin \theta}{\theta} & \frac{n_x n_z(1-\cos \theta)}{\theta^2} + \frac{n_y \sin \theta}{\theta} \\ \frac{n_x n_y(1-\cos \theta)}{\theta^2} + \frac{n_z \sin \theta}{\theta} & \cos \theta + \frac{n_y^2(1-\cos \theta)}{\theta^2} & \frac{n_y n_z(1-\cos \theta)}{\theta^2} - \frac{n_x \sin \theta}{\theta} \\ \frac{n_x n_z(1-\cos \theta)}{\theta^2} - \frac{n_y \sin \theta}{\theta} & \frac{n_y n_z(1-\cos \theta)}{\theta^2} + \frac{n_x \sin \theta}{\theta} & \cos \theta + \frac{n_z^2(1-\cos \theta)}{\theta^2} \end{pmatrix}$$

avec  $\theta = \|\mathbf{n}\|$ .

Nous aurons aussi besoin d'exprimer les matrices jacobiniennes de la matrice de rotation  $\mathbf{R}$  par rapport au vecteur de rotation instantanée  $\mathbf{n}$  dans l'annexe suivante :

$$\frac{\partial \mathbf{R}}{\partial n_k} = \begin{pmatrix} \frac{\partial r_{11}}{\partial n_k} & \frac{\partial r_{12}}{\partial n_k} & \frac{\partial r_{13}}{\partial n_k} \\ \frac{\partial r_{21}}{\partial n_k} & \frac{\partial r_{22}}{\partial n_k} & \frac{\partial r_{23}}{\partial n_k} \\ \frac{\partial r_{31}}{\partial n_k} & \frac{\partial r_{32}}{\partial n_k} & \frac{\partial r_{33}}{\partial n_k} \end{pmatrix} \quad (\text{H.1})$$

avec  $k = x, y$  ou  $z$ , et :

$$\begin{aligned} \frac{\partial r_{ij}}{\partial n_k} &= \frac{\partial r_{ij}}{\partial m_x} \frac{\partial m_x}{\partial n_k} + \frac{\partial r_{ij}}{\partial m_y} \frac{\partial m_y}{\partial n_k} + \frac{\partial r_{ij}}{\partial m_z} \frac{\partial m_z}{\partial n_k} + \frac{\partial r_{ij}}{\partial \theta} \frac{\partial \theta}{\partial n_k} \\ \frac{\partial \mathbf{R}}{\partial m_x} &= \begin{pmatrix} 2m_x(1 - \cos \theta) & m_y(1 - \cos \theta) & m_z(1 - \cos \theta) \\ m_y(1 - \cos \theta) & 0 & -\sin \theta \\ m_z(1 - \cos \theta) & \sin \theta & 0 \end{pmatrix} \\ \frac{\partial \mathbf{R}}{\partial m_y} &= \begin{pmatrix} 0 & m_x(1 - \cos \theta) & \sin \theta \\ m_x(1 - \cos \theta) & 2m_y(1 - \cos \theta) & m_z(1 - \cos \theta) \\ -\sin \theta & m_z(1 - \cos \theta) & 0 \end{pmatrix} \\ \frac{\partial \mathbf{R}}{\partial m_z} &= \begin{pmatrix} 0 & -\sin \theta & m_x(1 - \cos \theta) \\ \sin \theta & 0 & m_y(1 - \cos \theta) \\ m_x(1 - \cos \theta) & m_y(1 - \cos \theta) & 2m_z(1 - \cos \theta) \end{pmatrix} \\ \frac{\partial \mathbf{R}}{\partial \theta} &= \begin{pmatrix} (m_x^2 - 1) \sin \theta & m_x m_y \sin \theta - m_z \cos \theta & m_x m_z \sin \theta + m_y \cos \theta \\ m_x m_y \sin \theta + m_z \cos \theta & (m_y^2 - 1) \sin \theta & m_y m_z \sin \theta - m_x \cos \theta \\ m_x m_z \sin \theta - m_y \cos \theta & m_y m_z \sin \theta + m_x \cos \theta & (m_z^2 - 1) \sin \theta \end{pmatrix} \\ \frac{\partial \mathbf{m}}{\partial \mathbf{n}} &= \frac{1}{\theta} \begin{pmatrix} 1 - m_x^2 & -m_y m_x & -m_z m_x \\ -m_x m_y & 1 - m_y^2 & -m_z m_y \\ -m_x m_z & m_y m_z & 1 - m_z^2 \end{pmatrix} \\ \frac{\partial \theta}{\partial \mathbf{n}} &= (m_x \quad m_y \quad m_z) \end{aligned}$$

