HAL
open science

Mesure de vibrations par vision 3D

Thomas Durand-Texte

To cite this version:

Thomas Durand-Texte. Mesure de vibrations par vision 3D. Acoustique [physics.class-ph]. Le Mans Université, 2019. Français. NNT : 2019LEMA1003 . tel-02316314

HAL Id: tel-02316314

https://theses.hal.science/tel-02316314v1

Submitted on 15 Oct 2019

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

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

THESE DE DOCTORAT DE

LE MANS UNIVERSITE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 602
Sciences pour l'Ingénieur
Spécialité : Acoustique

Par
Thomas DURAND-TEXTE

« Mesure de vibrations par vision 3D »

Thèse présentée et soutenue à Le Mans, le 11 janvier 2019
Unité de recherche : Laboratoire d'Acoustique de L'Université du Mans
Thèse N° : 2019LEMA1003

Rapporteurs avant soutenance :

Jean-Luc Dion	Professeur, SUPMECA
François Hild	Directeur de Recherche, CNRS

Composition du Jury :

Président :	Jean-José Orteu	Professeur, École des Mines d'Albi
Examinateurs :	Jean-Luc Dion	Professeur, SUPMECA
	François Hild	Directeur de Recherche, CNRS
	Fabrice Morestin	Professeur, INSA de Lyon
Dir. de thèse :	Manuel Melon	Professeur, Le Mans Université
Co-Encadrant :	Élisabeth Simonetto	Maître de Conférences, ESGT
Co-Encadrant :	Marie-Hélène Moulet	Ingénieure de Recherche, CTTM

Table des matières

Remerciements **vii**

Avant propos **ix**

Introduction générale **1**

I Contexte, méthodes et outils **7**

1 Vibrations **9**
1.1 Introduction . . . 9
1.2 Systèmes vibrants . . . 10
1.2.1 Système masse-ressort . . . 10
1.2.2 Vibrations d'une plaque rectangulaire . . . 11
1.2.3 Vibrations d'un haut-parleur . . . 14
1.2.4 Ordres de grandeur à mesurer . . . 17
1.3 Outils utilisés pour le traitement . . . 18
1.3.1 Outils de traitement du signal . . . 18
1.3.2 Outils d'analyse modale . . . 19
1.4 Conclusion . . . 21

2 Etat des lieux des méthodes de mesure de vibrations **23**
2.1 Introduction . . . 23
2.2 Techniques ponctuelles de référence . . . 24
2.3 Méthodes d'interférométrie . . . 24
2.4 Méthodes acoustiques . . . 25
2.5 Bref aperçu d'autres approches . . . 25
2.5.1 Déflectométrie . . . 25
2.5.2 Projection de franges . . . 25
2.5.3 Vibromètre laser à balayage continu . . . 26
2.5.4 Méthode RIFF . . . 26
2.6 Photogrammétrie et vision 3D . . . 26
2.6.1 Un peu d'histoire... . . . 26
2.6.2 Quelques applications . . . 27
2.6.3 Avantages et inconvénients . . . 27
2.7 Conclusion . . . 28

3 Stéréo-vision **29**
3.1 Introduction . . . 29
3.2 Modèle de projection perspective : le modèle sténopé . . . 30
3.2.1 Image sur une caméra . . . 30
3.2.2 Modèle de caméra et repère caméra associé . . . 31
3.2.3 Calibrage d'une caméra . . . 34

3.2.4 Distortions optiques . . . 36
3.2.5 Repère lié à la mire de calibrage . . . 38
3.3 Principe du capteur stéréoscopique . . . 38
3.3.1 Lien géométrique entre deux repères de l'espace . . . 38
3.3.2 Calibrage d'un capteur stéréoscopique . . . 41
3.3.3 Triangulation de la position 3D d'un point dans l'espace . . . 43
3.3.4 Passage vers le repère objet . . . 46
3.4 Conclusion . . . 48

4 Traitement d'images 49
4.1 Introduction . . . 49
4.2 Filtrage convolutif . . . 50
4.3 Appariement des points . . . 51
4.4 Détection de cibles circulaires . . . 52
4.4.1 Centroïde du niveau de gris ou moyenne pondérée . . . 52
4.4.2 Détecteur de blobs de la bibliothèque OpenCV . . . 53
4.4.3 Coniques et ellipses . . . 54
4.4.4 Estimateur d'ellipses de Fitzgibbon . . . 55
4.4.5 Estimateur d'ellipses duales de Ouellet . . . 57
4.4.6 Détection initiale de cibles . . . 61
4.4.7 Tri et appariement initial . . . 61
4.5 Suivi de cibles . . . 63
4.6 Motif aléatoire . . . 65
4.6.1 Outils de corrélation d'images (DIC) . . . 66
4.6.2 Appariement initial - mesure de forme . . . 70
4.7 Suivi de subsets . . . 73
4.8 Conclusion . . . 74

II Montages, méthodes et applications 77

5 Protocole de mesure 79
5.1 Introduction . . . 79
5.2 Caméras ultra-rapides . . . 79
5.3 Mire de calibrage . . . 80
5.4 Protocole de mesure . . . 82
5.5 Traitements et logiciels . . . 83
5.6 Conclusion . . . 85

6 Pseudo stéréo-vision pour la mesure de vibrations 87
6.1 Introduction . . . 87
6.2 Système pseudo stéréoscopique avec jeu de quatre miroirs . . . 88
6.2.1 État des lieux des utilisations du montage . . . 88
6.2.2 Principe et dimensionnement . . . 89
6.2.3 Calibrage . . . 92
6.3 Étude de sensibilité . . . 93
6.3.1 Approches et critères d'évaluation . . . 93
6.3.2 Étude de sensibilité liée à la détection . . . 94
6.3.3 Étude de sensibilité liée à l'agencement du système pseudo stéréoscopique . . . 97
6.4 Application à une plaque . . . 100
6.4.1 Matériel utilisé . . . 100
6.4.2 Acquisition des images, calibrage et bruit de fond . . . 102
6.4.3 Mesure de vibrations - évaluation de la précision . . . 102
6.4.4 En bref... . . . 105
6.5 Application à un haut-parleur . . . 107

6.5.1 Description du montage . . . 107
6.5.2 Données de calibrage . . . 108
6.5.3 Mesure de vibrations - évaluation de la précision . . . 109
6.5.4 Bilan . . . 115
6.6 Conclusion . . . 116

7 Mesure par corrélation d'images 119
7.1 Introduction . . . 119
7.2 Méthode IRIs . . . 120
7.2.1 Préambule . . . 120
7.2.2 Principe général . . . 122
7.3 Validation expérimentale . . . 125
7.4 Validation numérique de la méthode . . . 129
7.5 Non-planéité des miroirs . . . 130
7.6 Automatisation de l'appariement initial de la méthode IRIs . . . 131
7.7 Conclusion . . . 136

8 Montage pseudo stéréoscopique à deux miroirs 137
8.1 Introduction . . . 137
8.2 Principe . . . 138
8.3 Application à un haut-parleur . . . 138
8.3.1 Fonctions de réponse en fréquence . . . 138
8.3.2 Déformées modales opérationnelles . . . 139
8.4 Application à la mesure plein champ d'une plaque . . . 140
8.4.1 Montages . . . 141
8.4.2 Calibrage . . . 141
8.4.3 Fonctions de réponse en fréquence . . . 143
8.4.4 Déformées modales opérationnelles . . . 143
8.5 Conclusion . . . 146

9 Méthode de mesure de vibrations mono-axiales 147
9.1 Introduction . . . 147
9.2 Principe du montage sans miroir . . . 148
9.3 Application à un tambourin . . . 149
9.3.1 Montage et matériel . . . 149
9.3.2 Protocole . . . 149
9.3.3 Calibrage . . . 150
9.3.4 Mesure de forme . . . 150
9.3.5 Validation sur un point avec un laser . . . 151
9.3.6 Mesure plein champ . . . 152
9.3.7 En bref... . . . 153
9.4 Application à la mesure plein champ d'une plaque . . . 154
9.5 Application à la mesure plein champ d'un haut-parleur . . . 154
9.6 Conclusion . . . 157

10 Méthode à caméras asynchrones 159
10.1 Introduction . . . 159
10.2 Principe . . . 160
10.3 Application à une plaque . . . 162
10.3.1 Fonctions de réponse en fréquence . . . 163
10.3.2 Déformées modales opérationnelles . . . 163
10.3.3 En bref... . . . 163
10.4 Application industrielle : cas du capot d'une voiture . . . 165
10.4.1 Montage et protocole . . . 165
10.4.2 Déformées modales opérationnelles . . . 166

10.4.3 En bref... 166
10.5 Application industrielle : cas du test d'un haut-parleur 168
10.5.1 Montage et protocole 168
10.5.2 Fonctions de réponse en fréquence du haut-parleur avant dégradation 169
10.5.3 Déformées modales opérationnelles du haut-parleur avant dégradation 169
10.5.4 Résultats après dégradation du haut-parleur 172
10.5.5 En bref... 173
10.6 Conclusion 174

Conclusion générale **175**

Liste des images **183**

A Liste des publications **185**

B Acte de la 144e convention de l'AES (Audio Engineering Society) **187**

Remerciements

« *Tous les hommes pensent que le bonheur est au sommet de la montagne alors qu'il est dans l'art de la gravir.* »
Confucius

Lorsque l'on commence une thèse, c'est un peu (et même beaucoup) l'inconnu. Les commentaires fusent : "c'est une réalisation de soi", "tu verras, c'est un apprentissage fantastique", ou encore "ce seront les meilleures années de ta vie !". Pour ma part, cette étape a été un travail passionnant, une expérience très enrichissante et surtout une sacrée épreuve. Adepte des sports extrêmes depuis mon plus jeune âge, et inconditionnel de *snowboard*, j'ai eu la sensation de me retrouver sur une piste bien noire, sans la remontée en téléphérique. De quoi déclencher des pics d'adrénaline. Avec la conscience qu'il va falloir se dépasser et repousser ses limites pour y arriver. Et en bas de l'Everest, le sommet parait très, mais vraiment très, très, loin. Il n'y a qu'un mot d'ordre finalement : avancer. Malgré les doutes, les pauses nécessaires pour reprendre son souffle, la météo parfois fluctuante. Avancer, continuer, serrer les dents, rester focalisé, avancer, encore et encore, en espérant que l'on arrive bel et bien au sommet et en rêvant secrètement que la-haut, là où c'est encore inaccessible, le paysage soit aussi magnifique qu'on le prétend. Sur la route, deux camps de base, deux CST, pour relancer la montée lorsque l'oxygène se fait plus rare, vérifier que l'acclimatation se fait correctement, et surtout pour récupérer du ravitaillement. Et tout le long de l'ascension, il y a eu ces sherpas qui ont préparé le chemin, orienté et aiguillonné, ces partenaires qui ont accepté eux aussi de faire cette montée vers le sommet. Des guides, qui, sous un soleil incandescent mais aussi parfois au cœur des avalanches, ont manifesté leur présence et ont permis de tenir le cap. Bleu, bleu, bleu est le ciel... Et un jour, un matin, un soir, un moment : le temps se suspend. La dernière ligne, le dernier mot, le dernier signe... Tout y est. Un instant d'éternité où la réalisation fugace s'infiltre tout doucement dans la perception : on voit le sommet. Ca y est, on y est. Il reste encore quelques mètres, des relectures, des vérifications... mais l'ascension est finie. Un sentiment d'apaisement et de quiétude se répand. Les endorphines font leur travail. Il reste à savourer, lentement pour en profiter pleinement, avant de boucler l'expérience. Les sherpas sont là, souriants. Leur aide a été précieuse, qu'ils en soient remerciés.

Je tiens à remercier :

Manuel Melon, Elisabeth Simonetto, Marie-Hélène Moulet et Stéphane Durand pour leur encadrement tout au long de ces trois années, leur aide pendant les phases expérimentales et leurs relectures avisées. J'espère qu'ils trouvent le paysage aussi beau que moi au sommet. Un merci particulier à Manuel Melon pour son soutien au moment des avalanches.
Charles Pézerat et Jean-José Orteu pour leurs avis précieux et leurs encouragements dans le cadre du CST.
Pascal Picart pour la mesure des miroirs à Palaiseau, le prêt de matériel et pour avoir refait le monde pendant quelques mètres de cordée avec un mille-feuille dans le sac à dos.
François Hild et Jean-Luc Dion pour avoir accepté d'être les rapporteurs de ce manuscrit, pour leur

temps de relecture ainsi que leur participation au jury.
Fabrice Morestin et Jean-José Orteu pour leur participation au jury.
Antonin Novak pour les discussions sur les non-linéarités des haut-parleurs.
James Blondeau et Emmanuel Brasseur pour la gestion du matériel de mesure et des caméras.
Pierre-Arnaud Lecomte pour la gestion du matériel informatique.
Eric Labergerie pour son temps, les discussions et le prêt d'appareils photo.
Mathieu Secail-Geraud pour les mesures effectuées avec le vibromètre laser à balayage de l'ENSIM.
Les techniciens de Le Mans Université pour les différentes réalisations mécaniques.
L'IRT Jules Vernes et Pascal Audrain, pour le prêt de la caméra Photron.
Le CTTM pour l'accès à la salle semi-anéchoïque.
Les personnels de l'ESGT, du CTTM et du LAUM en général pour leur accueil et leur aide.

... Et tous ceux qui, à un niveau ou un autre, m'ont aidé et soutenu lors de cette ascension.

Des remerciements particuliers à « Le Mans Acoustique », la région « Pays de la Loire » et le « Fonds européen de développement économique et régional » pour avoir cru en ce projet et l'avoir financé.

Avant propos

« Les hommes se plaisent à penser qu'ils peuvent se débrouiller seuls, mais l'homme, le vrai, sait que rien ne vaut le soutien et les encouragements d'une bonne équipe. »
Tim Allen

Une des particularités de ce sujet de thèse est son côté "touche-à-tout". Il mobilise des compétences diverses, corrélées à des champs disciplinaires tout aussi variés. Prenez un grand chaudron, et remplissez-le à moitié d'une bonne base en métrologie. Ajoutez une bonne dose de vibrations et une autre d'acoustique (n'hésitez pas à être généreux). N'oubliez pas une tasse bien pleine de vision 3D, une louche de photogrammétrie et laissez frémir à petits bouillons. Lorsque le mélange commence à parfumer l'air ambiant, ajoutez en une seule fois le bol de traitement de signal et les cubes de programmation. Remuez sans arrêt. Au bout de trois années, la potion (magique ?) est prête. Servez sans attendre et dégustez entre connaisseurs.

Des connaisseurs, qui, nécessairement, appartiennent à des communautés différentes, de par leurs formations et domaines de recherche. Des spécialistes, qui, chacun, portent un regard différent sur le réel : chaque champ disciplinaire tend à construire ses propres outils d'analyse, ses propres normes et codes, ses propres *modus operandi* en liaison avec les finalités affichées et l'historique de la communauté qui les porte. Comme le rappelait Protagoras : « ce que l'homme appelle vérité, c'est toujours sa vérité, c'est-à-dire l'aspect sous lequel les choses lui apparaissent ». L'expérience de quelques congrès au sein de différents groupes a très vite mis en exergue le cloisonnement de certaines niches, et la potentielle non-communication entre chercheurs de domaines pourtant connexes, mais perçus comme à des années-lumières les uns des autres. Ces rencontres avec les différentes communautés m'ont fait prendre conscience très rapidement de la force de notre équipe, pluri-disciplinaire certes, et justement multi-compétente. Finalement, en citant le Mahatma Gandhi, « les vérités différentes en apparence sont comme d'innombrables feuilles qui paraissent différentes et qui sont sur le même arbre ». Le prisme de chacun d'entre nous altère le réel de façon différente mais cette réalité, scientifique qui plus est, reste nécessairement unique et invariante.

Comment, dès lors, réussir à faire converger les perceptions de cette même réalité afin de tendre vers la vérité scientifique que ce sujet de thèse doit permettre de toucher du doigt ? Comment, à ce point de confluence, harmoniser les voix de ce groupe de chercheurs, ouverts et bienveillants, grands spécialistes de leurs domaines, qui doivent s'insérer dans un sujet aux multiples facettes, mobilisant des savoir-faire éclectiques ? Descartes m'a mis sur la voie : « pour atteindre la vérité, il faut une fois dans sa vie se défaire de toutes les opinions que l'on a reçues et reconstruire de nouveau, dès le fondement, tout le système de ses connaissances ».

Ce manuscrit est ainsi l'aboutissement de cette reconstruction. Les outils à disposition au sein des différentes communautés ont été mobilisés de façon parfois inhabituelle, et appliqués au contexte transversal qui cadrait cette recherche. L'efficacité et le pragmatisme ont été les mots d'ordre, aux dépens du purisme et de l'enfermement. Ce travail s'est positionné dans l'ouverture, à l'intersection des domaines qui le sous-tendent, en cherchant le plus possible la simplification qui favorise la concorde.

Introduction générale

« *Définir des objectifs est la première étape dans la transformation de l'invisible dans le visible.* »
Tony Robbins

Au commencement était... la vibration

À l'heure de la société numérique et hautement technologique, tout un chacun, au sein des pays développés, aspire à collectionner les appareils électriques et électroniques conçus pour assurer le confort de la vie quotidienne. Du réfrigérateur-congélateur-distributeur-de-glaçons au home-cinéma-dolby-surround-avec-lecteur-Bluray-intégré, nos maisons se transforment au fur et à mesure que de nouvelles inventions sont lancées sur le marché. La société n'est pas en reste : les nouvelles LGV permettent de replier l'espace et le temps (Paris-Bordeaux en à peine 2h ou encore Paris-Rennes en 1h25), les immeubles s'érigent vers le ciel pour tenter de toucher les étoiles (la tour Burj Khalifa à Dubaï, avec ses 830 mètres, sera théoriquement bientôt dépassée par la Jeddah Tower saoudienne et ses 1001 mètres à l'horizon 2020), et les voitures électriques se développent voire se démocratisent afin de nous libérer du pétrole et de limiter notre impact environnemental. La technologie s'est fait une place de choix au coeur de nos vies et peu d'entre nous envisageraient sereinement de se séparer de leurs biens et de revenir un siècle ou deux en arrière.

Mais toutes les plus belles roses ont des épines. Lorsque, dans votre petit studio, le bruit du réfrigérateur vous empêche de dormir, lorsque les avions, qui décollent et atterrissent près de chez vous, vous empêchent de vivre pleinement la finale de la coupe du monde de football, ou lorsque les trains passent toutes les dix minutes et font vibrer les vitrines de votre nouveau meuble IKEA, un sentiment aigre-doux d'irritation risque de vous saisir. Le revers de la médaille ne s'est pas fait attendre : nous évoluons désormais *de facto* au milieu de pollutions sonores et vibratoires. Les vibrations sont partout : structurelles ou acoustiques, elles sont la plupart du temps perçues comme parasites. L'enjeu est alors de les comprendre et surtout de les maîtriser au mieux, afin de les atténuer, voire de les éliminer, ou *a contrario*, de changer le point de vue et de les utiliser sciemment afin qu'elles contribuent au but recherché.

La problématique n'est pourtant pas nouvelle : les Grecs et les Romains, par exemple, s'intéressaient déjà de très près à l'acoustique des théâtres antiques. Les grands bâtisseurs chrétiens avaient eux aussi acquis une maîtrise fine des lois physiques afin d'améliorer l'acoustique des cathédrales. Finalement, les habitants de sites à haut risque sismique se sont très vite rendu compte de l'impact catastrophique de vibrations fortes et imprévisibles sur les structures et ont élaboré des stratégies afin d'assurer un maximum de sécurité à leurs communautés. Plus simplement, dès que l'homme a décidé d'ériger des bâtiments à un ou plusieurs étages au cours de son processus de développement, la question centrale des vibrations structurelles et sonores s'est imposée et est devenue un champ d'investigation majeur.