# Bibliographie générale

- [AG92] P. ASCHWANDEN et W. GUGGENBÜHL. Experimental results from a comparative study on correlation-type registration algorithms. Dans *Robust Computer Vision*, 1992.
- [Aya89] Nicolas AYACHE. *Vision stéréoscopique et perception multisensorielle*. Collection Science Informatique. InterEditions, 1989.
- [BD94] T. BLASZKA et R. DERICHE. Recovering and Characterizing Image Features Using an Efficient Model Based Approach. Rapport de recherche 2422, INRIA, novembre 1994.
- [BDL95] F. BRÉMAND, J. C. DUPRÉ, et A. LAGARDE. Mesure de déformations sans contact par analyse d'images. Dans *Photomécanique'95*, ENS Cachan (France), mars 1995.
- [Bey92] Horst A. BEYER. Accurate Calibration of CCD Cameras. Dans *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'92)*, 1992.
- [BF82] S. T. BARNARD et M. A. FISCHLER. Computational Stereo. *Computing Surveys*, 14(4):553–572, décembre 1982.
- [Bla97] Thierry BLASZKA. *Approches par modèles en vision précoce*. Thèse de doctorat, Université de Nice Sophia-Antipolis (France), février 1997.
- [BMSP89] H. BRUCK, S. MCNEILL, M. SUTTON, et W. PETERS. Digital Image Correlation Using Newton-Raphson Method of Partial Differential Correction. Dans *Experimental Mechanics*, volume 29, pp. 261–267, avril 1989.
- [Bra95] Pascal BRAND. *Reconstruction tridimensionnelle à partir d'une caméra en mouvement : de l'influence de la précision*. Thèse de doctorat, Université Claude Bernard – Lyon I (France), octobre 1995.
- [Cai01] Emmanuel CAILLEUX. *Microstructure et comportement thermomécanique d'un béton réfractaire renforcé par des fibres métalliques*. Thèse de doctorat, École des Mines de Paris (France), mai 2001.
- [CCLB99] T. CUTARD, E. CAILLEUX, P. LOURS, et G. BERNHART. Structural and Mechanical Properties of a Refractory Concrete for Superplastic Forming Tools. *Industrial Ceramics*, 19(2):100–102, 1999.
- [CMB99] P. CLERC, F. MORESTIN, et M. BRUNET. Measurement of Deformation Fields Using Numerical Images Correlation. Dans *International Conference on Advances in Materials and Processing Technologies (AMPT'99)*, pp. 1889–1897, Dublin (Irlande), août 1999.