Dans quel contexte se pose-t-elle de nos jours ? Les champs d'étude sont très variés mais tendent globalement à améliorer le confort et la sécurité de tous, la longévité des objets et des structures, ou

encore à réduire l'impact environnemental. Citons aussi tout l'enjeu de la parole dans notre société de l'information et de la communication : les haut-parleurs des gares doivent transmettre des messages très clairs et audibles par tous, les microphones des smartphones doivent certes capter la voix de l'utilisateur mais aussi estomper les bruits de fond, tandis que les différents systèmes mains-libres se doivent d'offrir une qualité de communication optimale.

La maîtrise des vibrations a déjà donné des résultats innovants : des théâtres dont l'acoustique "intelligente" est reconfigurable en quelques clics grâce à un système électro-acoustique révolutionnaire et des panneaux partiellement mobiles, des salles de spectacle qui affichent un niveau sonore uniforme sur une grande surface, des casques audio avec insonorisation active par rapport à l'environnement extérieur ou encore l'absorption des vibrations d'un très grand bâtiment par l'eau d'une piscine en U construite sous le toit [130]. Des systèmes audio ultra-directifs, utilisés dans les musées, aux ultrasons focalisés pour les opérations chirurgicales, nos vies ont changé grâce à la compréhension des phénomènes vibratoires.

Les matériaux, techniques et structures évoluent néanmoins constamment et apportent chaque jour leur lot de nouveaux problèmes. Les matériaux plus légers réduisent par exemple la consommation en énergie des moyens de transports construits avec, mais la quantité de vibrations associées est d'autant plus grande et doit être amortie afin d'éliminer les nuisances sonores pour les passagers. Notons finalement qu'à l'ère des *smart buildings*, la réduction des coûts et la préservation de l'environnement sont des priorités qui génèrent de nouvelles problématiques pour lesquelles la compréhension des comportements vibratoires demeure un enjeu clef.

Génèse du sujet

Le constat s'impose donc : le nombre croissant de matériaux et de systèmes vibrants innovants va de pair avec les bruits et vibrations parasites potentiellement associés, d'où l'enjeu, sans cesse renouvelé, d'améliorer le confort et la sécurité de tous. Le point de départ de la démarche est la caractérisation de ces structures vibrantes, afin de comprendre leur comportement dynamique, ce qui implique des simulations et surtout des mesures, et donc des méthodes de mesure efficaces.

Actuellement, un nombre certain de techniques de mesure est à disposition, avec diverses limites associées. Ne seront donnés ici que quelques exemples représentatifs, l'état des lieux plus détaillé des méthodes de mesure de vibrations étant présenté ultérieurement (*cf.* **chapitre 2**). Les accéléromètres par exemple sont intrusifs, et les vibromètres laser simples ou à balayage ne permettent que des mesures ponctuelles. Les méthodes inverses, la méthode RIFF par exemple, ne ciblaient pas, initialement, les phénomènes instationnaires. Une grande partie des techniques de mesure plein champ dynamiques, quant à elles, s'accompagnent de fortes contraintes sur le signal du phénomène mesuré (l'holographie optique et la shearographie numérique ne sont applicables, par exemple, qu'à de petits déplacements de la surface mesurée), sur l'environnement (typiquement l'holographie acoustique ne peut fonctionner avec du bruit de fond) ou sur l'objet même (une taille limitée ou un albedo élevé notamment, pour les méthodes interférométriques ou la déflectométrie, par exemple).

Au vu des inconvénients et contraintes cités ci-dessus, l'idée a été de se tourner vers un autre domaine : celui de la photogrammétrie dynamique et plus précisément de la vision 3D. En effet, les méthodes de vision 3D donnent des mesures de forme plein champ. En y associant des caméras ultra-rapides afin d'avoir accès à la dimension temporelle du déplacement de la structure étudiée, on a alors la possibilité d'effectuer des mesures plein champ, synchrones, applicables à des phénomènes transitoires, non répétables, possiblement non linéaires.

Afin d'étudier la pertinence de cette approche, une étude de faisabilité a été menée dès 2010 [142]. Elle atteste que la photogrammétrie, et notamment les méthodes de vision 3D, permettent, avec deux points de vue, de mesurer de très petits déplacements. Par la suite, un stage de fin d'études de cycle ingénieur [16] a mis en pratique les conclusions de cette étude en utilisant des appareils photos.

L'aboutissement de ce travail est la mise en oeuvre d'un montage utilisé initialement dans le cadre de la robotique principalement, mais transférable à ce contexte, et qui permet de n'utiliser qu'une seule caméra rapide avec un jeu de quatre miroirs. La mesure réussie de forme et de déformations statiques a ouvert la voie à la mesure de vibrations, en affinant le protocole.

De ces travaux préliminaires est issu le sujet de ce présent travail : mener à bien la **« Mesure de vibrations par vision 3D »**. Ce projet est soutenu par « Le Mans Acoustique » et financé par la région « Pays de la Loire » et le « Fonds européen de développement économique et régional ».
L'objectif global de cette thèse est ainsi d'étudier la pertinence et les limites de la vision 3D couplée à une caméra ultra-rapide pour la mesure 3D de champs vibratoires, sans contact et de manière synchrone, couvrant le domaine des fréquences associées à la vibro-acoustique, soit de 20 Hz à plusieurs kHz.

Motivations et cahier des charges

Premièrement, il s'agit de caractériser et comprendre les déplacements de structures. La démarche s'inscrit en rupture par rapport à la mesure de forme déjà très étudiée avec cette technique : l'intérêt se porte sur la mesure de petits déplacements, *i.e.* sub-millimétriques. Ainsi, des simulations et expérimentations ont été effectuées pour en étudier la faisabilité, estimer la précision que l'on peut obtenir en fonction des conditions de prise de vue, du matériel optique et de la taille de l'objet et finalement en évaluer l'applicabilité aux problématiques industrielles.

Deuxièmement, il s'est avéré très vite nécessaire de concevoir des outils et méthodes adaptés, le mot d'ordre étant la simplification des protocoles et leur adaptation à la mesure de vibrations. Un travail essentiel a ciblé les méthodes d'analyse possibles en traitement et post-traitement, afin d'optimiser le lien entre les données de mesure et leur analyse tout en réduisant le temps de calcul. Finalement, l'utilisation de logiciels payants ayant été écartée puisque la chaine complète de traitement devait être maitrisée afin d'optimiser le temps de traitement, de nouveaux codes adaptés, spécifiques et souples ont été développés.

Réalisations

Tout d'abord, il a fallu ***élaborer une méthodologie de reconstruction de champs de déplacements à partir des images acquises selon les positions, orientations et caractéristiques de deux caméras rapides.***
Ceci comprend :

- la compréhension du fonctionnement de la stéréo-vision et de la mesure de forme avec deux points de vue ;
- l'élaboration d'un protocole stéréo avec deux points de vue pour l'appliquer à la mesure de vibrations ;
- l'utilisation de la méthode de suivi de cibles.

Ensuite, ***la méthode élaborée ci-dessus a dû être adaptée pour une utilisation avec une seule caméra et un jeu de quatre miroirs.***
Ceci inclut :

- le dimensionnement du système pseudo stéréoscopique utilisant une seule caméra et un jeu de quatre miroirs ;
- l'adaptation des codes du protocole deux caméras au protocole une caméra ;
- une étude de sensibilité par simulation pour optimiser différents paramètres du protocole (l'angle et le positionnement des points de vue, le motif utilisé pour les cibles, le choix du détecteur utilisé) afin de maximiser la précision de la mesure.

En parallèle, des besoins spécifiques sont apparus, conduisant à la ***création d'outils de traitement.***
Ceci correspond à :
- la création des outils numériques de traitement des images afin d'obtenir la mesure de forme à partir du suivi de cibles ;
- la création des outils d'analyse modale pour l'extraction des fréquences de résonance et l'affichage des déformées modales opérationnelles ;
- l'adaptation de la détection à une mesure plein champ, et son optimisation, en utilisant un motif aléatoire, les outils de corrélation d'image et en développant les codes adaptés.

Cette phase a mené à la conception de la Méthode Itérative de Rectification d'Images, notée Méthode IRIs.

Une deuxième phase de travail s'est focalisée sur la ***simplification du protocole et du montage pseudo stéréoscopique avec un jeu de quatre miroirs.***
Cette étape a abouti à :
- la conception du montage pseudo stéréoscopique utilisant une caméra rapide et un jeu de deux miroirs uniquement, plus simple d'utilisation que le montage avec quatre miroirs ;
- la conception d'une méthode de mesure de vibrations n'utilisant aucun miroir adaptée à la mesure de déplacements mono-axiaux dans l'espace 3D ;
- la conception d'une méthode à caméras asynchrones, n'utilisant aucun miroir, adaptée à la mesure de déplacements multi-axiaux.

La mise au point de ces techniques constitue l'aboutissement de ce travail de recherche.

Finalement, ***des validations expérimentales des méthodes ont été obtenues successivement, en :***
- testant expérimentalement les méthodes avec les objets d'étude suivants, en fonction des montages : une plaque, un haut-parleur, un tambourin, un capot de voiture ;
- comparant les résultats avec ceux obtenus par des méthodes de référence ;
- menant une étude de faisabilité ciblée afin d'évaluer l'applicabilité de la démarche à certains enjeux industriels (les objets d'étude choisis, dans ce cas, sont un capot de voiture et un haut-parleur).

Plan du manuscrit

La partie I vise à donner des clefs au lecteur. Le **chapitre 1** présente tout d'abord la notion centrale de vibrations à travers quelques exemples ciblés, en fonction des objets d'étude choisis lors des phases expérimentales décrites ultérieurement. Les principaux outils de traitement du signal et d'analyse modale utilisés sont aussi expliqués. Le **chapitre 2** propose ensuite un état des lieux des méthodes de mesure de vibrations à disposition actuellement, afin de mieux cerner le contexte global dans lequel ce travail s'inscrit. Cette section explique et évalue les principales approches, afin de clarifier les atouts des méthodes de vision 3D, ainsi que leurs potentiels inconvénients, un des enjeux implicites de cette thèse consistant à repousser ces limites, si nécessaire, par des solutions innovantes. Ensuite, comme expliqué précédemment, le sujet et le contexte théorique de cette thèse s'inscrivent au point de confluence de plusieurs domaines : la notion de vibrations ayant déjà été abordée, les deux chapitres suivants ont pour but d'exposer le plus brièvement possible les fondements théoriques inhérents aux deux autres disciplines qui sous-tendent cette recherche, afin de faciliter la compréhension des travaux exposés. Le **chapitre 3** donne ainsi les bases géométriques et mathématiques de la stéréo-vision, sur lesquels reposent les méthodes de vision 3D conçues. Les notions de modèle sténopé, de capteur stéréoscopique, ou encore de repères, sont définies, tout en expliquant les processus de calibrage et de triangulation. Le **chapitre 4** cible enfin quelques concepts et outils de traitement d'images avec notamment la détection, l'appariement et le suivi de points de mesure, obtenus à partir de cibles ou d'un motif aléatoire. Quelques détecteurs de cibles sont passés en revue ; des outils de corrélation d'images, utilisés pour ce travail, sont présentés. Les principes méthodologiques ayant été posés, la deuxième partie de ce manuscrit s'articule autour des différents montages testés et des résultats obtenus.

Ainsi, la partie II détaille l'approche expérimentale et met à jour les apports de ce travail. En introduction, le **chapitre 5** explique le protocole de mesure appliqué, les choix opérés, les principaux problèmes rencontrés ainsi que les solutions trouvées. Le **chapitre 6** fait suite aux travaux de M. Biggio [16] et aborde la pseudo stéréo-vision appliquée à la mesure de vibrations. Le système pseudo stéréoscopique à caméra unique et jeu de quatre miroirs est décrit, ainsi que son dimensionnement. Ce chapitre présente aussi l'étude de sensibilité à certains paramètres menée avec une méthode de suivi de cibles, à l'aide de simulations réalisées avec le logiciel Blender [18]. La phase de mesures associée, sur un objet à géométrie complexe, a finalement mis à jour la nécessité d'utiliser un autre système de détection pour une mesure plein champ. Le **chapitre 7** expose ainsi la méthode de traitement d'images (dite méthode IRIs) conçue spécifiquement pour ce contexte, permettant d'initialiser la mesure de vibrations d'un objet présentant une forte profondeur, en utilisant un motif aléatoire et des outils de corrélation d'images. Les expérimentations présentées dans les deux chapitres précédents ayant montré les limites du montage pseudo stéréoscopique en termes de maniabilité, des simplifications du montage et du protocole ont été tentées. Le **chapitre 8** présente tout d'abord la simplification opérée sur le montage et décrit le montage pseudo stéréoscopique conçu avec une caméra ultra-rapide et un jeu de deux miroirs uniquement. Les résultats expérimentaux de mesure de vibrations sur un seul point de mesure et plein champ pour un haut-parleur sont affichés et valident une première fois la méthode. Une seconde validation est présentée, sur une mesure plein champ d'une plaque. Le **chapitre 9** décrit la simplification du protocole, basée sur l'hypothèse de la nature localement mono-axiale des déplacements mesurés. La méthode correspondante a été construite autour d'un montage n'utilisant aucun miroir. Les principes théoriques sont exposés, ainsi que les résultats expérimentaux, pour un seul point de mesure et/ou pour une mesure plein champ, obtenus en choisissant comme objets d'étude un tambourin, une plaque et un haut-parleur. L'approche est ainsi validée et s'avère *in fine* particulièrement adaptée à la mesure de déplacements mono-axiaux dans l'espace 3D. Finalement le **chapitre 10** présente une version améliorée de la méthode sans miroir, permettant la mesure de vibrations non plus mono-axiales mais multi-axiales. Cette dernière méthode, utilisant deux caméras asynchrones, est décrite, ainsi que les résultats expérimentaux, obtenus sur une plaque, qui la valident. Une étude de faisabilité ciblant deux cas industriels est enfin présentée, dans la logique du travail effectué, à savoir en prenant pour objet d'étude un capot de voiture pour la partie vibratoire et un haut-parleur volontairement dégradé pour la partie vibro-acoustique. La **Conclusion** de ce travail permet de mieux cerner les caractéristiques de chacune de ces méthodes, en présentant notamment une comparaison des résultats obtenus pour chacune d'elles lors d'une même phase expérimentale, sur le même objet d'étude (ici une plaque).

En bref...

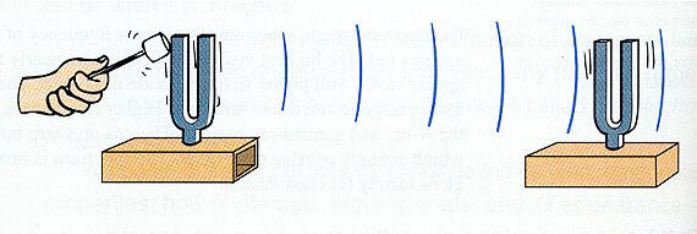
Ce travail se situe à l'intersection de différentes disciplines : la vibro-acoustique, la vision 3D et la photogrammétrie dynamique, la métrologie, le traitement de signal et d'images, ou encore la programmation. Rappelons que la finalité de cette thèse est d'étudier la pertinence et les limites de la vision 3D couplée à une caméra ultra-rapide pour la mesure 3D de champs vibratoires, sans contact et de manière synchrone, dans le domaine des fréquences associées à la vibro-acoustique. La question du montage à utiliser s'est rapidement posée et a abouti, dans une logique de simplification, à la mise en oeuvre de quatre montages ainsi qu'à la conception de trois méthodes de mesure de vibration plein champ et d'une technique de traitement d'images, adaptées à notre contexte. Les objets d'étude choisis vont d'une plaque à un capot de voiture, en passant par un tambourin et un haut-parleur, afin d'explorer au mieux différentes applications du domaine de la vibro-acoustique. Une comparaison des protocoles de mesure conçus et des résultats associés conclut ce manuscrit et met à jour les potentialités respectives de ces méthodes de vision 3D appliquées à la mesure de vibrations.

Première partie

Contexte, méthodes et outils

Chapitre 1

Vibrations



« Si vous voulez trouver les secrets de l'univers, pensez en termes d'énergie, de fréquence et de vibration. »
Nikola Tesla

Sommaire

1.1 Introduction **9**
1.2 Systèmes vibrants **10**
1.2.1 Système masse-ressort 10
1.2.2 Vibrations d'une plaque rectangulaire 11
1.2.3 Vibrations d'un haut-parleur 14
1.2.4 Ordres de grandeur à mesurer 17
1.3 Outils utilisés pour le traitement **18**
1.3.1 Outils de traitement du signal 18
1.3.2 Outils d'analyse modale 19
1.4 Conclusion **21**

1.1 Introduction

Comme exposé en introduction générale, nous vivons désormais au milieu de vibrations potentiellement perçues comme nuisibles : certains bruits parasitent le quotidien et les vibrations structurelles engendrent, par exemple, une fatigue des matériaux qui impacte leur longévité et notre sécurité. La caractérisation des propriétés mécaniques et vibratoires des structures s'avère ainsi vitale, et vise à améliorer la compréhension des phénomènes en jeu afin de concevoir des modèles numériques plus performants. L'enjeu est de réussir à trouver, grâce à ces modèles, des solutions innovantes qui amènent à une maitrise plus complète des phénomènes vibratoires et permettent d'éviter les trop fortes amplitudes de déformations. Il est alors possible de réduire voire d'éliminer les nuisances et d'améliorer le confort et la sécurité. Gardons aussi à l'esprit que la maîtrise des vibrations sonores est un atout essentiel dans le domaine de la musique ou de la communication : par exemple pour l'industrie du haut-parleur, la caractérisation vibratoire permet de développer de nouveaux produits à coûts moindres mais affichant un meilleur rendu audio.

Ainsi que le rappelle le **chapitre 2**, les systèmes de mesure de vibrations utilisés actuellement pour la caractérisation des propriétés vibratoires des structures sont très performants pour les phénomènes répétables, et souvent au stade de l'étude approfondie pour les techniques de mesure des phénomènes transitoires et/ou non-linéaires et/ou non-répétables, sachant qu'afin de concevoir des méthodes adaptées, il est nécessaire de comprendre *in extenso* le comportement des structures étudiées. Dans le contexte de ce manuscrit, la présentation du comportement du système masse-ressort apportera les bases nécessaires à la compréhension du comportement des objets plus complexes utilisés au cours de cette thèse : une plaque et un haut-parleur. Rappelons enfin qu'une fois les mesures effectuées et l'acquisition des signaux vibratoires opérée, une phase de traitement est indispensable avant d'obtenir des résultats exploitables. L'analyse des signaux dans le domaine temporel étant complexe, il est habituel de se tourner vers le domaine fréquentiel afin d'extraire des informations (par exemple les fréquences de résonance ou les déformées modales) voire vers l'analyse temps-fréquence, très utile

dans certains cas. Ce traitement est effectué grâce à des outils de traitement du signal et d'analyse modale.

Ce chapitre présente des résultats certes classiques pour les mécaniciens ou encore les acousticiens. Ils sont néanmoins sciemment rappelés ici, ce manuscrit de thèse étant destiné à une large communauté incluant des personnes issues d'autres domaines, notamment de la vision. Une première section s'applique ainsi à présenter des systèmes vibrants (**section 1.2**) : le comportement d'un système masse-ressort, d'une plaque et d'un haut-parleur, tout en définissant la notion de résonance. Y est inclue la quantification des ordres de grandeur des déplacements mentionnés et donc à mesurer. La deuxième section présente et explique les outils utilisés en traitement du signal, notamment la transformée de Fourier et la fonction de réponse en fréquence (**section 1.3.1**), et les outils permettant de détecter les fréquences de résonance et d'afficher les déformées modales (**section 1.3.2**).