- [CRSP86] T. C. CHU, W. F. RANSON, M. A. SUTTON, et W. H. PETERS. Application of Digital-Image-Correlation Techniques to Experimental Mechanics. Dans *Experimental Mechanics*, volume 25, pp. 232–244, 1986.
- [DA89] U. R. DHOND et J. K. AGGARWAL. Structure from Stereo — A Review. *IEEE Transactions on Systems, Man and Cybernetics*, 19(16):1489–1510, novembre 1989.
- [Dev97] Frédéric DEVERNAY. *Vision stéréoscopique et propriétés différentielles des surfaces*. Thèse de doctorat, École Polytechnique (France), février 1997.
- [DF94] F. DEVERNAY et O. FAUGERAS. Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models. Rapport de recherche 2304, INRIA, juillet 1994.
- [Dol96] Jürgen DOLD. Influence of large targets on the results of photogrammetric bundle adjustment. *International Archives of Photogrammetry and Remote Sensing*, 31(B5), 1996.
- [Dou00] Pascal DOUMALIN. *Microextensométrie locale par corrélation d'images numériques. Application aux études micromécaniques par microscopie électronique à balayage*. Thèse de doctorat, École Polytechnique (France), juin 2000.
- [Fau93] Olivier FAUGERAS. *Three-Dimensional Computer Vision : A Geometric Viewpoint*. The MIT Press, 1993. ISBN 0-262-06158-9.
- [FHM<sup>+</sup>93] O. FAUGERAS, B. HOTZ, H. MATHIEU, T. VIÉVILLE, Z. ZHANG, P. FUA, E. THÉRON, L. MOLL, G. BERRY, J. VUILLEMIN, P. BERTIN, et C. PROY. Real time correlation-based stereo : algorithm, implementations and applications. Rapport de recherche 2013, INRIA, août 1993.
- [FT87] Olivier FAUGERAS et Giorgio TOSCANI. Camera Calibration for 3D Computer Vision. Dans *International Workshop on Machine Vision and Machine Intelligence*, pp. 240–247, Tokyo (Japon), février 1987.
- [Fua91] Pascal FUA. Combining Stereo and Monocular Information to Compute Dense Depth Maps that Preserve Depth Discontinuities. Dans *International Joint Conference on Artificial Intelligence (IJCAI'91)*, Sydney (Australie), août 1991.
- [Fus00] Andrea FUSIELLO. Uncalibrated Euclidean Reconstruction : A Review. *Image and Vision Computing*, 18:555–563, septembre 2000.
- [Gar96] Vincent GARRIC. *Vision pour la robotique de manipulation : calibration, localisation et saisie d'objets*. Thèse de doctorat, Institut National des Sciences Appliquées de Toulouse (France), mai 1996.
- [GDO<sup>+</sup>01] D. GARCIA, B. DAVID, J.-J. ORTEU, T. CUTARD, et E. CAILLEUX. Application de la stéréovision à l'étude du comportement de bétons renforcés de fibres métalliques. Dans *Photomécanique'2001*, pp. 24–26, Futuroscope, Poitiers (France), avril 2001.
- [GH99] K. GALANULIS et A. HOFMANN. Determination of Forming Limit Diagrams Using an Optical Measurement System. Dans *7th International Conference on Sheet Metal (SheMet'99)*, pp. 245–252, Erlangen (Allemagne), septembre 1999.
- [GMN<sup>+</sup>98] B. GALVIN, B. McCANE, K. NOVINS, D. MASON, et S. MILLS. Recovering Motion Fields : An Evaluation of Eight Optical Flow Algorithms. Dans *British Machine Vision Conference (BMVC'98)*, septembre 1998.