1.2 Systèmes vibrants

1.2.1 Système masse-ressort

Un système masse-ressort est constitué d'une masse m accrochée à un ressort de raideur k, avec un amortissement ξ (*cf.* **figure 1.1**). Il s'agit d'un système à un degré de liberté où seule la masse est mobile.

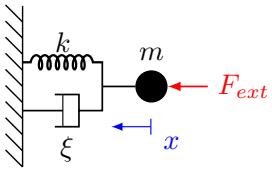


FIGURE 1.1 – Schéma de principe du système masse-ressort

Quand la masse est déplacée en exerçant une force statique, constante dans le temps, le ressort exerce en retour une force sur la masse, proportionnelle au déplacement par rapport à la position d'équilibre, afin de ramener la masse vers la position d'équilibre, selon l'équation :

$$F = -kx \tag{1.1}$$

avec F la force exercée par le ressort sur la masse, k la raideur du ressort et x la distance de la masse à sa position d'équilibre.

Lorsque l'on met ce système en vibration, en exerçant une force dynamique sinusoïdale, la masse oscille alors autour de sa position d'équilibre. Le nombre d'oscillations par seconde, soit la fréquence, est noté f. Son inverse est la période, ou temps d'une oscillation, et est notée T avec $T = 1/f$.

L'expression du déplacement sous forme d'équation s'appuie sur le principe fondamental de la dynamique : « la somme des forces est égale à la masse multipliée par l'accélération », selon l'équation (1.2) :

$$\sum^{i} \overrightarrow{F_i} = m\overrightarrow{a} \tag{1.2}$$

avec m la masse et $\overrightarrow{a}$ l'accélération.

En appliquant ce principe au système masse-ressort et en identifiant les forces exercées, on obtient :

$$-kx - \xi v + F_{ext} = ma \tag{1.3}$$

avec ξ l'amortissement, F_{ext} la force exercée sur la masse, x, v et a le déplacement, la vitesse et l'accélération de la masse, notés sous forme non-vectorielle, le système présentant ici un seul degré de liberté.

La vitesse et l'accélération correspondant respectivement à la dérivée temporelle du déplacement et de la vitesse, l'équation (1.3) devient :

$$m\frac{\partial^2}{\partial t^2}x + \xi\frac{\partial}{\partial t}x + kx = F_{ext} \quad . \tag{1.4}$$

En basculant dans le domaine fréquentiel, une dérivée temporelle correspond à une multiplication par $j\omega$ avec j l'unité imaginaire (telle que $j^2 = -1$) et ω la pulsation égale à $2\pi f$. L'équation (1.4) devient alors :

$$(-\omega^2 m + j\omega\xi + k) * x(f) = F_{ext}(f) \quad . \tag{1.5}$$

La fonction de réponse en fréquence (*cf.* **section 1.3.1**) qui relie le déplacement $x(f)$ et la force appliquée $F_{ext}(f)$ est donc :

$$\frac{x(f)}{F_{ext}(f)} = \frac{1}{-\omega^2 m + j\omega\xi + k} \quad . \tag{1.6}$$

La fréquence correspondant au maximum de déplacement fait entrer le système en résonance. Le déplacement étant maximal lorsque $-\omega^2 m + k = 0$, la fréquence de résonance est exprimée par la formule suivante :

$$f_r = \frac{\sqrt{k/m}}{2\pi} \quad . \tag{1.7}$$

En illustration, la **figure 1.2** présente le tracé de la fonction de réponse en fréquence d'un système masse-ressort pour différentes valeurs d'amortissement. En statique, pour $f = 0$ ou f proche de zéro, ω est négligeable par rapport à k. En reprenant l'équation (1.6), on aboutit à $x(f)/F_{ext}(f) \approx 1/k$, avec $1/k$ la souplesse du ressort. À la fréquence de résonance, les termes $-\omega^2 m$ et k de l'équation (1.6) s'annulent, seul reste l'amortissement $1/j\omega\xi$. Plus l'amortissement ξ est petit, plus le déplacement est grand. Si ξ est grand, le déplacement à la fréquence de résonance peut même être inférieur à $1/k$, la réponse en statique. Après la fréquence de résonance, le terme $\omega^2 m$ prédomine, on a donc une décroissance selon le carré de la fréquence.

1.2.2 Vibrations d'une plaque rectangulaire

À première vue, une plaque est une structure différente d'un système masse-ressort car sa vibration correspond à une onde qui se propage *a priori* dans toute la structure. Pourtant, même si un tel solide présente en effet, en théorie, six degrés de liberté à chaque point de la plaque, dans beaucoup de cas les rotations sont considérées, à raison, comme négligeables. De plus, les déplacements sont globalement hors du plan de l'objet : ainsi ne sont pas non plus pris en compte les deux degrés de liberté correspondant aux déplacements dans le plan de l'objet. Au final, la plaque peut donc aussi être considérée comme se déformant selon un seul axe. Néanmoins, contrairement au système masse-ressort où un seul point est mobile, dans le cas d'une plaque, toute la surface vibre de manière non-uniforme. Quant au comportement mécanique de la plaque, il présente une forte similitude avec

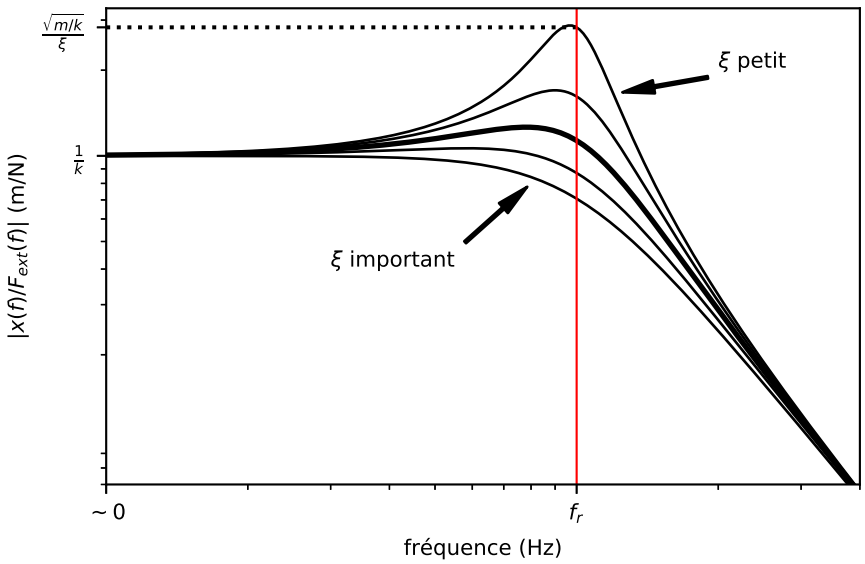


FIGURE 1.2 – Fonction de réponse en fréquence du système masse-ressort (déplacement sur force), tracée pour différentes valeurs d'amortissement

celui du système masse-ressort : en effet, il y a toujours une masse (surfacique), une raideur (de matériau), et un amortissement. Pour rappel, la rigidité de flexion est donnée par l'équation :

$$D = \frac{Eh^3}{12(1-\nu^2)} \tag{1.8}$$

avec h l'épaisseur, E le module d'Young et ν le coefficient de Poisson de la plaque.

En analysant l'équation (1.8), on note que plus la plaque est fine, plus elle est souple. La rigidité de la plaque est l'équivalent de la raideur du système masse-ressort défini dans la section précédente. Comme vu précédemment, selon l'équation (1.6), plus le système est souple, plus les amplitudes de vibrations sont grandes et les fréquences de résonance basses. D'où le compromis à trouver pour l'industrie : soit augmenter l'épaisseur des plaques pour diminuer les vibrations mais obtenir des structures plus lourdes et plus coûteuses, soit diminuer l'épaisseur des plaques pour diminuer le poids et les coûts mais obtenir davantage de vibrations potentiellement parasites. Des solutions sont donc à l'étude afin de concevoir des structures légères tout en limitant la quantité et l'amplitude des vibrations, comme le concept de trou noir acoustique par exemple [118].

Comment les résonances se créent-elles dans une plaque ? Une onde s'y déplace avec une certaine célérité. Avec une excitation mono-fréquentielle, le déplacement décrit un sinus dans le plan de la plaque. On définit alors la longueur d'onde λ qui correspond à une période spatiale de sinus. On a :

$$\lambda = c/f \tag{1.9}$$

avec c la célérité de l'onde dans la plaque.

Contrairement au système masse-ressort où seule la masse oscille, ici, toute une surface se déplace de manière non-homogène. Quand la géométrie de l'objet et la longueur d'onde sont en adéquation, on obtient une fonction de déplacement spécifique, ou mode, et une forte amplitude de vibrations peut être obtenue pour une faible excitation. On parle alors de résonance. Les modes sont directement liés aux caractéristiques de la plaque et aux conditions aux limites. Pour certaines conditions aux

limites uniformes, si les caractéristiques mécaniques sont connues, les modes peuvent être calculés. La **figure 1.3** présente quelques exemples, avec des conditions aux limites uniformes du type simplement supporté, de déformées de plaque selon la longueur, lorsque cette dernière contient un multiple entier de demi-longueurs d'onde ($\lambda = 2l$, $\lambda = l$ et $\lambda = 2l/3$). Les maxima de vibrations sont appelés ventres de vibrations et les points fixes nœuds de vibrations.

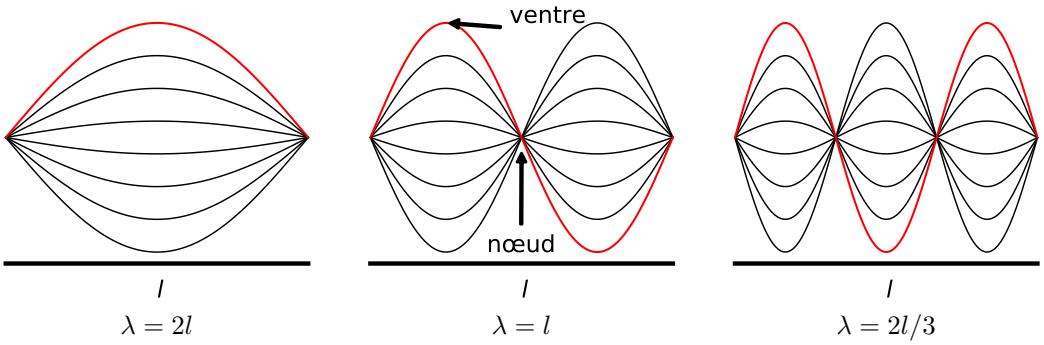


FIGURE 1.3 – Exemples de modes d'une plaque selon la longueur pour des conditions aux limites du type simplement supporté

Si l'on s'intéresse au déplacement de toute la surface de la plaque, et non plus seulement sur sa longueur, des couplages de modes sont observables dans les différentes directions. En illustration, la **figure 1.4** présente trois modes d'une plaque rectangulaire avec des conditions aux limites du type simplement supporté, avec de gauche à droite :

— 1 longueur d'onde dans la largeur et 0,5 dans la longueur
— 0,5 longueur d'onde dans la largeur et 1 dans la longueur
— 1 longueur d'onde dans la largeur et 1,5 dans la longueur

Les ventres de vibration (positifs en *rouge* et négatifs en *bleu*) et les lignes nodales (*i.e.* l'équivalent des nœuds de vibrations en 2D, ici en *noir*), génèrent différents motifs symétriques. La barre de couleurs à droite de l'image correspond à la valeur du déplacement, ici normé par rapport à la valeur maximale.

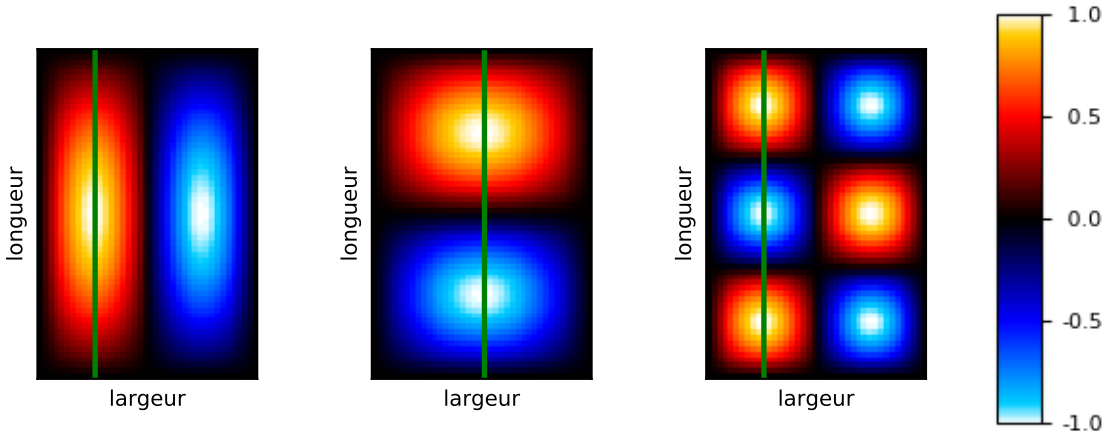


FIGURE 1.4 – Exemples de modes d'une plaque rectangulaire simplement supportée dont les vues en coupe (lignes *vertes*) sont affichées sur la **figure 1.3**

Pour des conditions aux limites non-uniformes, lorsque la plaque est encastrée aux quatre coins par exemple, les solutions analytiques deviennent plus complexes. En effet, la raideur et l'amortissement

du système (plaque + support) vont dépendre de la position où l'on se situe sur la surface. On a alors recours à des simulations et/ou des mesures afin de comprendre le comportement vibratoire du système et de caractériser les paramètres mécaniques aux différents endroits de la structure. Lors des mesures, l'objet étudié subit une excitation spécifique à un endroit donné, avec un pot vibrant par exemple. La réponse à cette excitation est alors mesurée. Les fonctions de déformation spécifiques mesurées sont appelées déformées modales opérationnelles.

En illustration, la **figure 1.5** donne un exemple de déformée modale opérationnelle d'une plaque carrée encastrée aux quatre coins, mesurée avec un vibromètre laser à balayage. Il est difficile d'obtenir des conditions aux limites parfaitement identiques aux quatre coins, on constate donc que les positions des ventres de vibration (positifs en *rouge* et négatifs en *bleu*) et des lignes nodales (en *noir*) génèrent un motif non symétrique. On retrouve cependant une certaine cohérence : une alternance de valeurs positives et négatives est visible de chaque côté d'une ligne nodale.

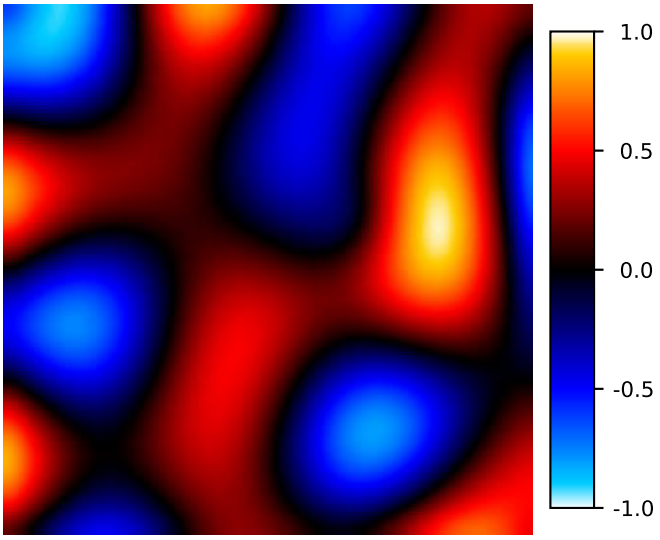


FIGURE 1.5 – Exemple de déformée modale opérationnelle d'une plaque carrée de 330 mm × 330 mm, encastrée aux quatre coins, mesurée avec un vibromètre laser à balayage pour une zone de mesure de 300 mm × 300 mm

1.2.3 Vibrations d'un haut-parleur

Un haut-parleur est un transducteur électro-acoustique qui permet de transformer un signal électrique en signal acoustique. Son comportement vibratoire ressemble à celui d'un système masse-ressort, auquel se rajoute un couplage électromagnétique qui génère, selon un signal électrique défini, une force sur la partie mobile [82]. Cependant, contrairement au système masse-ressort, le couplage électromagnétique ou les suspensions peuvent entrainer des comportements non-linéaires, partiellement caractérisables, par exemple, par une mesure de déplacement effectuée avec un vibromètre laser sur un point, en parallèle de la mesure électrique [137]. De plus, en hautes fréquences la membrane se fractionne et on observe donc des déformées modales opérationnelles que diverses méthodes permettent de mettre à jour, afin d'améliorer la compréhension du comportement dynamique du haut-parleur.

Afin de mieux visualiser le fonctionnement d'un haut-parleur électro-dynamique, une vue en coupe est présentée sur la **figure 1.6**. Un puissant aimant permanent (1) crée un flux magnétique. Ce flux est guidé par un circuit magnétique afin de traverser une bobine (2). Si cette bobine est traversée par un courant, une force de Laplace est générée et s'applique sur la membrane du haut-parleur (4). Le mouvement de la membrane va ensuite exercer une pression sur l'air : un signal acoustique. La membrane du haut-parleur est fixée à deux suspensions (3).

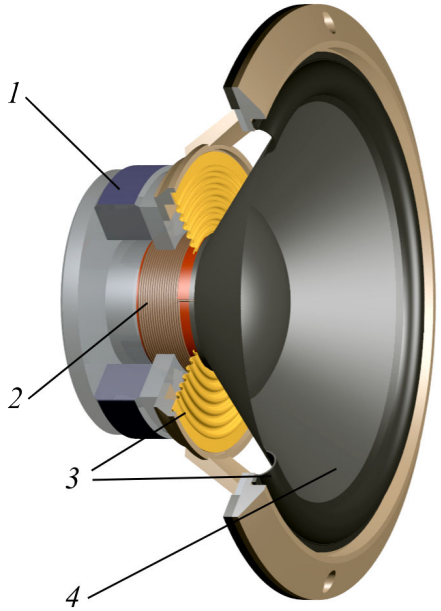


FIGURE 1.6 – Vue en coupe d'un haut-parleur électro-dynamique ; (*1*) aimant permanent, (*2*) bobine, (*3*) suspensions, (*4*) membrane ; wikimedia commons, par *Svjo*, le 26/12/2013

En première approximation, la membrane est supposée indéformable, on a l'équivalent d'un système masse-ressort : la membrane, le support de la bobine et la bobine correspondent à une masse et la suspension amène la raideur et l'amortissement. Cependant, la membrane n'est pas parfaitement rigide, elle peut se déformer lorsque le haut-parleur chante. On observe alors des déformées modales opérationnelles correspondant à des modes de membrane. De plus, le système n'est jamais parfaitement équilibré et symétrique : il peut y avoir des mouvements de balancier ou *rocking modes*. Ces modes vont influencer la réponse acoustique, il est donc important d'arriver à les caractériser pour les limiter.

Afin d'obtenir un ordre de grandeur du déplacement à mesurer sur la membrane du haut-parleur, il est nécessaire de calculer sa fonction de réponse en fréquence (déplacement sur tension). Le modèle utilisé est détaillé ci-dessous, et s'appuie sur le modèle linéaire de Thiele & Small [157] en incluant implicitement l'impédance acoustique de rayonnement dans l'impédance mécanique pour simplifier l'exposé. La force de Laplace exercée sur la membrane est :

$$F = Bli \tag{1.10}$$

avec B le flux magnétique traversant la bobine, l la longueur du fil de la bobine et i le courant la traversant. Généralement, on parle du facteur Bl, le facteur de force. La notation Bl^2 correspond à $(Bl)^2$. La bobine est fixée à la membrane, donc lorsque la bobine se déplace, la membrane vibre aussi avec une certaine vitesse v. Selon la loi de Faraday, une tension est générée au niveau de la bobine :

$$U = Blv \quad . \tag{1.11}$$

Lorsque l'on applique une tension aux bornes du haut-parleur, l'impédance mécanique (masse-ressort-amortissement) va induire une modification de l'impédance électrique du système. Pour rappel, l'impédance électrique correspond au rapport tension sur courant ($Z_e = U/I$) et l'impédance mécanique au rapport force sur vitesse ($Z_m = F/v$). En utilisant les équations (1.10), (1.11), et les équations de l'impédance électrique et mécanique, on obtient une impédance électrique Z_{me} correspondant à la modification de l'impédance électrique de la bobine induite par l'impédance mécanique couplée électro-magnétiquement à la bobine. On a :

$$Z_{me} = \frac{Bl^2}{Z_m} \quad . \tag{1.12}$$

Afin d'illustrer le phénomène, la **figure 1.7** reprend les courbes calculées par simulation à partir des données réelles mesurées pour le haut-parleur étudié dans la **partie II**. L'impédance théorique de la bobine est tracée en *rouge* et l'impédance du haut-parleur en *noir*. Avant 10 Hz et après 200 Hz, l'impédance du haut-parleur est très proche de celle de la bobine. Entre 10 Hz et 200 Hz, l'impédance mécanique va modifier l'impédance électrique mesurée aux bornes de la bobine par couplage électromagnétique (loi de Faraday). En effet, la valeur absolue de l'impédance du haut-parleur est nettement augmentée autour de sa fréquence de résonance (ici ≈ 45 Hz).

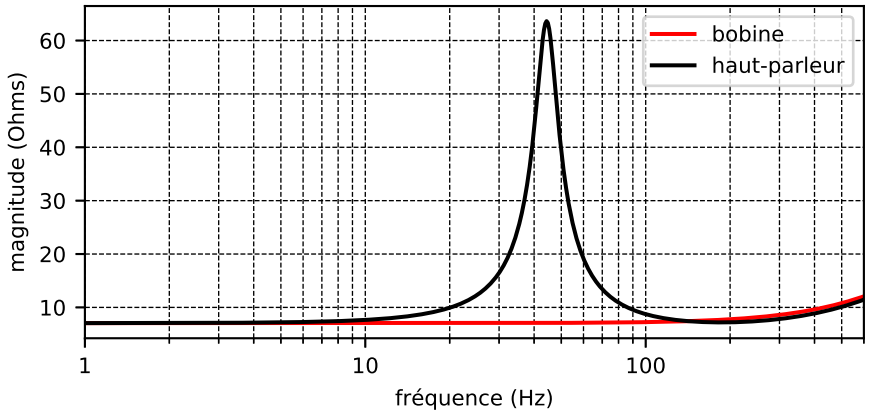


FIGURE 1.7 – Courbes simulées de l'impédance électrique du haut-parleur étudié ; (*rouge*) impédance de la bobine bloquée, (*noir*) impédance du haut-parleur

L'impédance électrique d'une bobine, notée Z_{Le}, peut être modélisée comme une résistance en série avec une inductance. Elle est calculée à l'aide de l'équation :

$$Z_{Le} = R_e + j\omega L_e \tag{1.13}$$

avec R_e la valeur de la résistance de la bobine, L_e son inductance et j le nombre imaginaire (afin d'éviter les confusions avec le courant i). L'impédance mécanique Z_m de la partie mobile du haut-parleur correspond à une résistance en série avec une masse et une souplesse (*i.e.* l'inverse d'une raideur) et est calculée avec l'équation :

$$Z_m = R_m + j\omega M_m + \frac{1}{j\omega C_m} \tag{1.14}$$

avec R_m la résistance mécanique, M_m la masse mobile et C_m la souplesse. L'impédance électrique du haut-parleur Z_{hp} est finalement la somme des impédances Z_{Le} et Z_{me}, et est obtenue avec l'équation :

$$Z_{hp} = Re + j\omega Le + \frac{Bl^2}{Z_m} \quad . \tag{1.15}$$

On peut alors calculer la fonction de réponse en fréquence du haut-parleur, $X(f)/U(f)$ (déplacement sur tension). La vitesse étant la dérivée du déplacement, on a :

$$X(f) = V(f)/(j\omega) \quad . \tag{1.16}$$

La tension $U(f)$ étant égale à l'impédance du haut-parleur multipliée par le courant ($U = iZ_{hp}$), on obtient :

$$\frac{X(f)}{U(f)} = \frac{Bl * v}{j\omega Bl * i * Z_{hp}} = \frac{Bl}{j\omega F/v * Z_{hp}} = \frac{Bl}{j\omega Z_m Z_{hp}} \quad . \tag{1.17}$$

En illustration, la courbe simulée de la fonction de réponse en fréquence (déplacement sur tension) du haut-parleur étudié est présentée sur la **figure 1.8**. On observe que, avant 10 Hz, pour ce haut-parleur, l'amplitude est constante à environ 0,6 mm/V. Aux alentours de la fréquence de résonance [1] ($\approx$ 45 Hz), on note une légère décroissance. Après la fréquence de résonance, la décroissance suit le carré de la fréquence (ceci est dû aux termes de masse et d'inductance dans les impédances Z_m et Z_{hp}).

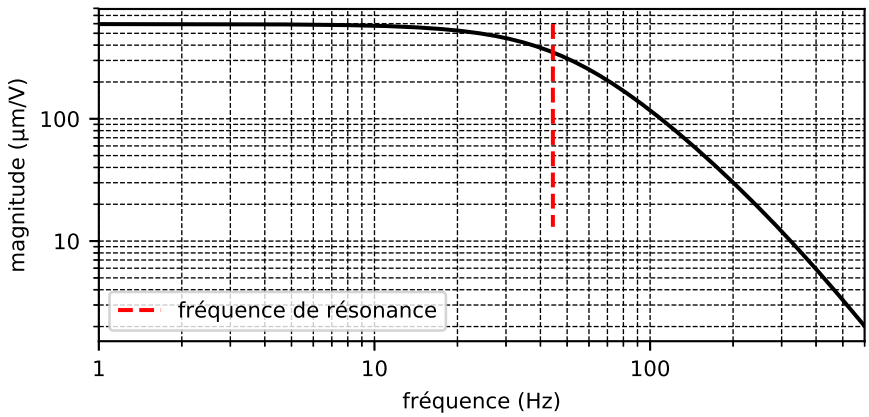


FIGURE 1.8 – Courbe simulée de la fonction de réponse en fréquence (déplacement sur tension) d'un haut-parleur électro-dynamique à bobine mobile

1.2.4 Ordres de grandeur à mesurer

L'amplitude des déformées vibratoires dépend de la taille, de l'amortissement et de la raideur de l'objet ou de la structure, ainsi que de la fréquence du signal d'excitation. Pour une plaque non encastrée de 300 mm $\times$ 300 mm, un signal d'excitation de quelques dizaines de Hertz génère facilement des déplacements de l'ordre du millimètre. Comme l'amplitude du déplacement décroit avec le carré de la fréquence, les déplacements dans la plaque avec un signal d'excitation supérieur à 1 kHz franchissent le seuil du micromètre. En ce qui concerne le haut-parleur, le modèle étudié ici est un woofer, *i.e.* un système conçu pour émettre des sons entre 100 Hz et quelques kHz. La fonction de réponse en fréquence présentée sur la **figure 1.8** indique que pour un signal d'excitation de 1 V, en dessous de 20 Hz, le déplacement est de 0,6 mm. Après la fréquence de résonance ($\approx$ 45 Hz), la courbe décroit selon le carré de la fréquence : à 330 Hz le déplacement est égal à $\approx$ 10 μm, à 660 Hz il est de $\approx$ 2,5 μm, vers 1 kHz il est de $\approx$ 1 μm, etc. toujours pour un signal d'excitation de 1 V. On observe que, pour rester dans le domaine linéaire, l'amplitude maximale de l'excitation doit être de l'ordre du Volt (en sachant que cette valeur dépend du haut-parleur). Il est à noter que les amplitudes données correspondent à des déplacements sans déformation de membrane. Des mesures de déformées ont donné des aperçus des comportements dynamiques de membranes de haut-parleurs [25, 21, 83], mais aucune amplitude de déformées modales opérationnelles n'est précisée, à notre connaissance, dans la littérature. Ce point sera donc à tester empiriquement.

En conclusion, les méthodes qui ont été développées dans cette thèse prennent en compte ces ordres de grandeur afin d'être adaptées aux objets d'étude choisis. Notons que ces données ne sont que des points de départ. Pour rappel, un des objectifs de ce travail est aussi d'amorcer la caractérisation de l'incertitude des méthodes de mesure proposées et donc de définir par la même, avec précision, les ordres de grandeur mesurables avec les protocoles mis en place.

1. Le système étant fortement amorti, la résonance n'est pas visible sur la courbe d'amplitude, tel que présenté sur la **figure 1.2**.

1.3 Outils utilisés pour le traitement

La campagne de mesure permet d'obtenir, pour chaque point de mesure sur l'objet, une série temporelle de déformation, liée aux caractéristiques de vibration de l'objet. Une fois ce signal vibratoire obtenu, il est nécessaire d'en extraire l'information pertinente en fonction du but recherché. Les données temporelles s'avèrent certes utiles ; néanmoins, dans le cadre d'un travail avec des signaux contenant plusieurs fréquences, voire une large bande de fréquences, par exemple pour un signal aléatoire ou une impulsion, l'analyse du signal brut est difficile, peu intuitive, et demeure insuffisante pour obtenir des données complètes et fiables. Le travail dans le domaine fréquentiel simplifie ainsi l'analyse : on observe un signal qui contient de l'énergie dans une bande de fréquences données. On a donc recours aux outils de traitement du signal afin de passer dans le domaine fréquentiel, d'obtenir des informations exploitables et de mieux comprendre le phénomène observé. Notons qu'il est aussi possible de coupler les deux approches et d'effectuer une analyse temps-fréquence, particulièrement pertinente pour l'analyse de signaux transitoires. Une fois obtenues les données fréquentielles de centaines voire de milliers de points de mesure, une analyse modale peut être effectuée afin de déterminer les fréquences de résonance et d'afficher les déformées modales associées. Cette section détaille les outils de traitement du signal et d'analyse modale utilisés.

1.3.1 Outils de traitement du signal

Transformée de Fourier Selon Fourier, tout signal peut être représenté comme une somme infinie de sinus de différentes fréquences. La transformée de Fourier (notée TF) est un outil qui permet de passer du domaine temporel au domaine fréquentiel (*cf.* équation (1.18)). Ainsi, à partir d'un signal temporel, un spectre est calculé, à chaque fréquence considérée, en comparant le signal à un sinus et un cosinus, afin d'extraire la fréquence, l'amplitude et la phase des composantes d'un signal périodique. On a :

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-2j\pi f t} dt \tag{1.18}$$

avec $X(f)$ le spectre, f la fréquence, t le temps, et $x(t)$ le signal temporel. Cet outil est une des bases de l'analyse fréquentielle, sert de point d'appui aux outils suivants et est utilisé tout au long de ce travail. Notons que le fenêtrage du signal temporel permet de lisser le spectre correspondant [145].

Inter-corrélation L'inter-corrélation permet de comparer deux signaux temporels, en faisant glisser, dans le temps, un signal sur l'autre afin de calculer une corrélation à chaque instant. Cette opération est utilisée dans d'autres calculs, par exemple la densité spectrale de puissance. Elle permet aussi d'obtenir une estimation du délai temporel entre deux signaux, à partir de la position de la valeur maximale, correspondant au maximum de ressemblance. L'inter-corrélation $R_{xy}(\tau)$ est estimée avec l'estimateur biaisé, afin de calculer ultérieurement des densités spectrales de puissance (*cf. infra*), avec la formule suivante :

$$R_{xy}(\tau) = \int_{-\infty}^{+\infty} x(t) y(t+\tau) dt \tag{1.19}$$

avec $x(t)$ et $y(t)$ les signaux comparés et τ le retard appliqué pour faire glisser le signal $y(t)$.

Densité spectrale de puissance La Densité Spectrale de Puissance (notée DSP) permet d'obtenir l'auto-spectre S_{xx} d'un signal et l'inter-spectre S_{xy} de deux signaux selon les équations suivantes :

$$S_{xx}(f) = TF\left[R_{xx}\right] \tag{1.20}$$

et

$$S_{xy}(f) = TF\left[R_{xy}\right] \quad . \tag{1.21}$$

Cet outil est notamment utilisé pour calculer une fonction de réponse en fréquence (*cf. infra*), qui est un des outils les plus utilisés pour caractériser un système, par exemple un haut-parleur ou une plaque.

Fonction de réponse en fréquence Une Fonction de Réponse en Fréquence (notée FRF) donne un lien direct entre une excitation donnée et une oscillation résultante, sous hypothèse d'un système linéaire. On peut la calculer principalement par l'intermédiaire de trois estimateurs différents :

— l'estimateur H1, utilisé quand le rapport signal à bruit est meilleur en entrée qu'en sortie ;
— l'estimateur H2, utilisé quand le rapport signal à bruit est meilleur en sortie qu'en entrée ;
— l'estimateur H3, moyenne de H1 et H2, utilisé quand le bruit est similaire en entrée et en sortie.

Dans le cadre de ce travail, la mesure de déplacement s'avère généralement plus bruitée que la mesure de référence. L'estimateur H1 a donc été utilisé avec l'équation suivante :

$$H1(f) = \frac{S_{xy}(f)}{S_{xx}(f)} \tag{1.22}$$

avec S_{xy} l'inter-spectre entre les signaux d'entrée et de sortie et S_{xx} l'auto-spectre du signal d'entrée.

Spectrogramme Le spectrogramme est un outil d'analyse temps-fréquence, qui consiste à calculer des transformées de Fourier sur différentes portions du signal, afin de suivre l'évolution temporelle du spectre. Cet outil est particulièrement intéressant pour extraire des informations de signaux non-stationnaires comme un sinus glissant ou une impulsion.

1.3.2 Outils d'analyse modale

Détermination des fréquences de résonance à partir de l'écart-type de l'auto-spectre

Comme rappelé précédemment, à une fréquence de résonance, une structure présente une déformée modale avec un ou plusieurs ventres de vibration et une ou plusieurs lignes nodales (*cf.* **figure 1.3**). Si l'on calcule l'auto-spectre des signaux de vibrations mesurés sur une surface, les amplitudes résultantes à une fréquence de résonance sont alors liées aux ventres et aux nœuds de vibrations de la déformée modale correspondante. Par conséquent, en calculant l'écart-type des différents auto-spectres à chaque fréquence, des pics apparaissent aux fréquences de résonance, correspondant à des déformées modales. La **figure 1.9** présente un exemple de courbe calculée à partir de mesures réalisées sur un tambourin. Des pics sont visibles à environ 130 Hz, 230 Hz, 250 Hz, 410 Hz, 480 Hz ou 730 Hz par exemple. Chacun correspond à une déformée modale de la membrane du tambourin, qui peut alors être affichée avec les outils présentés ci-dessous.

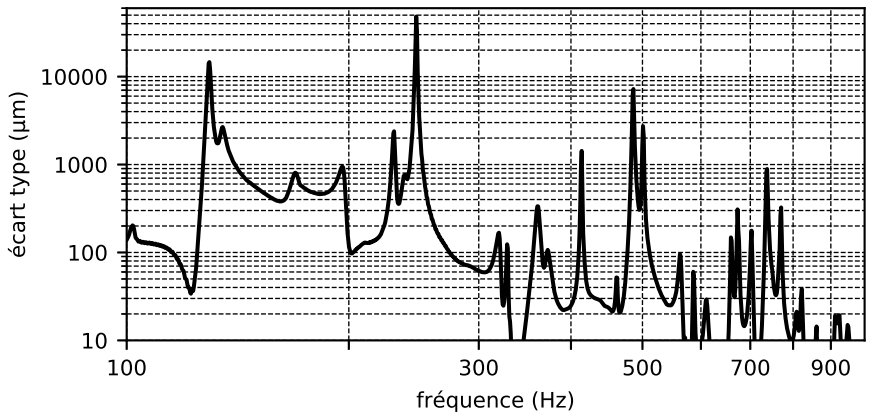


FIGURE 1.9 – Écart-type de l'auto-spectre des signaux mesurés pour un tambourin

Méthodes de tracé d'une déformée modale

Une fois la fréquence d'une déformée modale obtenue, l'affichage se fait soit à partir de la transformée de Fourier (**section 1.3.1**) des signaux mesurés à chaque point, si aucun signal de référence

n'est disponible, soit à partir des fonctions de réponse en fréquence si un signal de référence est à disposition. Les données spectrales présentent, dans tous les cas, une amplitude et une phase pour chaque point de mesure à chaque fréquence. Trois méthodes d'affichage sont présentées ici :

— l'affichage du module et de la phase ;
— l'affichage d'un module signé avec les données de phase ;
— l'animation de la déformée modale, qui génère plusieurs clichés pour une vidéo par exemple.

Afin d'expliquer le tracé d'une déformée modale opérationnelle, à titre d'illustration, un mode théorique d'une plaque rectangulaire à conditions aux limites libres sera utilisé (*cf.* **figure 1.4**).

Affichage du module et de la phase Il s'agit d'une manière simple d'afficher les résultats, mais elle aboutit à deux tracés ou deux cartes, une pour le module et l'autre pour la phase, qui ne sont pas toujours évidentes à interpréter. Sur la **figure 1.10** correspondante, la carte du module (à gauche) affiche les deux ventres (en *couleur*) et la ligne nodale (en *noir*). La carte de phase (à *droite*) permet de visualiser les valeurs des deux ventres de vibration, soit 0°(en *noir*) au dessus de la ligne nodale et 180°(en *blanc*) en dessous.

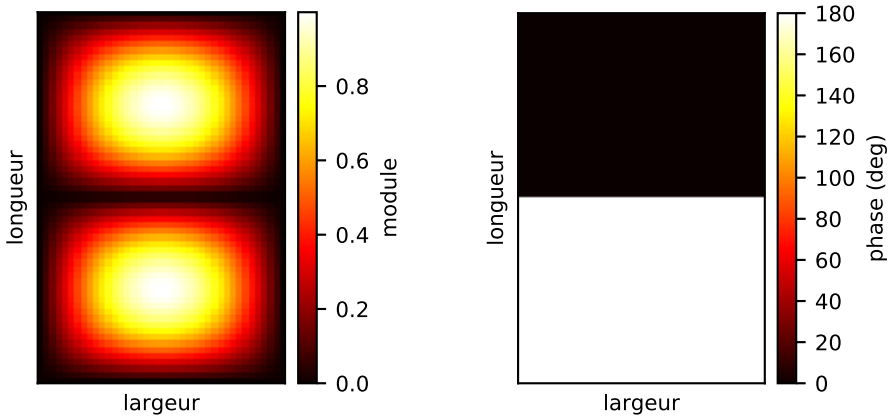


FIGURE 1.10 – (*gauche*) module du mode, (*droite*) phase du mode

Module signé Dans les cas où la phase est simple, du type 0 ou π, on observe des variations franches, non lisses. Le module peut alors être multiplié par -1 dans le cas où la phase est décalée de plus de 90° par rapport à la phase du point pour lequel le module est maximal. Sur la **figure 1.11**, le module est inchangé au-dessus de la ligne nodale (par rapport à la **figure 1.10**) et multiplié par -1 lorsque la phase est égale à 180°. On choisit en même temps un code couleur afin d'établir la barre de couleurs associée : une échelle de couleurs pour le positif et une échelle de couleurs pour le négatif, la ligne nodale, à 0, étant généralement en noir par simplicité.