- [GO99] D. GARCIA et J.-J. ORTEU. 3D Deformation Measurement using Stereo-Correlation applied to the Forming of Metal or Elastomer Sheets. Dans *International Workshop on Video-Controlled Materials Testing and In-Situ Microstructural Characterization*, Nancy (France), novembre 1999.
- [GO01] D. GARCIA et J.-J. ORTEU. 3D Deformation Measurement using Stereocorrelation applied to Experimental Mechanics. Dans *10th FIG International Symposium on Deformation Measurements*, pp. 19–22, Orange, Californie (USA), mars 2001.
- [GOD00] D. GARCIA, J.-J. ORTEU, et M. DEVY. Accurate Calibration of a Stereovision Sensor : Comparison of Different Approaches. Dans *5th Workshop on Vision, Modeling, and Visualization (VMV'2000)*, pp. 25–32, Saarbrücken (Allemagne), novembre 2000.
- [GRMT98] M. GASPÉRINI, A. RAZAKAINAVO, M. MOREL, et C. TEODOSIU. Mesure par microgrilles des déformations en étirage plan dans l'épaisseur de tôles minces d'alliage d'aluminium. Dans *Photomécanique'98*, Marne la Vallée (France), avril 1998.
- [Gro93] Patrick GROS. *Outils géométriques pour la modélisation et la reconnaissance d'objets polyédriques*. Thèse de doctorat, Institut National Polytechnique de Grenoble (France), juillet 1993.
- [Har93] Richard I. HARTLEY. Computing Matched-epipolar Projections. Dans *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'93)*, pp. 549–555, New York (USA), juin 1993.
- [Har95] Richard I. HARTLEY. In defence of the 8-point algorithm. Dans *IEEE International Conference on Computer Vision (ICCV'95)*, pp. 1064–1070, Cambridge (USA), juin 1995.
- [HEG01] J. M. HIVER, J. ETIENNE, et C. G'SELL. Mesure de l'endommagement plastique dans une striction par un système d'analyse d'image. Dans *Photomécanique'2001*, Futuroscope, Poitiers (France), avril 2001.
- [HM95] R. HORAUD et O. MONGA. *Vision par ordinateur : outils fondamentaux*. Hermès, 1995. Deuxième édition.
- [HS81] K. P. HORN et B. G. SCHUNCK. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [HS97] R. I. HARTLEY et P. STURM. Triangulation. Dans *Computer Vision and Image Understanding (CVIU'97)*, volume 68 :2, pp. 146–157, novembre 1997.
- [KJC90] Z. L. KHAN-JETTER et T. C. CHU. Three-dimensional displacement measurements using digital image correlation and photogrammic analysis. *Experimental Mechanics*, 30(1):10–16, 1990.
- [KW97] K. KRAUS et P. WALDHÄUSL. *Manuel de photogrammétrie — principes et procédés fondamentaux*. Hermès, 1997. ISBN 2-8660-1656-4.
- [LA98] Fazilay LARABA-ABBES. *Étude des comportements hyperélastique et viscohyperélastique de deux élastomères de type NR et PDMS par extensométrie optique bidimensionnelle*. Thèse de doctorat, École Centrale de Paris (France), février 1998.

- [LAE00] N. LANGERAK, E. ATZEMA, et J. ESSING. Strain measurements with the PHAST™ system. Dans *International Deep Drawing Research Group (ID-DRG'2000) — Meeting of the Working Groups*, Ann Arbor, Michigan (USA), 2000.
- [Lan97] Zhong-Dan LAN. *Méthodes robustes en vision : application aux appariements visuels*. Thèse de doctorat, Institut National Polytechnique de Grenoble (France), mai 1997.
- [LCS94] P. F. LUO, Y. J. CHAO, et M. A. SUTTON. Application of Stereo Vision to 3-D Deformation Analysis in Fracture Mechanics. *Optical Engineering*, 33:981, 1994.
- [LCSP93] P. F. LUO, Y. J. CHAO, M. A. SUTTON, et W. H. PETERS. Accurate measurement of three-dimensional deformable and rigid bodies using computer vision. *Experimental Mechanics*, 33(2):123–132, 1993.
- [LD99] Manolis I.A. LOURAKIS et Rachid DERICHE. Camera self-calibration using the singular value decomposition of the fundamental matrix : From point correspondences to 3D measurements. Rapport de recherche 3748, INRIA, août 1999.
- [LD00] Manolis I.A. LOURAKIS et Rachid DERICHE. Camera self-calibration using the Kruppa equations and the SVD of the fundamental matrix : The case of varying intrinsic parameters. Rapport de recherche 3911, INRIA, 2000.
- [LDFP93] Q.-T. LUONG, R. DERICHE, O. FAUGERAS, et T. PAPADOPOULO. On determining the fundamental matrix : analysis of different methods and experimental results. Dans *Israeli Conf. on Artificial Intelligence and Computer Vision*, Tel-Aviv (Israel), 1993.
- [LEF97] A. LORUSSO, D. W. EGGERT, et R. B. FISHER. A Comparison of Four Algorithms for Estimating 3-D Rigid Transformations. Dans *Machine Vision and Applications*, volume 9, pp. 272–290, 1997.
- [LG95] P. LASSERRE et P. GRANDJEAN. Stereovision improvements. Dans *7th International Conference on Advanced Robotics (ICAR'95)*, St Feliu de Guixols (Espagne), septembre 1995.
- [LGA92a] J. L. LE GOËR et J. AVRIL. Capteurs à jauges extensométriques. Dans *Mesures et contrôle — Grandeurs mécaniques*, volume R7, p. R1860. Techniques de l'Ingénieur, 1992.
- [LGA92b] J. L. LE GOËR et J. AVRIL. Extensométrie. Dans *Mesures et contrôle — Grandeurs mécaniques*, volume R7, p. R1850. Techniques de l'Ingénieur, 1992.
- [LH81] H. C. LONGUET-HIGGINS. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, 293:133–135, 1981.
- [Luo92] Quang-Tuan LUONG. *Matrice fondamentale et calibration visuelle sur l'environnement — Vers une plus grande autonomie des systèmes robotiques*. Thèse de doctorat, Université de Paris-Sud, Centre d'Orsay (France), décembre 1992.
- [LVD99] J.M. LAVEST, M. VIALA, et M. DHOME. Quelle précision pour une mire d'étalement ? *Traitement du Signal*, 16(3), 1999.
- [LZ99] C. LOOP et Z. ZHANG. Computing Rectifying Homographies for Stereo Vision. Rapport technique MSR-TR-99-21, Microsoft Research, avril 1999.

- [Mar63] Donald W. MARQUARDT. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11:431–441, 1963.
- [MBB95] F. MORESTIN, P. BOGAERT, et M. BOIVIN. Mesure de déformations par imagerie : détermination des courbes limites de formage. Dans *Photomécanique'95*, ENS Cachan (France), mars 1995.
- [MBL93] D. W. MANTHEY, R. M. BASSETTE, et D. LEE. Comparison of Different Surface Strain Measurement Techniques Used for Stamped Sheet Metal Parts. Dans *International Body Engineering Conference : Body Assembly & Manufacturing*, pp. 106–111, Detroit, MI (USA), septembre 1993.
- [McD81] M. J. MCDONNELL. Box-Filtering Techniques. *Computer Graphics and Image Processing*, 17:65–70, 1981.
- [MD92] Z. MARCINIAK et J. L. DUNCAN. *The Mechanics of Sheet Metal Forming*. Edward Arnold, 1992.
- [Mon98] Jérôme MONTEIL. *Traitement et analyse par une technique adaptée de flux optique du champ de déformation de matériaux à partir de séquence d'images de microscopie*. Thèse de doctorat, Université Denis Diderot — Paris VII (France), mars 1998.
- [NGIG95] M. NÉMOZ-GAILLARD, P. IENNY, et J. GERMAIN. Extensométrie bidimensionnelle par analyse d'images. Dans *Photomécanique'95*, ENS Cachan (France), mars 1995.
- [OAHL98] J.J. ORTEU, H. AL-HADDAD, et S. LACROIX. Mesure de déformations 3D par stéréocorrélation. Dans *Photomécanique'98*, Marne la Vallée (France), avril 1998.
- [OGD97] J.J. ORTEU, V. GARRIC, et M. DEVY. Camera calibration for 3D reconstruction : application to the measure of 3D deformations on sheet metal parts. Dans *European Symposium on Lasers, Optics and Vision in Manufacturing*, Munich (Allemagne), juin 1997.
- [Ort91] Jean-José ORTEU. *Application de la vision par ordinateur à l'automatisation de l'abattage dans les mines*. Thèse de doctorat, Université Paul Sabatier de Toulouse (France), novembre 1991.
- [Per85] Ken PERLIN. An Image Synthesizer. Dans *SIGGRAPH'85*, pp. 287–296, 1985.
- [Peu94] Bernard PEUCHOT. Utilisation de détecteurs subpixels dans la modélisation d'une caméra. Dans *Actes du 9ème Congrès AFCET (RFIA'94)*, pp. 691–695, Paris (France), janvier 1994.
- [Pho95] Actes du colloque Photomécanique'95. ENS Cachan (France), mars 1995.
- [Pho98] Actes du colloque Photomécanique'98. Marne la Vallée (France), avril 1998.
- [Pho01] Actes du colloque Photomécanique'2001. Futuroscope, Poitiers (France), avril 2001.
- [Pol00] Humbert POLICELLA. *Jauges extensométriques : l'expérience au service de la qualité*. GAMAC, édition TEC & DOC, juin 2000. ISBN 2-84107-444-7.
- [Pér94] José-Philippe PÉREZ. *Optique géométrique et ondulatoire*. Masson, édition quatrième, 1994. ISBN 2-225-84270-1.