Animation de la déformée modale et obtention d'un cliché (ou d'une vidéo) Si la phase du mode ne présente pas des délimitations franches, pour un mode tournant sur une surface circulaire par exemple, alors le module signé ne sera pas approprié. La manière la plus parlante pour visualiser la déformée modale, dans ce cas, est de réaliser plusieurs clichés à différentes phases du mode. A chaque point, la donnée spectrale complexe, obtenue par la fonction de réponse en fréquence ou la transformée de Fourier, est multipliée par $e^{j\theta}$ pour θ allant de 0 à 2π. On utilise alors la partie réelle du complexe résultant. Sur la **figure 1.12**, le déplacement de chaque point décrit une sinusoïde dans le temps. Il reste à générer la barre de couleurs selon le principe expliqué ci-dessus (ici, la barre de couleurs est la même que pour la **figure 1.11**).

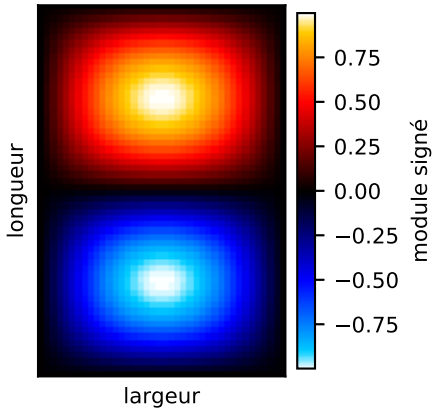


FIGURE 1.11 – Module signé du mode

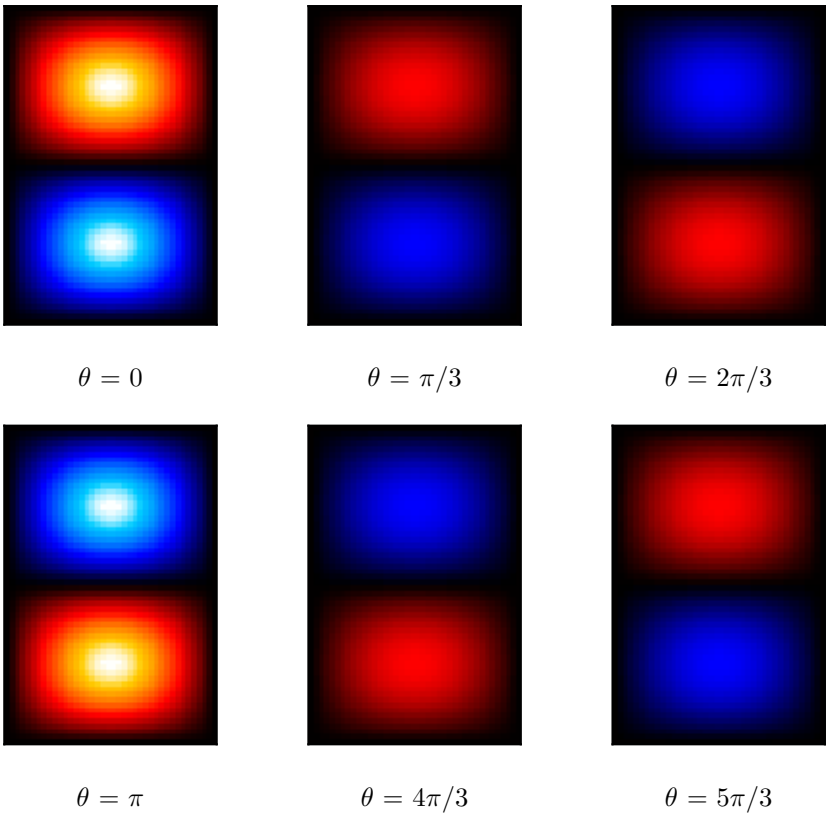


FIGURE 1.12 – Exemple de clichés d'un mode pour différentes valeurs de θ

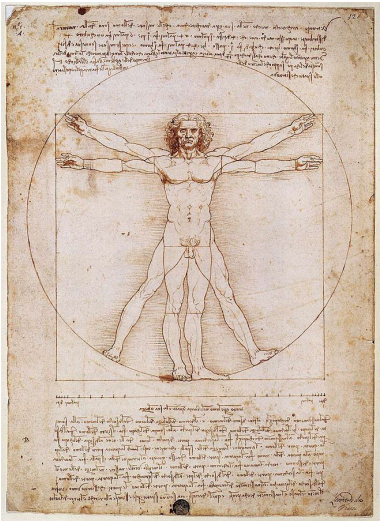
1.4 Conclusion

Les vibrations répondent à des lois précises, plus ou moins complexes. La compréhension du comportement dynamique des structures utilisées est essentielle et passe par leur caractérisation et l'extraction d'informations. Le traitement et post-traitement de ces données sont une clef fondamentale pour l'analyse. En effet, tout l'enjeu consiste à obtenir des informations viables et compréhensibles, d'où l'importance accordée ici aux outils de traitement du signal et d'affichage des déformées modales.

Enfin, comme l'exposera le chapitre suivant, les méthodes basées sur la stéréo-vision mesurent des positions, et donc des déplacements (*cf.* **section 2.6**). Le déplacement vibratoire diminuant en fonction du carré de la fréquence, leur applicabilité est logiquement meilleure dans les basses fréquences. Même si la plage de vibration des objets d'étude s'étend de 20 Hz à 20 kHz (*i.e.* au domaine audible), on s'attend ainsi à une limite de mesure aux alentours de quelques kHz maximum, en fonction de l'objet étudié et des caractéristiques des caméras actuelles. De plus, les objets choisis (pour rappel : une plaque, un haut-parleur, un tambourin et un capot de voiture) ont des déplacements allant de moins d'un micromètre jusqu'au millimètre. Ces ordres de grandeur sont à prendre en compte lors de l'application des méthodes de vision choisies pour ce travail. Rappelons finalement que les mesures de vibrations sont souvent utilisées pour remonter aux paramètres des objets de mesure. Dans le cadre de ce travail, l'étape finale est volontairement la visualisation des déformées modales, le sujet de cette thèse n'étant pas l'analyse modale *per se*. Le chapitre suivant propose un bref état des lieux, non exhaustif, des techniques de mesure de vibrations actuellement utilisées.

Chapitre 2

Etat des lieux des méthodes de mesure de vibrations



« Sans la curiosité de l'esprit, que serions-nous ? Telle est bien la beauté et la noblesse de la science : désir sans fin de repousser les frontières du savoir, de traquer les secrets de la matière et de la vie sans idée préconçue des conséquences éventuelles. »
Marie Curie

Sommaire

2.1 Introduction **23**
2.2 Techniques ponctuelles de référence **24**
2.3 Méthodes d'interférométrie **24**
2.4 Méthodes acoustiques **25**
2.5 Bref aperçu d'autres approches **25**
2.5.1 Déflectométrie 25
2.5.2 Projection de franges 25
2.5.3 Vibromètre laser à balayage continu 26
2.5.4 Méthode RIFF 26
2.6 Photogrammétrie et vision 3D **26**
2.6.1 Un peu d'histoire... 26
2.6.2 Quelques applications 27
2.6.3 Avantages et inconvénients 27
2.7 Conclusion **28**

2.1 Introduction

Comme rappelé en introduction générale, les vibrations sont omniprésentes au quotidien, avec les potentielles nuisances sonores ou vibratoires associées. Afin de les maîtriser au mieux, il est nécessaire de caractériser les structures vibrantes et donc, en amont, de comprendre les comportements dynamiques correspondants en mesurant des déplacements. Qui dit mesure dit, inévitablement... méthodes de mesure. Ce chapitre tente ainsi de faire un état des lieux des techniques, issues de différentes disciplines, actuellement utilisées pour la mesure de vibrations.

Dans un premier temps, les méthodes de mesure ponctuelles sont brièvement abordées (**section 2.2**). Sont ensuite décrites, de manière non-exhaustive, les principales techniques de mesure plein champ provenant de divers champs d'étude : interférométrie (**section 2.3**), holographie acoustique (**section 2.4**), déflectométrie, projection de franges ou encore méthode RIFF (**section 2.5**), chaque section s'appliquant à cerner les avantages et contraintes de chacune d'entre elles. Finalement, la **section 2.6** aborde

le domaine de la photogrammétrie et des méthodes de vision 3D, cadres méthodologiques de cette thèse. Un bref rappel historique précède une description des diverses applications de ces techniques, afin de cerner leurs potentiels, avantages et inconvénients et de mieux comprendre l'inscription de ce travail dans ce contexte.

2.2 Techniques ponctuelles de référence

Historiquement, la première technique utilisée pour la mesure de vibration a été une jauge de déformation de fil métallique collée sur une surface (1938) [120]. Lorsque le long fil qui la constitue est allongé, la résistance mesurée à ses extrémités varie, ce qui permet de mesurer indirectement la déformation de l'objet.
Ensuite, dans les années 1960, après le développement du laser He-Ne [170], sont apparus les premiers vibromètres laser à effet Doppler [179]. Limitée à la mesure de la vitesse d'un point, la méthode fut améliorée dans les années 1980 par l'invention du vibromètre laser Doppler à balayage [150], qui permet d'effectuer, successivement, la mesure sur un grand nombre de points de manière automatique, réduisant ainsi le temps de mesure. Le vibromètre laser Doppler est largement utilisé pour diverses applications industrielles et de recherche [29, 48]. Actuellement, un système de trois vibromètres laser, montés sur un bras robotisé, permet de mesurer, successivement, la vitesse vibratoire de surfaces à géométries complexes dans les trois axes de l'espace. La mesure est cependant faussée si l'objet subit un déplacement de corps rigide hors de l'axe de la vibration.
Enfin, l'accéléromètre a été développé dans les années 1920 et a rapidement été appliqué à divers contextes, pour les ponts, les dynamomètres ou encore l'aviation [163]. Aujourd'hui omniprésent dans les produits grand public, comme les téléphones, les air-bags des voitures ou encore les consoles de jeu, ce capteur permet de mesurer précisément l'accélération locale d'une surface suivant un ou plusieurs axes de l'espace.

De nature fondamentalement ponctuelle, la mesure effectuée par ces capteurs est très précise dans une grande bande de fréquences mais elle reste nécessairement limitée à des phénomènes stationnaires ou répétables si l'on veut effectuer une mesure plein champ. Le prix de ces systèmes de mesure est très variable : si le coût des accéléromètres et des jauges de déformation reste modique, celui des lasers à balayage 1D avoisine les cent cinquante mille euros, alors que le vibromètre laser 3D atteint le million d'euros s'il est monté sur un bras robotisé permettant la mesure d'objets aussi volumineux qu'une voiture. Finalement, ces techniques ponctuelles servent de référence pour la mesure de vibrations. Un accéléromètre et des vibromètres laser ont ainsi été utilisés pour évaluer la précision des méthodes présentées dans ce manuscrit. Les sections suivantes visent maintenant à rappeler succinctement les méthodes plein champ existant pour la mesure de vibrations.

2.3 Méthodes d'interférométrie

Dans les années 1980, le développement des caméras numériques a permis l'essor de méthodes d'interférométrie optique plein champ, initialement axées sur la mesure de forme, et dont l'applicabilité à la mesure de vibrations a tout d'abord été étudiée avec l'effet stroboscopique. Par la suite, lorsque les caméras ultra-rapides ont été suffisamment performantes, ces études ont abouti à des méthodes de mesure plein champ dynamique permettant de mesurer, sur toute une surface, des phénomènes transitoires, non-répétables, et possiblement non-linéaires. Les mesures sont effectuées grâce à l'utilisation d'une caméra ultra-rapide monochrome, d'un puissant laser et de composants optiques afin de calculer une différence de phase entre une onde lumineuse réfléchie par une surface ou traversant un fluide, et une onde de référence. Des mesures plein champ 3D sont aussi possibles en utilisant trois lasers de longueurs d'ondes différentes et une caméra ultra-rapide couleur [35]. Parmi les techniques existantes, on peut citer l'holographie numérique [124], l'interférométrie de speckle électronique [34, 132], l'interférométrie micro-ondes [42] ou encore la shearographie numérique [148].

Ces méthodes permettent d'effectuer des mesures plein champ 1D, voire 3D, de déplacements de l'ordre de quelques dizaines de nanomètres. Cependant, le montage est très délicat et l'installation des composants et les réglages associés très chronophages, même si l'utilisation d'éléments optiques holographiques ou diffractifs permet de simplifier un peu le montage [7]. Des objets d'étude suffisamment réfléchissants sont généralement requis (sauf pour l'interférométrie micro-ondes qui affiche moins de contraintes à ce niveau), de même qu'une table optique : ces techniques ne sont utilisables qu'en laboratoire, dans un cadre très réglementé de par l'utilisation des lasers. Les mesures sont limitées à des objets d'étude de moins de 30 cm, avec des déplacements maximaux de quelques micromètres entre deux images. Les mesures peuvent aussi être très perturbées si la surface mesurée bouge selon un axe autre que celui de la vibration. Ces techniques nécessitent un post-traitement lourd [124] et potentiellement très long si l'on n'opte pas pour une programmation fine sur un processeur GPU puissant, d'où une adéquation préférentielle avec des phénomènes de courte durée. Finalement, le coût des composants est élevé puisqu'un, voire trois lasers, et une caméra ultra-rapide, monochrome ou couleur, sont utilisés.

2.4 Méthodes acoustiques

Dans le domaine acoustique, différentes méthodes s'appliquent à caractériser des sources de bruits, en utilisant les données enregistrées par des microphones proches d'une structure vibrante. On peut notamment citer l'holographie acoustique, dans le domaine fréquentiel [100, 174] ou temporel [59, 159, 158]. D'autres méthodes acoustiques ont également été proposées, comme l'holographie acoustique statistiquement optimisée (SONAH) [60], les méthodes de sources équivalentes dans le domaine fréquentiel [133, 138] ou temporel [3] ou encore l'imagerie par retournement temporel [8, 91].

Ces techniques ont l'avantage de proposer des mesures plein champ avec une bonne sensibilité pour les phénomènes perceptibles, *i.e.* les bruits dans le domaine de l'audible, de par l'utilisation d'une antenne microphonique avec un nombre conséquent de microphones. Néanmoins, ce sont des méthodes de mesure indirectes : le principe de la rétro-propagation permet de remonter à la vibration de la structure mais le bruit acoustique ambiant ou encore le bruit intrinsèque des capteurs perturbent, de façon logique, très fortement les mesures. Les déplacements basses fréquences, qui ne génèrent que très peu de niveau sonore, ne sont, eux, pas mesurables. Finalement, une antenne microphonique est nécessaire et son prix varie en fonction du nombre et de la qualité des microphones insérés, de quelques dizaines à quelques centaines de milliers d'euros pour les plus performantes.

2.5 Bref aperçu d'autres approches

2.5.1 Déflectométrie

La déflectométrie [14] est utilisée pour mesurer la forme d'objets très réfléchissants, à partir du reflet d'un objet ou d'une lumière comportant une structure régulière, par exemple une grille. Cette contrainte forte sur l'objet semble, à première vue, limiter l'applicabilité à la mesure de vibrations et les tentatives en cours montrent la complexité du transfert des algorithmes pour cette finalité.

2.5.2 Projection de franges

Cette méthode consiste à projeter des franges sinusoïdales sur un objet en vibration. À l'aide de traitement d'images, une carte de déphasage, liée à la hauteur de l'objet, est obtenue pour chaque image acquise. Couplée à une caméra ultra-rapide, cette technique permet d'obtenir une mesure plein champ dynamique de profondeur, et a notamment été appliquée à la mesure de déformées vibratoires d'un haut-parleur [78]. Il est aussi possible d'utiliser l'effet stroboscopique et une caméra standard si le phénomène étudié est stationnaire [78]. Néanmoins, la mesure est nécessairement effectuée dans un seul axe de l'espace, et est facilement perturbée si l'objet subit des déplacements dans une direction autre que celle de l'axe de vibration.

2.5.3 Vibromètre laser à balayage continu

Actuellement à l'étude, la vibrométrie laser à balayage continu [24] utilise, comme son nom l'indique, un vibromètre laser qui effectue un balayage lent d'une surface afin d'estimer sa fonction continue de déformation. L'objectif est d'éviter la succession de mesures de milliers de points pour réduire le temps de mesure et obtenir des données spatialement continues de la vibration d'une surface, à condition que la vibration soit stationnaire.

2.5.4 Méthode RIFF

La méthode inverse RIFF [119] utilise des mesures effectuées sur quelques points, avec cinq accéléromètres en simultané par exemple pour une poutre, ou encore treize pour une plaque, afin d'identifier les efforts exercés sur une structure plane, caractériser la source vibratoire et calculer la distribution de la force excitatrice. La méthode a un coût très modique (le prix des capteurs) mais s'avère très sensible aux bruits de mesure. Elle doit donc être régularisée avec des outils de traitement du signal (fenêtrage et filtrage) et est actuellement limitée à des géométries simples.

2.6 Photogrammétrie et vision 3D

La photogrammétrie et *a fortiori* la vision 3D étant les cadres méthodologiques de ce travail, cette section est volontairement plus détaillée afin de permettre au lecteur de cerner plus précisément le contexte dans lequel ce travail de thèse s'est développé.

2.6.1 Un peu d'histoire...

Historiquement, les techniques de mesure utilisant des images appartiennent au domaine de la photogrammétrie [151]. Orienté vers la cartographie pour des applications originellement militaires puis industrielles, ce champ d'étude vise à obtenir une précision maximale en un minimum de clichés. Le perfectionnement des méthodes photographiques et l'apparition des caméras numériques ont permis son essor et ont donné naissance, dans les années 1980, à la branche de la photogrammétrie numérique. Les recherches correspondantes s'attachent alors à automatiser et optimiser les algorithmes de détection avec la méthode du suivi de cibles, notamment, et une ou deux caméras [143], ce qui laisse entrevoir des potentielles applications pour la mesure de vibrations.

En parallèle, on observe l'apparition d'un champ d'investigation connexe [64], lié au développement des photographies numériques : celui de la vision par ordinateur, qui cible, quant à lui, la rapidité des algorithmes d'extraction automatique de détails dans des images et les mesures volumétriques à partir d'images acquises de deux points de vue, permettant une mesure stéréoscopique par triangulation [64]. L'enjeu est le repérage dans l'espace proche, en temps réel, avec des applications directes dans le domaine de la robotique principalement [75, 54]. La stéréo-vision est mise au service de la rapidité de traitement de l'information, dans une logique de précision relative, au dépens de la précision absolue recherchée dans le cadre de la photogrammétrie. Se sont développés, en lien, des outils de corrélation 2D (DIC) et de stéréo-corrélation 3D (SDIC) d'images numériques [152, 101, 113, 151], dans une quête d'automatisation et de précision pour la métrologie mécanique. En découlent diverses techniques qui ont tenté d'appliquer ces éléments fondamentaux à la mesure industrielle de forme, de déformation et de contraintes de structures, et, à terme, à la mesure de vibrations et du comportement dynamique des matériaux par exemple [127, 67]. Les premiers essais avec des stroboscopes ont vite cédé la place à l'utilisation d'une caméra ultra-rapide pour la mesure de déplacements 2D ou de deux caméras pour obtenir des mesures de positions ou de formes 3D, avec l'utilisation de cibles pour des mesures synchrones ou des motifs aléatoires naturels ou peints pour des mesures plein champ. Différentes méthodes sont alors apparues pour accélérer le traitement du grand nombre d'images et de réduire la quantité données en sortie [55, 77, 13]

2.6.2 Quelques applications

On note un nombre certain d'études menées afin d'appliquer ces techniques de mesure de forme ou de vibrations à divers objets d'étude. Citons tout d'abord l'utilisation du suivi de cibles, dans le cadre de la photogrammétrie, pour la mesure, en soufflerie, de différents systèmes rotatifs, constitués d'ailes ou de pales : pales d'hélices de moteur d'avion [146], d'hélicoptères [5, 95, 140], ou encore d'éoliennes [115] ou de turbines [167].

On peut aussi mentionner la vidéogrammétrie avec une caméra, qui, couplée à différentes techniques de détection, s'intéresse aux vibrations parallèles au plan image de la caméra. Le suivi de cibles a là encore été testé, cette fois pour la mesure de vibrations de cordes d'un violon [80] ou encore la mesure de vibrations de ponts [111, 88]. La détection de contours, quant à elle, a été utilisée pour les mesures sur des structures comme des poutres et des sections de tuyaux, ou encore des structures à plusieurs étages, voire couplée à un algorithme de magnification de mouvement visant à améliorer la précision de la mesure de phénomènes stationnaires [32, 178].

Des essais numériques ont enfin montré que la stéréo-vision, grâce au développement des caméras (ultra) rapides, était une approche pertinente pour la mesure de champs de déplacement d'objets [142].

Combinés à l'utilisation de deux caméras ultra-rapides, les outils de stéréo-corrélation d'images présentent un potentiel certain pour la mesure de déplacements 3D. Les travaux actuels ciblent notamment la recherche des sources d'erreurs liées aux caméras rapides [160], ou encore les sources de bruit [10]. La stéréo-corrélation d'images permet aussi d'effectuer des mesures de forme et de contraintes, liées à l'encastrement des objets d'étude par exemple, ce qui mène à l'amélioration des modèles numériques pour une adéquation plus fine entre les simulations et les données recueillies expérimentalement [166]. Les méthodes SDIC ont finalement été appliquées à la vibration pour des déplacements rotatifs, par exemple à la mesure 3D de déformées des pales d'une double hélice coaxiale [27], à la mesure de déformées modales ou de la réponse non-linéaire à des excitations aléatoires, ou encore à la mesure dynamique de stress sur une poutre encastrée d'un côté [10]. Ces techniques sont validées par la comparaison des résultats avec ceux obtenus par des méthodes de référence, soit des accéléromètres [167], soit des vibromètres laser (simples ou à balayage) [10, 66], avec des coefficients de corrélation pouvant atteindre plus de 99% [167]. La validation a aussi été obtenue par comparaison avec les résultats mesurés grâce à l'interférométrie de speckle électronique [141].

2.6.3 Avantages et inconvénients

La photogrammétrie et les méthodes de vision permettent ainsi d'effectuer une mesure d'un déplacement en 2D (avec une caméra) ou en 3D (avec deux caméras) et d'obtenir une mesure synchrone d'un certain nombre de points, avec l'utilisation de cibles, ou d'obtenir une mesure plein champ, possiblement sur toute la surface inscrite dans le champ de vision, si un motif aléatoire et des outils de corrélation d'images sont utilisés. Comme les autres techniques décrites dans les sections précédentes (à l'exception des capteurs physiques que sont l'accéléromètre ou la jauge de déformation), elles sont sans contact, à la condition que les cibles ou la peinture appliquées sur l'objet d'étude pour la corrélation d'images aient un impact négligeable sur le comportement dynamique de l'objet en question. La mesure de grandes dynamiques de déplacements est possible, sur des objets petits ou grands. La précision de la mesure dépend du montage réalisé et de la taille de l'objet mesuré, puisque la précision obtenue par ces méthodes de vision est relative à la taille de l'objet. Même si la précision devient plus faible pour des objets plus grands, les vibrations associées sont, dans ce cas, généralement plus fortes, donc la mesure est toujours possible, avec une précision relative à l'amplitude du déplacement.

Comme vu précédemment, la mesure est sensible aux déplacements (et non à la vitesse, comme c'est le cas, par exemple, pour un vibromètre laser), qui diminuent lorsque la fréquence augmente. Ces techniques sont donc plutôt adaptées à la mesure de basses fréquences. On notera, au passage, la possibilité de repliement spectral si les fréquences du signal mesuré excèdent la moitié de la cadence d'acquisition des images. Les caméras ont un coût élevé (environ soixante mille euros pour la Photron SA-X2, et les accessoires associés, utilisée pour les phases expérimentales), mais qui reste relativement équivalent voire en deçà du coût d'un vibromètre laser à balayage (de l'ordre de cent cinquante mille

euros pour celui utilisé en référence dans cette thèse). Notons néanmoins que les moyens financiers mis à disposition ne permettent pas toujours d'acquérir deux caméras, d'où les tentatives d'adaptation des protocoles à des montages pseudo-stéréoscopiques n'utilisant qu'une seule caméra ultra-rapide (cet aspect est traité au **chapitre 6**).

Finalement, un problème se pose généralement : ces mesures plein champ impliquent une grande quantité de données et nécessitent un post-traitement adapté, tel que proposé par Wang, Mottershead *et al.* par exemple [166]. Le temps de calcul est très long car des quantités colossales de données sont mobilisées par des logiciels de photogrammétrie souvent mal adaptés, n'intégrant pas forcément d'outils de traitement de signal [6]. Il est alors nécessaire d'exporter les données, souvent un fichier par image [70], contenant les données de mesure de position, vers un logiciel adapté à l'analyse des signaux vibratoires. La solution la plus simple consiste fréquemment à développer des codes de post-traitement adaptés [70]. D'autres approches ont été tentées, par exemple en utilisant les polynômes de Tchebichef afin de décrire les déformées modales et de réduire ainsi la quantité de données [166].

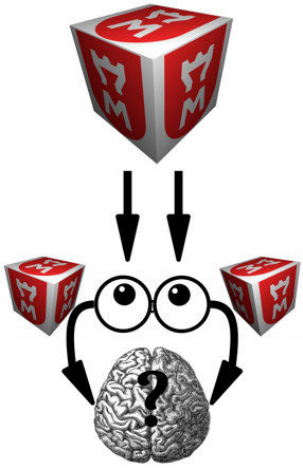
2.7 Conclusion

Comme mentionné précédemment, ce travail de thèse s'inscrit dans le cadre de la photogrammétrie dynamique et plus précisément de la vision 3D : les méthodes associées donnent des mesures de forme plein champ, qui, couplées à une ou deux caméras ultra-rapides, permettent d'avoir accès à la dimension temporelle du déplacement de la structure étudiée. En découle la possibilité d'effectuer des mesures plein champ, synchrones, applicables à des phénomènes transitoires, non répétables, possiblement non-linéaires. Ce cadre méthodologique est ainsi en adéquation avec la finalité de ce travail de recherche, à savoir la mesure de vibrations.

Les outils de corrélation d'images développés dans le cadre des méthodes de vision sont potentiellement très intéressants, mais les logiciels existants ne sont pas nécessairement adaptés à l'analyse des vibrations. En prenant comme point de départ les avantages indéniables présentés *supra*, une partie de ce travail a consisté à adapter finement le principe de la vision 3D et les outils associés à la mesure de vibration dynamique, et ainsi à simplifier les protocoles, trouver des solutions au problème des données et optimiser le traitement et le post-traitement. Le chapitre suivant vise à expliquer le principe sur lequel reposent ces méthodes : la stéréo-vision.

Chapitre 3

Stéréo-vision



« La vision est l'art de voir les choses invisibles. »
Jonathan Swift

Sommaire

3.1 Introduction **29**
3.2 Modèle de projection perspective : le modèle sténopé **30**
3.2.1 Image sur une caméra 30
3.2.2 Modèle de caméra et repère caméra associé 31
3.2.3 Calibrage d'une caméra 34
3.2.4 Distortions optiques 36
3.2.5 Repère lié à la mire de calibrage 38
3.3 Principe du capteur stéréoscopique **38**
3.3.1 Lien géométrique entre deux repères de l'espace 38
3.3.2 Calibrage d'un capteur stéréoscopique 41
3.3.3 Triangulation de la position 3D d'un point dans l'espace 43
3.3.4 Passage vers le repère objet 46
3.4 Conclusion **48**

3.1 Introduction

Comme rappelé aux **chapitres 1 et 2**, l'enjeu est de mesurer avec précision des petits déplacements rapides avec une méthode de stéréo-vision. Le principe consiste à imiter au mieux la vision naturelle de l'espace en trois dimensions, que possède un certain nombre d'animaux ainsi que l'homme. En effet, à partir d'images fournies par deux yeux en position frontale, le cerveau est capable, en utilisant des processus complexes et le principe de triangulation, de retrouver la position 3D approximative, avec une précision relative, des objets qui l'entourent.

Afin d'effectuer une mesure de vibration, *i.e.* une mesure dynamique, avec la méthode de stéréo-vision, il est auparavant nécessaire de savoir mesurer une forme en 3D, c'est-à-dire d'effectuer une mesure statique, avec un capteur stéréoscopique. Il faut donc comprendre le fonctionnement de ce capteur et des caméras qui le composent. Il faut aussi comprendre comment se repérer dans l'espace des deux caméras et comment opérer une triangulation à partir de ces données.

Ce chapitre rappelle ainsi les principes géométriques et algorithmiques liés à la stéréo-vision. Le concept de modèle sténopé est défini en premier (**section 3.2**). Cette section inclut la définition du repère image, du principe géométrique qui unit un point 3D à un point 2D sur le plan image d'une caméra. Y sont aussi expliqués le calibrage de la caméra, visant à obtenir ses caractéristiques, et

la façon de gérer les distorsions optiques associées. Dans un deuxième temps, le principe du capteur stéréoscopique est rappelé (**section 3.3**), à savoir : comment passer d'un repère à l'autre dans l'espace (par exemple d'une caméra à l'autre), comment obtenir l'orientation et la position relative des deux caméras, comment faire la triangulation d'une position 3D à partir des données de calibrage et d'une position donnée d'un point dans chaque image et finalement comment modifier les paramètres de calibrage pour trianguler dans un repère autre que celui de la caméra.

Il est à noter que ce chapitre n'a pas pour vocation d'être un cours complet de stéréo-vision. Le but est d'avantage de permettre aux néophytes du domaine de bien comprendre les principes essentiels de la stéréo-vision. Il ne se veut donc ni exhaustif, ni trop détaillé, mais suffisamment explicite pour permettre une bonne compréhension du contenu du travail de thèse réalisé.

3.2 Modèle de projection perspective : le modèle sténopé

L'objectif de cette section est de répondre aux questions suivantes :
— Comment fonctionne une caméra et son capteur ?
— Comment sont définis les repères associés ?
— Quel est le lien entre une image obtenue avec une caméra et l'espace à trois dimensions ?

3.2.1 Image sur une caméra

Une caméra à capteur matriciel présente une surface photosensible qui permet de transformer l'énergie lumineuse en signal électrique. Cette surface est divisée en pixels formant une matrice. Les valeurs d'intensité mesurées par cette matrice de pixels définissent une image.

Bruit photonique

L'énergie lumineuse mesurée correspond à un nombre de photons arrivant sur le capteur. Cependant, l'énergie reçue n'est pas un flux continu dans l'espace et le temps : le nombre de photons impactant la surface du capteur, et ensuite convertis en électrons, suit la loi de Poisson [65]. Le nombre d'électrons ne mesuré par le capteur est échantillonné pour obtenir une intensité lumineuse et correspond à la somme d'une valeur moyenne et d'un bruit gaussien b_{ne} d'écart type σ_{ne}. On a :

$$ne = \overline{ne} + b_{ne} \tag{3.1}$$

avec $\overline{ne}$ le nombre moyen d'électrons, et, pour l'intensité lumineuse associée :

$$I = \frac{ne}{ne_{full}} I_{max} \tag{3.2}$$

avec ne_{full} le nombre d'électrons correspondant à une intensité maximale, et I_{max} la valeur de niveau de gris maximale. De manière générale et pour les travaux présentés, les images sont codées sous huit bits, pour une plage de niveaux de gris allant de 0 à 255.

L'écart-type du nombre d'électrons mesuré correspond à la racine du nombre moyen d'électrons :

$$\sigma_{ne} = \sqrt{\frac{\overline{I}}{I_{max}} ne_{full}} \tag{3.3}$$

avec $\overline{I}$ l'intensité moyenne. Cette valeur peut être convertie en niveaux de gris avec l'équation (3.2). On obtient :

$$\sigma_I = \sqrt{\frac{\overline{I} I_{max}}{ne_{full}}} \quad . \tag{3.4}$$

avec σ_I l'écart-type du bruit, exprimé en niveau de gris. Le bruit sur les images dépend donc de l'intensité lumineuse moyenne reçue. Notons que l'équation (3.4) peut être utilisée afin de générer un bruit Gaussien réaliste dans des images simulées, même si les autres sources de bruit ne sont, ici, pas prises en compte.

Repère image

Comme rappelé ci-dessus, la surface photosensible d'un capteur est divisée en pixels, qui forment une matrice. Pour chacune des cellules, une mesure du niveau de gris est effectuée. La juxtaposition de ces cellules, aux niveaux de gris variables, aboutit à une image.

À cette matrice de pixels est associé un repère de coordonnées 2D, dit repère image (*cf.* **figure 3.1**). La position d'un pixel est définie comme (u, v), avec l'axe $\overrightarrow{u}$ horizontal (de gauche à droite) et l'axe $\overrightarrow{v}$ vertical (de haut en bas). L'indice $(0, 0)$ correspond au centre du pixel en haut à gauche de l'image.

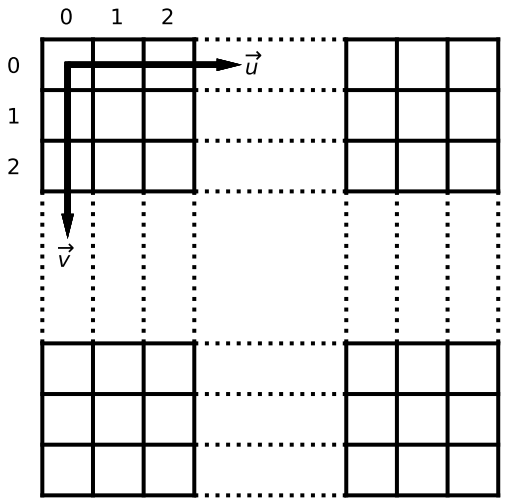


FIGURE 3.1 – Schéma définissant le repère image : l'axe $\overrightarrow{u}$ correspond au numéro de colonne, (de gauche à droite) et l'axe $\overrightarrow{v}$ correspond au numéro de ligne (de haut en bas) ; la numérotation des indices commence à 0 au centre du pixel en haut à gauche

3.2.2 Modèle de caméra et repère caméra associé

Une matrice de pixels photosensibles placée directement face à une scène ne donnerait qu'une valeur d'illumination globale. Afin d'obtenir des images nettes, un objectif optique est placé devant la matrice de pixels de la caméra (*cf.* **figure 3.2**). Il est constitué de lentilles et d'un diaphragme, centré sur l'axe optique, qui limite la quantité de lumière passant dans l'objectif.

La **figure 3.2** montre la lumière qui part de l'objet y et se concentre au centre de symétrie optique. Le diaphragme est placé proche de ce centre. La lumière se sépare ensuite et projette une image inversée y' de l'objet sur la surface photosensible. Le grossissement y'/y correspond au rapport s'/s, avec s la distance entre l'objet et le centre de symétrie et s' la distance entre le centre de symétrie et la surface photosensible. La distance s' correspond généralement à la focale de l'objectif afin de limiter le flou dans l'image (pour plus de précision *cf.* [129]). Les objectifs ne sont cependant pas parfaits et induisent des distortions optiques dans les images. Celles-ci sont supposées nulles dans cette section, pour simplifier l'exposé. Elles sont néanmoins expliquées en **section 3.2.4**, en même temps que la méthode pour les corriger.

Quel est le modèle géométrique correspondant ? Le modèle d'une caméra, dit modèle sténopé [43, 68, 64] (ou *pinhole* en anglais) correspond à une projection perspective, et suppose que le diaphragme soit extrêmement resserré et ne laisse passer qu'un rayon lumineux très fin. En pratique, le diaphragme est toujours un peu ouvert, ce qui peut, dans certaines conditions, induire un flou dans les images.

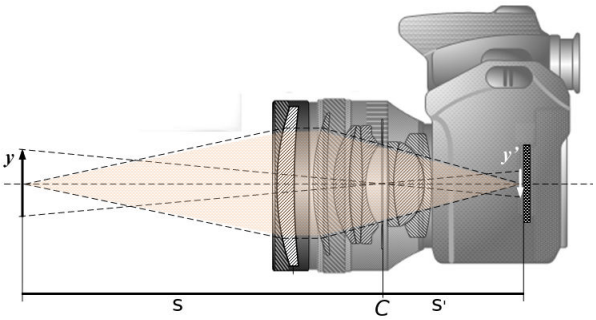


FIGURE 3.2 – Schéma de principe d'un objectif ; wikimedia commons, par *Tamasflex*, en 2011

Comme on peut le voir sur la **figure 3.2**, l'image qui est vue sur la surface photosensible est inversée par rapport à l'objet réel. Afin de simplifier la géométrie associée, le plan image est, de façon usuelle, placé de l'autre côté du centre de symétrie optique, aussi appelé centre de perspective, qui est utilisé comme centre du repère caméra et noté C. Sur la **figure 3.3**, le plan image est situé à une distance f (dite distance focale) du centre du repère caméra [1]. Selon le théorème de Thalès, le point (Z, Y) est projeté sur le plan image au point $(f, fY/Z)$.

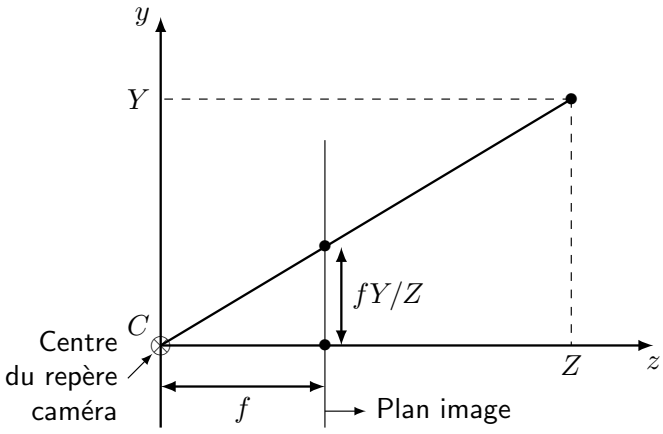


FIGURE 3.3 – Schéma du modèle de caméra sténopé

La position 2D (fx, fy), d'un point dans le plan image, correspondant en 3D à (fx, fy, f) dans le repère caméra, est obtenue à partir de la position (X, Y, Z) du point 3D correspondant dans le repère de la caméra avec l'équation suivante :

$$(fx, fy, f)^T = (fX/Z, fY/Z, f)^T \tag{3.5}$$

avec $.^T$ la transposée d'une matrice. Les paragraphes suivants expliquent, de façon plus détaillée, le lien entre le repère image et le repère caméra.

En photogrammétrie, deux points distincts servent à caractériser la caméra et ainsi définir le repère associé : le Point Principal d'Autocollimation (PPA) et le Point Principal de Symétrie (PPS) [143, 84]. Le PPS correspond au point présentant la distorsion optique minimale. Le PPA correspond à la

1. La distance focale, aussi appelée distance principale, dépend de la position de l'objectif par rapport au capteur matriciel et ne doit pas être confondue avec la longueur focale de l'objectif qui est lié à la disposition des lentilles qui le constituent.

projection orthogonale du centre de perspective sur le plan image, (u_0, v_0) dans le repère image, et est situé à une distance p nommée distance principale. Le centre de pesrspective est utilisé comme le centre du repère à trois dimensions lié à la caméra. La **figure 3.4** présente le schéma de principe (simplifié par rapport à la géométrie réelle, avec une inclinaison exagérée du plan image), valable si l'angle entre le plan image et l'axe optique est proche de 90°(selon l'approximation de Gauss).