- [PTVF92] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, et B. P. FLANNERY. *Numerical Recipes in C — The Art of Scientific Computing*. Cambridge University Press, édition 2nd, 1992.
- [RA94] V. RODIN et A. AYACHE. Stéréovision axiale : modélisation et calibrage du système de prise de vue, reconstruction 3D d'objets naturels. *Traitement du Signal (TSI)*, 11(5):373–391, novembre 1994.
- [RAS93] O. RAVN, N. A. ANDERSEN, et A. T. SORENSEN. Auto-calibration in Automation Systems using Vision. Dans *3rd International Symposium on Experimental Robotics (ISER'93)*, pp. 206–218, (Japon), 1993.
- [Rog96] William D. ROGATTO. *The Infrared and Electro-Optical Systems Handbook — second edition*, volume 3 : Electro-Optical Components. SPIE Press, 1996. ISBN 0-8194-1072-1.
- [RSLM<sup>+</sup>01] N. REUGE, F. M. SCHMIDT, Y. LE MAOULT, M. RACHIK, et F. ABBÉ. Elastomer Biaxial Characterization Using Bubble Inflation Technique. I : Experimental Investigations. *Polymer Engineering and Science*, 41(3):522, mars 2001.
- [Rut63] H. RUTISHAUSER. Algorithm 150 : syminv2. Dans *Communications of the ACM*, volume 6, pp. 67–68. ACM Press, New York, NY (USA), février 1963.
- [SBS00] H. SCHREIER, J. BRAASCH, et M. SUTTON. Systematic errors in digital image correlation caused by intensity interpolation. *Optical Engineering*, 39(11):2915–2921, novembre 2000.
- [SDC86] R. SOWERBY, J. L. DUNCAN, et E. CHU. The Modelling of Sheet Metal Stamping. *International Journal of Mechanical Science*, 28(7):415–430, 1986.
- [She96] Jonathan Richard SHEWCHUK. Triangle : Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. Dans *Applied Computational Geometry : Towards Geometric Engineering*, volume 1148 de *Lecture Notes in Computer Science*, pp. 203–222. Springer-Verlag, mai 1996.
- [SMHC00] M. A. SUTTON, S. R. MCNEILL, J. D. HELM, et Y. J. CHAO. Advances in Two-Dimensional and Three-Dimensional Computer Vision. *Photomechanics, Topics Applied Physics*, 77:323–372, 2000.
- [SMHS00] M. A. SUTTON, S. R. MCNEILL, J. D. HELM, et H. W. SCHREIER. Computer vision applied to shape and deformation measurement. Dans *International Conference on Trends in Optical Nondestructive Testing and Inspection, Elsevier Science*, pp. 571–589, Lugano (Suisse), mai 2000.
- [SMJB88] M. SUTTON, S. MCNEILL, J. JANG, et M. BABAI. Effects of subpixel image restoration on digital correlation error estimates. *Optical Engineering*, 27(10):870–877, octobre 1988.
- [SRLM<sup>+</sup>99] F. SCHMIDT, N. REUGE, Y. LE MAOULT, J.P. ARCENS, B. DAVID, et F. ABBÉ. Inflation technique optimisation for an elastomer biaxial characterisation. Dans *Polymer Processing Society - 15th annual meeting, 's-Hertogenbosch (Pays-Bas)*, juin 1999.
- [SS99] P. SYNNERGREN et M. SJÖDAHL. A stereoscopic digital speckle photography system for 3-D displacement field measurements. *Optics and Lasers in Engineering*, 31:425–443, 1999.

- [SWP<sup>+</sup>83] M. SUTTON, W. J. WOLTERS, W. H. PETERS, W. F. RANSON, et S. R. MCNEILL. Determination of Displacements Using an Improved Digital Correlation Method. *Image and Vision Computing*, 21:133–139, 1983.
- [Tar98] Fabrice TARDIF. *Analyse des mécanismes d'endommagement dans les composites 3D C/C — Mesures optiques des champs de déformation par corrélation d'images*. Thèse de doctorat, Université Bordeaux I (France), novembre 1998.
- [TBBD01] R. TIE BI, N. BRETAGNE, et J. C. DUPRÉ. Mesures par suivi de marqueurs et thermographie infrarouge de paramètres thermomécaniques. Dans *Photomécanique'2001*, Futuroscope, Poitiers (France), avril 2001.
- [Thé96] Jan THÉSING. *Mesure par corrélation des déplacements et des rotations à partir de figures granulaires*. Mémoire de fin d'études, LMT Cachan (France), 1996.
- [TM97] Siham TOUCHAL-MGUIL. *Une technique de corrélation directe d'images numériques : application à la détermination de courbes limites de formage et proposition d'un critère de striction*. Thèse de doctorat, Institut National des Sciences Appliquées de Lyon (France), juillet 1997.
- [Tos87] Giorgio TOSCANI. *Système de calibration optique et perception du mouvement en vision artificielle*. Thèse de doctorat, Université de Paris-Sud, Centre d'Orsay (France), 1987.
- [Tsa87] R.Y. TSAI. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. *IEEE Journal of Robotics and Automation*, RA-3(4), août 1987.
- [UAE93a] M. UNSER, A. ALDROUBI, et M. EDEN. B-spline processing : part I — theory. *IEEE Transactions on Signal Processing*, 41(2):821–833, 1993.
- [UAE93b] M. UNSER, A. ALDROUBI, et M. EDEN. B-spline processing : part II — efficient design and applications. *IEEE Transactions on Signal Processing*, 41(2):834–847, 1993.
- [Uns99] Michael UNSER. Splines, A perfect Fit for Signal and Image Processing. *IEEE Signal Processing Magazine*, 16(6):22–28, novembre 1999.
- [VL89] J. H. VOGEL et D. LEE. An Automated Two-View Method For Determining Strain Distributions on Deformed Surfaces. Dans *Journal of Material Shaping Technology*, volume 6, 1989.
- [Wat99] Bertrand WATTRISE. *Étude cinématique des phénomènes de localisation dans des aciers par intercorrélation d'image*. Thèse de doctorat, Université Montpellier II (France), février 1999.
- [WCH92] J. WENG, P. COHEN, et M. HERNIOU. Camera Calibration with Distorsion Models and Accuracy Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI'92)*, 14(10):965–980, octobre 1992.
- [Yan92] Ying YANG. *Mesure du champ de déplacement des matériaux sous charge mécanique par analyse des images*. Thèse de doctorat, École Nationale des Ponts et Chaussées (France), septembre 1992.
- [Zel96] Cyril ZELLER. *Calibration projective, affine et euclidienne en vision par ordinateur et application à la perception tridimensionnelle*. Thèse de doctorat, École Polytechnique (France), février 1996.

- [Zha93a] Zhengyou ZHANG. Le problème de la mise en correspondance : l'état de l'art. rapport de recherche 2146, INRIA, décembre 1993.
- [Zha93b] ChangSheng ZHAO. *Reconstruction de surfaces tridimensionnelles en vision par ordinateur*. Thèse de doctorat, Institut National Polytechnique de Grenoble (France), décembre 1993.
- [Zha98a] Zhengyou ZHANG. A Flexible New Technique for Camera Calibration. Rapport technique MSR-TR-98-71, Microsoft Research, décembre 1998. Mis à jour en mars 1999.
- [Zha98b] Zhengyou ZHANG. Understanding the Relationship Between the Optimization Criteria in Two-View Motion Analysis. Dans *IEEE International Conference on Computer Vision (ICCV'98)*, pp. 231–236, Bombay (India), janvier 1998.