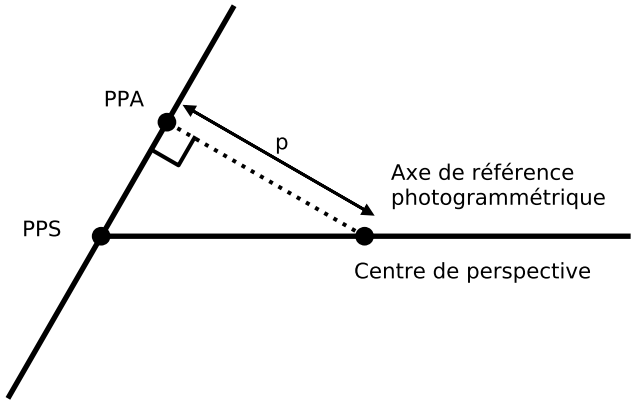


FIGURE 3.4 – Schéma de la position du PPS, du PPA et du centre de perspective

La manufacture des caméras et appareils photos de qualité est de nos jours très précise : ces deux points sont très proches. Dans le cadre de la vision 3D, ces deux points sont supposés confondus et appelés Point Principal (PP). Pour simplifier, on parle de distance focale f au lieu de distance principale p et d'axe optique au lieu d'axe de référence normal au plan image. Il est à noter que l'inclinaison du plan image est modélisée par l'utilisation de deux distances focales, f_u et f_v, soit une par axe dans le repère image. Les coordonnées (u, v) de l'image d'un point $(X, Y, Z)^T$ dans le repère de la caméra sont donc obtenues dans le repère image avec l'équation non-linéaire suivante :

$$(u, v) = (f_u X/Z + u_0, f_v Y/Z + v_0) \tag{3.6}$$

avec (u_0, v_0) les coordonnées du point principal dans le repère image, *i.e.* la projection du centre du repère caméra C dans le repère image. Les valeurs X/Z et Y/Z correspondent à des rapports de distances et sont donc sans dimension. Les focales f_u et f_v sont exprimées en pixels afin que u et v soient en pixels. Finalement, le repère 3D de la caméra est défini tel que (*cf.* **figure 3.5**) :

- les axes $\overrightarrow{x}$ et $\overrightarrow{y}$ soient colinéaires avec les axes $\overrightarrow{u}$ et $\overrightarrow{v}$ respectivement ;
- l'axe $\overrightarrow{z}$ soit aligné avec l'axe optique et intersecte le plan image au point (u_0, v_0) ;
- le centre du repère soit situé à une distance égale à la distance focale du point (u_0, v_0) du plan image.

De plus, afin de simplifier les écritures, les valeurs des distances focales ainsi que la position du centre du repère image sont stockées dans une matrice nommée matrice intrinsèque K avec :

$$K = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{3.7}$$

Dans ce manuscrit, les pixels sont supposés parfaitement carrés, le facteur de non-perpendicularité des pixels, c'est-à-dire le second terme de la première ligne de la matrice K, est donc fixé à zéro. L'équation (3.6) peut alors être réécrite :

$$Z * \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad . \tag{3.8}$$

Le vecteur $(u, v, 1)^T$ correspond aux coordonnées homogènes [43, 64] du point dans le repère image.

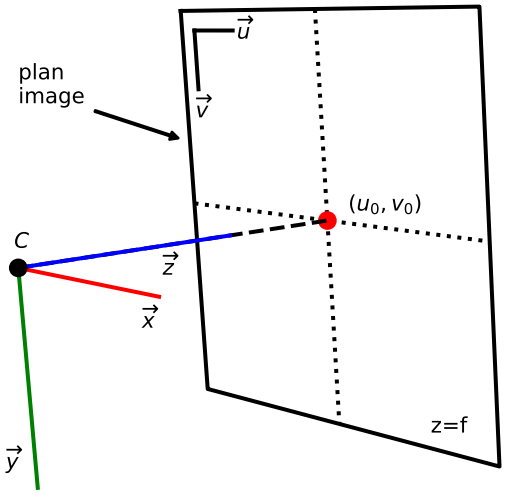


FIGURE 3.5 – Schéma présentant le repère 3D lié à la caméra ; l'axe $\vec{x}$ est aligné avec l'axe $\vec{u}$, l'axe $\vec{y}$ est aligné avec l'axe $\vec{v}$, l'axe $\vec{z}$ est perpendiculaire au plan image et la distance entre C et (u_0, v_0) est égale à la distance focale de l'objectif

Il s'agit maintenant de comprendre comment passer de la position 2D (u, v) d'un point à sa position 3D correspondante (X, Y, Z). Premièrement, on passe des coordonnées homogènes dans le repère image aux coordonnées homogènes dans le repère caméra selon l'équation :

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = K^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \tag{3.9}$$

avec :

$$K^{-1} = \begin{bmatrix} 1/f_u & 0 & -u_0/f_u \\ 0 & 1/f_v & -v_0/f_v \\ 0 & 0 & 1 \end{bmatrix} . \tag{3.10}$$

Deuxièmement, en reprenant l'équation (3.6), les coordonnées homogènes $(x, y, 1)^T$ d'un point dans le repère caméra correspondent à $(X/Z, Y/Z, 1)^T$. La position 3D (X, Y, Z) est alors obtenue en multipliant les coordonnées homogènes dans le repère caméra par Z. Si l'on ne connait pas la distance Z, un point (u, v) peut seulement être associé à la droite passant par le centre du repère C et le point $(X/Z, Y/Z, 1)^T$. La section suivante explique comment d'obtenir les valeurs des différents paramètres de la caméra, à savoir la ou les distances focales, la position du PP dans l'image ainsi que les coefficients de distortions optiques (*cf.* **section 3.2.4**).

3.2.3 Calibrage d'une caméra

Terminologie utilisée

La littérature internationale sur le sujet utilise le mot anglais *calibration*, dont la définition, donnée par le *Collins Cobuilt* est : « from *calibrate* - a technical term ; if you calibrate an instrument or tool 1. you mark it so that you can use it to measure something accurately 2. you correct or adjust it so that it measures accurately ». En consultant tout d'abord le *Manuel terminologique didactique de télédétection et photogrammétrie - français/anglais*, la traduction donnée est *étalonnage*. On peut lire : « [anglais] *calibration* Physique, métrologie - Détermination de la relation entre les indications de l'appareil de mesure et les valeurs du mesurande (grandeur à mesurer). 1. Le terme calibration ne doit pas être employé pour désigner l'étalonnage. 2. Cet étalonnage peut être effectué par rapport à une source de référence intérieure ou extérieure au système ». Le terme *étalonnage* est ainsi utilisé en photogrammétrie et renvoie à l'histoire de la discipline, même si un étalon certifié n'est plus obligatoirement

utilisé. *A contrario*, au sein de la communauté vision 3D, on trouve le plus souvent le terme *calibrage* dans des publications françaises émanant de spécialistes du domaine [49, 113]. En consultant le site *France*Terme [2], « consacré aux termes publiés au *Journal officiel* de la République française par la Commission d'enrichissement de la langue française [...] pour nommer en français les réalités nouvelles et les innovations scientifiques et techniques [...] et regroup[ant] un ensemble de termes recommandés de différents domaines scientifiques et techniques, [...] [dont] l'emploi [...] s'impose à l'administration », on trouve la rubrique suivante :
« *Journal Officiel* du 22/09/2000 - *calibrage*, n.m. du domaine physique et métrologie : détermination de l'appartenance d'une mesure à une classe d'intensité comprise entre deux valeurs limites prédéterminées ; équivalent étranger *calibration* (en) ».
Au vu des éléments présentés ci-dessus, et dans le mesure ou ce travail se rattache davantage au domaine de la vision 3D qu'à celui de la photogrammétrie, c'est ce néologisme qui sera utilisé dans ce manuscrit.

Techniques de calibrage

L'objectif est de déterminer les valeurs de la ou des focales, la position du PP dans le repère image ainsi que les coefficients de distorsions optiques. Ces informations sont regroupées sous le terme paramètres intrinsèques ou encore paramètres internes de la caméra. Il s'agit de calibrer la caméra en estimant ces paramètres, cette étape étant nécessaire pour obtenir ensuite des mesures métriques avec un capteur stéréoscopique (*cf.* **section 3.3**). Les méthodes les plus simples et les plus précises pour effectuer ce calibrage utilisent une mire [49, 113], même si d'autres techniques existent, par exemple les méthodes d'auto-calibrage [64, 38].

Le calibrage est possible si l'on connait des points 3D exprimés dans un repère quelconque, ainsi que les coordonnées 2D de ces points projetés dans le repère image. On parle alors de calibrage fort. En pratique, on utilise les images d'une mire de calibrage (ou d'étalonnage si elle est certifiée par un organisme reconnu), avec un marquage dont la position 3D est connue. Les premières méthodes proposées utilisaient une mire de calibrage précise dite volumique, avec des marqueurs répartis dans un volume (*cf.* **figure 3.6** où la position des marqueurs est définie par les croisements des lignes). On peut aussi citer les travaux de Besnard [15], où la mire de calibrage volumique est en livre ouvert. Ce type de mire permet de réaliser l'estimation avec une seule image, mais la complexité pour réaliser une telle mire avec précision a conduit à l'apparition d'autres méthodes.

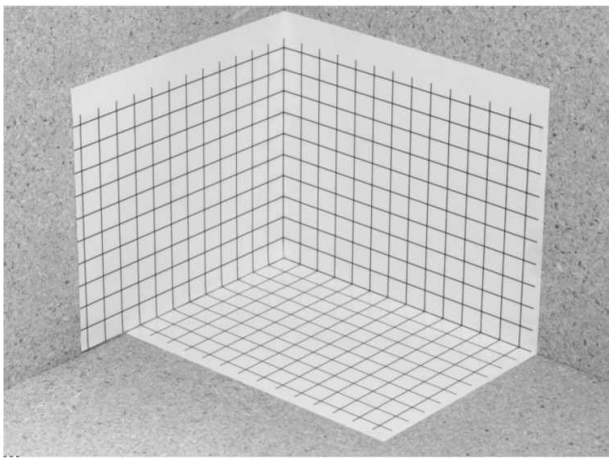


FIGURE 3.6 – Exemple de mire volumique (tiré de [49])

2. http ://www.culture.fr/Ressources/FranceTerme

Ce deuxième groupe de techniques utilise une mire plane. Plusieurs clichés de la mire, dans différentes positions et orientations (*cf.* **figure 3.7**) sont utilisés pour estimer les paramètres internes de la caméra, ainsi que sa position et son orientation relative au repère lié à la mire (*cf.* **section 3.2.5**) pour chaque image, ainsi que la position 3D des points de la mire (en posant certaines contraintes afin d'éviter les estimations aberrantes *cf. infra*).

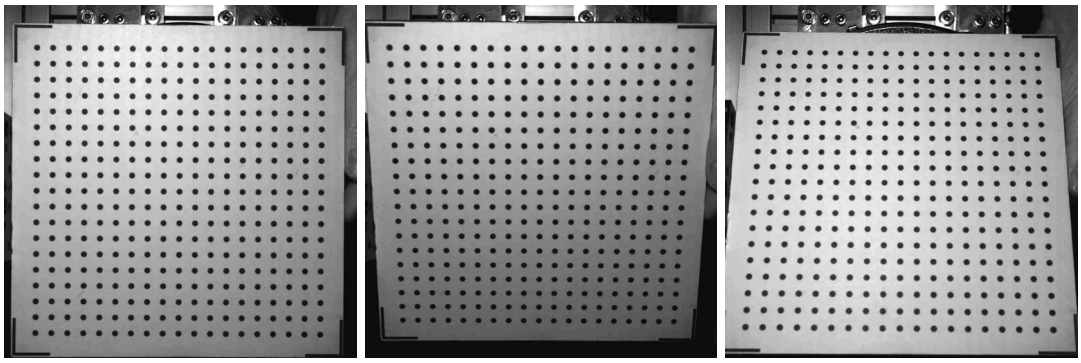


FIGURE 3.7 – Exemples de clichés d'une même mire utilisés pour le calibrage d'une caméra

Algorithme de Zhang

Une des méthodes utilisant une mire plane, déjà implémentée sous OpenCV [20], repose sur les algorithmes de Zhang [185] et Bouget [19]. Cette technique ayant été choisie pour ce travail de recherche, elle est exposée brièvement ci-dessous.

Les positions 2D, dans le repère image (u, v), des cibles d'une mire de calibrage plane sont préalablement estimées. Les positions 2D des cibles dans le plan de la mire sont connues avec une précision suffisante (par exemple de par leur impression sur une imprimante laser) et la troisième dimension (Z) est initialisée à zéro pour tous les points. À partir de ces données, la position et l'orientation relative de la caméra par rapport à la mire, ainsi que les distances focales et la position du PP sont estimées pour chaque image (pour plus de détails sur les calculs *cf.* [185]). Tous les coefficients de distortions sont initialisés à zéro. Ces données sont ensuite optimisées, à l'aide de la méthode numérique de Levenberg-Marquardt [125], afin de minimiser l'erreur de re-projection, *i.e.* la distance entre les points 2D détectés et les points 3D estimés de la mire, re-projetés dans le repère image pour chaque cliché. Trois contraintes sont alors imposées afin d'éviter les instabilités numériques et les solutions aberrantes :

— la position 3D d'une des cibles est fixe dans le processus d'optimisation ;
— la distance entre ce point fixé et un autre point est aussi fixe ;
— ces deux points et un troisième définissent le plan de la mire, *i.e.* leur altitude est fixée à zéro.

La précision de l'estimation dépend de la précision de la mire mais aussi du nombre de points d'appui (*i.e.* des cibles) et de leur volume dans l'espace (si les points ont des profondeurs différentes). Zhang recommande alors de prendre des clichés de la mire avec plusieurs angles différents entre la mire et la caméra [185].

3.2.4 Distortions optiques

Comme vu précédemment, les objectifs des caméras ne sont pas parfaits et induisent, dans les images, des distortions optiques, dont les coefficients doivent être estimés lors de la phase de calibrage de la caméra. Cette section définit les distortions optiques radiales et tangentielles, le modèle utilisé pour le calibrage et la méthode de correction appliquée.

Modélisation des distortions optiques

Différents modèles existent [41, 143, 169, 22]. Un modèle est, là encore, implémenté par OpenCV et prend en compte les distortions optiques radiales et les tangentielles. C'est cet outil qui a été choisi, d'une part par simplicité, d'autre part parce qu'il s'avère suffisant par rapport aux besoins de ce travail et qu'il ne génère pas d'erreurs visibles dans les résultats obtenus. Une présentation détaillée en est donnée ci-dessous.

Pour un point de coordonnées homogènes $(x, y, 1)$ du repère caméra (*cf.* équation (3.9)), considérées comme sans distorsion, on définit $r = \sqrt{x^2 + y^2}$, correspondant à la distance entre le point et la projection du PP sur le plan des coordonnées homogènes. La position de ce point, prenant cette fois en compte les distortions optiques, notée (x_d, y_d), est alors modélisée par les équations :

$$x_d = x\frac{1 + k_1r^2 + k_2r^4 + k_3r^6}{1 + k_4r^2 + k_5r^4 + k_6r^6} + 2p_1xy + p_2(r^2 + 2x^2) \tag{3.11}$$

$$y_d = y\frac{1 + k_1r^2 + k_2r^4 + k_3r^6}{1 + k_4r^2 + k_5r^4 + k_6r^6} + p_1(r^2 + 2y^2) + 2p_2xy \tag{3.12}$$

avec les coefficients k_1 à k_6 correspondant à la modélisation de les distortions optiques radiales et les coefficients p_1 et p_2 à la modélisation des distortions optiques tangentielles. Les coordonnées du point dans le repère image, prenant en compte les distorsions optiques, sont alors :

$$(u_d, v_d, 1)^T = K\,(x_d, y_d, 1)^T \quad . \tag{3.13}$$

Dans son article [185], Zhang affirme qu'il est fort probable que la fonction de distortions soit totalement dominée par les composantes radiales, et plus spécifiquement par le premier terme. Il rappelle qu'avec les objectifs optiques actuels il a été prouvé que des modèles plus élaborés non seulement n'amélioraient pas la précision (leur apport étant négligeable par rapport à la résolution du capteur), mais aussi qu'ils menaient à des instabilités numériques si les valeurs des distortions supplémentaires modélisées étaient petites [161, 56]. Zhang préconise alors de n'utiliser que les coefficients k_1 et k_2 du modèle présenté sur les équations (3.11) et (3.12).

Le modèle finalement utilisé dans ce travail est le suivant :

$$x_d = xr_d/r \tag{3.14}$$

et

$$y_d = yr_d/r \tag{3.15}$$

avec :

$$r_d = r(1 + k_1r^2 + k_2r^4) \quad . \tag{3.16}$$

Correction des distortions optiques

À partir des coordonnées homogènes (u_d, v_d) mesurées dans le repère image, la méthode consiste à :

1. calculer les coordonnées homogènes $(x_d, y_d, 1)^T$ avec l'équation (3.9) ;
2. initialiser $x_u^{[0]} = x_d$ et $y_u^{[0]} = y_d$, $i = 1$ et $r^{[0]} = \sqrt{x_u^{[0]2} + y_u^{[0]2}}$;
3. calculer $r_u^{[i]} = r_d/(1 + k_1r_u^{[i-1]2} + k_2r_u^{[i-1]^4})$;
4. vérifier si $f_u * |r_u^{[i]} - r_u^{[i-1]}| > pixLim$; si oui, incrémenter i de 1 et boucler sur l'étape 3 ;
5. calculer $(u_u, v_u) = f_ur_u^{[i]}/r_d(x_d, y_d) + (u_0, v_0)$,

avec $|.|$ correspondant à la valeur absolue, $pixLim$ la condition limite de sortie (*i.e.* la différence limite entre deux itérations, en pixels) et $.^{[i]}$ la valeur d'une variable à l'itération i. Le processus converge après quelques itérations [49]. Le critère d'arrêt $pixLim$ peut alors être défini à 1/1000 pixel ou 1/10 000 pixel par exemple, en fonction de la précision recherchée.

3.2.5 Repère lié à la mire de calibrage

Cette section définit maintenant le repère lié à la mire, par le biais des choix opérés dans ce travail de recherche et des exemples de mires utilisées (*cf.* **section 5.3**). Notons que ces mires sont constituées d'un autocollant collé sur une plaque en aluminium, dont l'image imprimée a été créée au préalable sur ordinateur. Les positions des cibles sont alors connues, à l'erreur de collage et d'impression près, soit $\approx 0{,}2$ mm pour 300 dpi. Deux principaux motifs sont utilisés pour tracer des cibles : des lignes croisées (le centre d'un point de mesure est dans ce cas le centre de la croix correspondante) ou des cibles circulaires (le centre d'un point de mesure est dans ce cas le centre de chaque disque correspondant).

En illustration, la **figure 3.8** montre un motif qui peut être utilisé comme mire de calibrage. Le centre du repère associé est situé au barycentre des centres des cibles, l'axe $\overrightarrow{x}$ est horizontal vers la droite et l'axe $\overrightarrow{y}$ est vertical vers le bas. Le travail étant effectué avec des repères ortho-normés, l'axe $\overrightarrow{z}$ est dirigé vers la mire. L'altitude des points est supposée égale à zéro puisqu'ils appartiennent tous au même plan.