## Mesure de formes et de champs de déplacements tridimensionnels par stéréo-corrélation d'images

De nombreux domaines concernant le comportement mécanique des matériaux posent le problème de la mesure des déplacements ou des déformations. Pour ce type de mesure, les méthodes optiques se sont largement imposées en raison de leur caractère non intrusif, de leur grande résolution spatiale, de leur sensibilité élevée, de l'importance du champ examiné à tout instant et des progrès de l'informatique qui permet le traitement automatique d'un grand volume d'information. Dans ce contexte, nous avons développé un système de mesure de formes 3D ou de champs de déplacements 3D par stéréovision (en particulier par stéréo-corrélation). Cette technique permet : (1) la mesure de la forme 3D d'un objet à partir d'une simple paire d'images stéréoscopiques de l'objet, (2) la mesure de champs de déplacements 3D à partir d'au moins 2 paires d'images correspondant à différents instants de déformation de l'objet (en général analyse d'une séquence de paires d'images acquises en cours de déformation). Les points développés dans la thèse sont : le calibrage fort d'une caméra ou d'un capteur de vision stéréoscopique, la reconstruction 3D par stéréovision (en particulier par stéréo-corrélation), la mesure de champs de déplacements 3D à partir du couplage de la stéréo-corrélation et du suivi de pixels dans une séquence d'images par corrélation. Compte tenu de la finalité métrologique de ces travaux, nous accordons une attention toute particulière à la précision des méthodes mises en oeuvre (qualité du calibrage, qualité de la mise en correspondance des images, corrélation subpixel,...). Ces travaux ont été appliqués à l'emboutissage de tôles minces (mesure de formes 3D d'emboutis et mesure de champs de déformations à la surface d'emboutis 3D), à la mise en forme de polymères (mesure de champs de déplacements 3D sur des membranes en élastomère soufflées), et à l'étude du comportement mécanique de bétons réfractaires renforcés de fibres métalliques (mesure de champs de déplacements 3D lors d'essais de traction).

**Mots-clés** : calibrage de caméras, stéréovision, suivi de pixels, appariement d'images par corrélation, métrologie 3D sans contact, photomécanique.

## Measurement of 3-D Shapes or 3-D Displacement Fields Using a Stereo-correlation Based Method

Many applications dealing with the mechanical behaviour of materials require the measurement of displacement fields or deformation fields. In order to attain these measurements, optical methods have become unavoidable due to their non-intrusive approach, their high spatial resolution, their high sensitivity, the large size of their examined field, and the increasing power of the computers that now allow the processing of huge quantities of data. In this context we have developed a system for measuring 3-D shapes or 3-D displacement fields using a stereo-vision approach (by stereo-correlation). This method allows: (1) the measurement of the 3-D object shape using a single pair of stereoscopic images of this object and (2) the measure of the 3-D displacement field using at least two pairs of stereoscopic images of an object corresponding to two states of its deformation (or the processing of a sequence of pairs of images acquired during the deformation). The main topics developed in this thesis are: the strong calibration of a camera or a stereo-vision sensor, the 3-D reconstruction by stereo-vision, the measurement of the 3-D displacement field using a combined stereo-correlation and tracking approach in a sequence of images. Due to the metrology aspect of this research we have paid particular attention to the accuracy of the underlying methods (quality of the calibration, quality of the image matching, subpixel correlation,...). This work has been successfully applied to the sheet metal forming (measurement of 3-D shapes and 3-D displacement fields of the stamped parts), to polymer forming (measurement of the 3-D displacement field on a blown elastomere membrane), and to the study of the mechanical behaviour of a refractory ceramic reinforced with metallic fibers (measurement of 3-D displacement fields during a tensile test).

**Keywords** : camera calibration, stereovision, correlation based image matching, non-contacting 3-D metrology, computer vision based experimental mechanics.