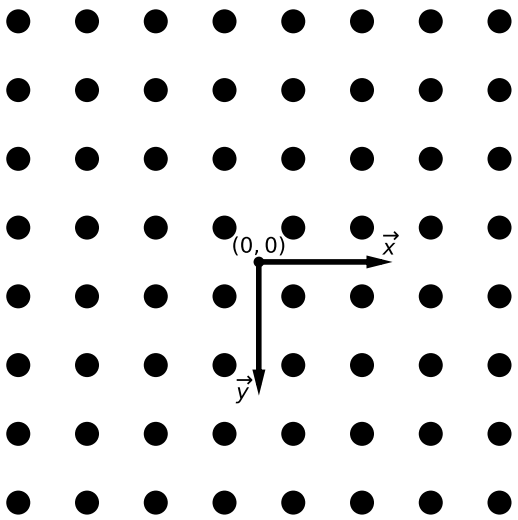


FIGURE 3.8 – Définition du repère lié à une mire : le centre du repère est situé au barycentre des centres des cibles, l'axe $\overrightarrow{x}$ est horizontal vers la droite et l'axe $\overrightarrow{y}$ est vertical vers le bas

3.3 Principe du capteur stéréoscopique

Un capteur stéréoscopique est un système constitué de deux caméras montées rigidement. Par définition, en vision stéréoscopique, les plans images sont parallèles, mais le terme capteur stéréoscopique est aussi utilisé si les caméras ne sont pas parallèles [113, 49]. Ces caméras imitent les yeux : elles permettent d'obtenir deux images depuis deux points de vue différents, puisque les caméras ont des positions et des angles différents par rapport à l'objet d'étude. En effet, les caméras occupent un certain volume, il est impossible de les aligner parfaitement et, fondamentalement, il est nécessaire d'obtenir des points de vue décalés afin d'effectuer la triangulation de la position 3D des points de mesure (*cf. infra*). Les repères des deux caméras sont ainsi nécessairement décalés. La première étape consiste à définir le lien mathématique et géométrique qui les unit, avant d'aborder le calibrage du capteur dans son ensemble et la phase de triangulation.

3.3.1 Lien géométrique entre deux repères de l'espace

Le passage d'un repère 3D à un autre s'effectue via quatre opérations : trois translations, selon chaque axe, une rotation autour d'un axe quelconque, et un facteur d'échelle. Les translations sont

réalisées à l'aide d'un vecteur, soustrait aux positions 3D d'un point. De plus, la rotation peut être séparée en trois sous-rotations, ce qui simplifie la compréhension (*cf. infra*). Finalement, dans le contexte de ce manuscrit, en travaillant avec différents repères métriques, le facteur d'échelle est unitaire, et n'est donc pas exprimé explicitement.

Translation entre deux repères

Supposons tout d'abord que deux repères (1 et 2) sont parfaitement alignés (*i.e.* les axes $\overrightarrow{x}$, $\overrightarrow{y}$ et $\overrightarrow{z}$ des deux repères sont colinéaires) mais décalés dans l'espace. Si l'on a les coordonnées 3D d'un point dans le repère 1, comment calculer sa position 3D dans le repère 2 ? Si l'on prend le repère 1 comme base, on définit $C_2 = (x_{C2}, y_{C2}, z_{C2})^T$ la position du centre du repère 2 exprimé dans le repère 1. La position d'un point $(x_1, y_1, z_1)^T$ dans le repère 1 correspond, dans le repère 2, à :

$$(x_2, y_2, z_2)^T = (x_1 - x_{C2},\ y_1 - y_{C2},\ z_1 - z_{C2})^T \quad . \tag{3.17}$$

On définit alors le vecteur de translation T qui permet de passer du repère 1 au repère 2 selon l'équation :

$$T = -C_2 \quad . \tag{3.18}$$

Rotation entre deux repères

En pratique, les repères des deux caméras ne sont jamais parfaitement alignés (sauf si les images sont obtenues par simulation). Il faut donc définir la rotation existant entre les deux repères, à savoir : la matrice de rotation. Il existe une matrice de rotation R telle que :

$$\mathrm{X}_2 = R\mathrm{X}_1 \tag{3.19}$$

avec X_1 et X_2 les positions 3D d'un point dans les repères 1 et 2 respectivement.

Afin de définir cette matrice R, deux outils mathématiques sont requis. La matrice identité Id :

$$Id = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.20}$$

et la matrice antisymétrique $[.]_x$, produite par un opérateur qui s'applique à un vecteur a, avec $a = (a_1, a_2, a_3)^T$:

$$[a]_x = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \quad . \tag{3.21}$$

La matrice de rotation R, de dimension (3×3), correspondant à une rotation d'angle θ autour d'un axe $\hat{\mathrm{t}}$, est alors définie, selon la formule d'Euler, telle que [64] :

$$R(\theta, \hat{\mathrm{t}}) = Id + \sin(\theta)[\hat{\mathrm{t}}]_x + (1 - \cos(\theta))[\hat{\mathrm{t}}]_x^2 \quad . \tag{3.22}$$

Les matrices de rotation ont pour propriété que leur inverse est égale à leur transposée. On a :

$$R^{-1} = R^T \quad . \tag{3.23}$$

Ainsi, $R^T R$ correspond à une rotation nulle, *i.e.* la matrice identité Id.

Modèle de projection à deux caméras

Soient deux caméras non alignées. Il y a alors deux repères caméras. Une convention usuelle est de définir le repère 1 de la caméra de gauche comme référence. Soient R et T respectivement la matrice de rotation et le vecteur de translation permettant de passer du repère 1 au repère 2. Les coordonnées 3D X_1 d'un point dans le repère 1 s'expriment dans le repère 2 selon l'équation :

$$\mathrm{X}_2 = R(\mathrm{X}_1 + T) \quad . \tag{3.24}$$

Pour la caméra 1, on a l'équation de projection suivante :

$$k_1 * (u_1, v_1, 1)^T = K\mathrm{X}_1 \tag{3.25}$$

avec k_1 la distance entre le point X_1 et le centre du repère caméra 1. Pour la caméra 2, l'équation est similaire, avec X_1 remplacé par X_2. En utilisant l'équation (3.24), on obtient :

$$k_2 * (u_2, v_2, 1)^T = KR(\mathrm{X}_1 + T) \tag{3.26}$$

avec k_2 la distance entre le point X_1 et le centre du repère caméra 2. La matrice de rotation R et le vecteur de translation T liant les repères de deux caméras sont appelés paramètres d'orientation relative. Ils sont obtenus en calibrant le capteur stéréoscopique (*cf.* **section 3.3.2**).

Formalisation et analyse de la matrice de rotation

Si l'on connait la matrice de rotation R permettant de passer du repère caméra 1 au repère caméra 2, il est possible de retrouver les angles donnant une valeur chiffrée compréhensible, permettant de reconstruite cette matrice et de mieux comprendre l'agencement entre les caméras. On découple alors la matrice R en trois sous-rotations. La convention habituelle, empruntée à l'aviation, est : *yaw, pitch and roll* (*i.e.* lacet, tangage et roulis). Le principe utilisé ici est identique, à ceci près que l'axe correspondant à la vue n'est pas l'axe $\overrightarrow{x}$ mais l'axe $\overrightarrow{z}$. On a alors :

$$R = R_{zxy} = R_z R_x R_y \tag{3.27}$$

avec R_x, R_y et R_z trois matrices de rotation telles que $R_y = R(\theta_y, (0,1,0)^T)$,
$R_x = R(\theta_x, R_y(1,0,0)^T)$ et $R_z = R(\theta_z, R_x R_y(0,0,1)^T)$.

On utilise les coordonnées polaires 3D, en définissant l'axe $\overrightarrow{y}$ vers le haut du schéma. En prenant le repère caméra 1 comme référence, l'axe $\overrightarrow{x}$ est fixé de la gauche vers la droite de l'image, l'axe $\overrightarrow{y}$ du haut vers le bas, et l'axe $\overrightarrow{z}$ sort de la caméra. On tourne d'abord autour de l'axe vertical $\overrightarrow{y}$ (regard gauche/droite), puis on tourne autour de l'axe horizontal $\overrightarrow{x}$ obtenu (inclinaison de la vue vers le haut ou le bas), puis autour du nouvel axe $\overrightarrow{z}$. Comme l'illustre la **figure 3.9**, l'axe $\overrightarrow{z}$ (*bleu*) subit une première rotation d'angle θ_y autour de l'axe $\overrightarrow{y}$ (*vert*) puis une seconde rotation d'angle θ_x autour de l'axe $\overrightarrow{x}$ (*rouge*) ayant subi la première rotation (*orange*).

En posant la contrainte $|\theta_x| \leqslant 90°$(cet angle θ_x étant généralement petit en stéréo-vision), d'après la **figure 3.9**, le vecteur $R\overrightarrow{z} = (v_x, v_y, v_z)^T$, correspondant à la dernière colonne de la matrice R, peut être défini comme :

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} \cos(\theta_x)\sin(\theta_y) \\ -\sin(\theta_x) \\ \cos(\theta_x)\cos(\theta_y) \end{pmatrix} \quad . \tag{3.28}$$

Les valeurs de θ_x et θ_y peuvent alors être déterminées. En développant l'expression de $R_z R_x R_y$, on obtient :

$$R_{zxy} = \begin{bmatrix} \cos(\theta_y)\cos(\theta_z) + \sin(\theta_x)\sin(\theta_y)\sin(\theta_z) & \sin(\theta_x)\sin(\theta_y)\cos(\theta_z) - \cos(\theta_y)\sin(\theta_z) & \cos(\theta_x)\sin(\theta_y) \\ \cos(\theta_x)\sin(\theta_z) & \cos(\theta_x)\cos(\theta_z) & -\sin(\theta_x) \\ \sin(\theta_x)\cos(\theta_y)\sin(\theta_z) - \sin(\theta_y)\cos(\theta_z) & \sin(\theta_y)\sin(\theta_z) + \sin(\theta_x)\cos(\theta_y)\cos(\theta_z) & \cos(\theta_x)\cos(\theta_y) \end{bmatrix} \tag{3.29}$$

À partir de l'expression (3.29), on déduit la valeur de θ_z, avec les deux premiers termes de la deuxième ligne, si l'on a préalablement déterminé la valeur de l'angle θ_x à partir de l'expression (3.28).

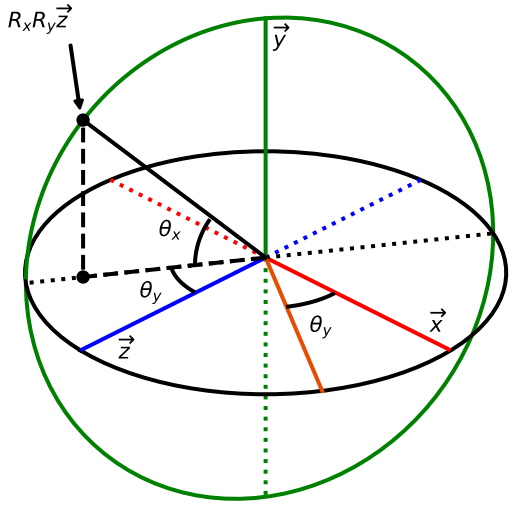


FIGURE 3.9 – Illustration de la rotation autour des axes $\vec{y}$ puis $\vec{x}$; θ_y=-33°et θ_x=-40°

3.3.2 Calibrage d'un capteur stéréoscopique

Le calibrage d'un système stéréoscopique consiste à estimer les paramètres intrinsèques, les paramètres d'orientation relative et les paramètres externes liés aux deux caméras utilisées. Certaines méthodes proposent d'estimer les paramètres d'orientation relative à partir de points appairés dans différentes paires d'images, sans utiliser de mire [64, 149, 38]. Néanmoins, comme mentionné précédemment (*cf.* **section 3.2.3**) les méthodes utilisant une mire sont considérées comme plus précises [151, 49] et sont donc utilisées dans cette thèse.

Rappelons que pour le calibrage d'une caméra réalisé avec une méthode de calibrage fort, l'objectif est d'obtenir les paramètres intrinsèques de la caméra et accessoirement la position et l'orientation de la caméra par rapport à la mire (*cf.* **section 3.2.3**). Pour un capteur stéréoscopique, les paramètres d'orientation relative (*i.e.* la rotation et la translation entre les deux caméras) font partie des informations nécessaires et doivent être déterminés le plus précisément possible, car la précision des reconstructions réalisées avec le capteur en dépend. Les méthodes usuelles d'estimation des incertitudes sont présentées en **section 6.3**. Même s'il est possible d'estimer simultanément les paramètres intrinsèques et extrinsèques, le grand nombre de paramètres à optimiser réduit la précision de l'estimation. Il est alors préférable de calibrer au préalable les caméras, indépendamment, avant d'estimer uniquement les paramètres extrinsèques du capteur, en fixant les valeurs des paramètres intrinsèques des deux caméras [151, 49, 20].

Ainsi, plusieurs paires d'images d'une mire visible dans le champ de vision des deux caméras du système permettent d'obtenir les paramètres d'orientation relative, *i.e.* la matrice de rotation R et le vecteur de translation T, qui minimisent les erreurs de re-projection (*cf.* **section 3.2.3**) dans toutes les images, ainsi que les paramètres d'orientation absolue par rapport au repère de la mire [3]. Un exemple en est donné sur la **figure 3.10**. À l'instar de la méthode de calibrage de caméra, la position 3D des cibles de la mire est optimisée dans le processus. Encore une fois, il est recommandé de prendre plusieurs images de la mire selon différentes orientations et positions afin que les points d'appui soient situés dans un plus grand volume [185, 151].

3. Dans ce manuscrit, le vecteur T est exprimé dans le repère source. L'équation (3.26) est donc utilisée pour projeter un point de l'espace vers le repère image de la caméra 2. Habituellement, le vecteur de translation fourni par les outils de calibrage, dont openCV [20], correspond à une valeur dans le repère de la caméra 2, *i.e.* ici à RT.

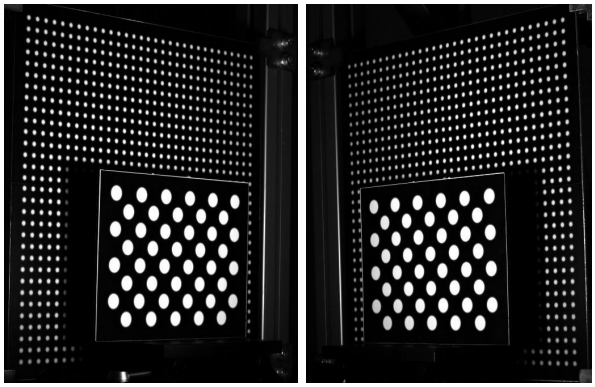


FIGURE 3.10 – Exemple de paire d'images utilisée pour le calibrage d'un système stéréoscopique à angles convergents

Paramètres extrinsèques et repère de la mire

Avec deux caméras différentes, les deux matrices intrinsèques K sont différenciées en ajoutant un indice : K_1 pour la caméra 1 (de référence) et K_2 pour la caméra 2.

Comme vu ci-dessus, le calibrage d'un système stéréoscopique permet d'obtenir la matrice de rotation et la position relative de deux caméras, nommées paramètres d'orientation relative. On peut alors utiliser le repère d'une des caméras comme référence. En utilisant une mire de calibrage, le processus permet aussi d'obtenir la position et l'orientation des deux caméras par rapport à la mire, appelées paramètres externes, ce qui facilite le traitement puisque la mire peut être, assez aisément, alignée au plan moyen d'un objet, à quelques degrés près. On a alors deux matrices de rotation et deux vecteurs de translation par paire d'images utilisée pour le calibrage du système, permettant de passer du repère de la mire (*cf.* **section 3.2.5**), au repère de chaque caméra (*cf.* **section 3.2.2**). Afin d'utiliser les données optimisées lors du calibrage du système stéréoscopique, on utilise un seul des couples rotation-translation afin de calculer une matrice de rotation et un vecteur de translation par caméra, soit R_1 et T_1 pour la caméra de référence et R_2 et T_2 pour la seconde caméra [4]. Les paramètres extrinsèques correspondent alors aux deux matrices de rotation et aux deux vecteurs de translation.

Afin de calculer ces paramètres extrinsèques, on part des équations (3.8) et (3.26) :

$$Z * (u_1, v_1, 1)^T = K_1 \mathrm{X}_1$$

avec $\mathrm{X}_1 = (X, Y, Z)^T$ la position 3D d'un point dans le repère caméra 1, et

$$k_2 * (u_2, v_2, 1)^T = K_2 R(\mathrm{X}_1 + T)$$

avec k_2 la distance entre le point X_1 et le centre du repère caméra 2 (*cf.* **section 3.2.2**). En utilisant les paramètres externes de le caméra 1, R_1 et T_1, obtenus pour l'une des paires d'images du calibrage, permettant de passer du repère de la mire à celui de la caméra 1, on a alors :

$$\mathrm{X}_1 = R_1(\mathrm{X}_{mire} + T_1) \tag{3.30}$$

avec X_{mire} les coordonnées du point X_1 exprimées dans le repère de la mire.
L'équation (3.30) permet ensuite de modifier les équations (3.8) et (3.26), tel que :

$$k_1 * (u_1, v_1, 1)^T = K_1 R_1(\mathrm{X}_{mire} + T_1) \tag{3.31}$$

avec k_1 la distance du point X_{mire} au centre du repère caméra 1 et :

4. Habituellement, le repère de la caméra de gauche est utilisé comme référence ; ce n'est pas toujours le cas dans ce manuscrit.

$$k_2 * (u_2, v_2, 1)^T = K_2 R(R_1(\mathrm{X}_{mire} + T_1) + T) \tag{3.32}$$
$$= K_2 R R_1(\mathrm{X}_{mire} + T_1 + R_1^T T) \tag{3.33}$$
$$= K_2 R_2(\mathrm{X}_{mire} + T_2) \quad . \tag{3.34}$$

Les paramètres extrinsèques sont alors R_1 et T_1 pour la caméra 1 et

$$R_2 = R R_1 \tag{3.35}$$
$$T_2 = T_1 + R_1^T T \tag{3.36}$$

pour la caméra 2.

Utilisation de la forme 3D mesurée - affinage des paramètres externes

Avec un capteur stéréoscopique calibré à l'aide d'une mire de calibrage, les positions 3D d'un nuage de points peuvent être estimées à partir de positions 2D dans les repères images de chaque caméra du capteur. Il est alors possible d'affiner l'estimation de la rotation et de la translation liant les repères des deux caméras en utilisant ces positions 3D estimées [12, 38]. Les positions 3D du nuage de points et les paramètres externes du capteur stéréoscopique sont ainsi estimés successivement, de manière itérative : les positions 3D sont triangulées à partir des positions 2D dans chaque repère image et des données de calibrage, puis ces positions 3D et 2D sont utilisées pour estimer la rotation et la translation entre les deux repères caméras, avant de trianguler de nouvelles positions 3D, etc... Lors du processus, les positions 2D dans les repères images et les paramètres internes de chaque caméra ne sont pas modifiés. Dans ce travail de recherche, cette méthode est utilisée pour les capteurs stéréoscopiques à deux caméras, mais pas avec les montages pseudo-stéréoscopiques.

3.3.3 Triangulation de la position 3D d'un point dans l'espace

Pour trianguler la position 3D métrique d'un point dans l'espace avec un capteur stéréoscopique, il est nécessaire de connaitre :

- les paramètres intrinsèques de chaque caméra ;
- les paramètres extrinsèques du système stéréoscopique ;
- la position 2D des points images dans les repères images des deux caméras.

À partir de ces données, il existe différentes méthodes de triangulation [63, 64, 151]. Dans le contexte de ce travail de recherche, l'objectif est de mesurer des vibrations sub-millimétriques. Il est alors nécessaire d'utiliser une méthode qui fournit des coordonnées 3D correctes pour de tels déplacements. En conséquence, la méthode géométrique proposée par Hartley et Zisserman [64] semble bien adaptée. Elle consiste à projeter dans l'espace le point 2D (u_1, v_1), du repère image de la caméra 1, à une distance z_{cam} du centre du repère caméra 1, en inversant l'équation (3.31). Ce point projeté dans l'espace est ensuite projeté dans le plan image de la caméra 2 pour obtenir les coordonnées $(u_{1\to2}, v_{1\to2})$. Hartley et Zisserman précisent que la distance z_{cam} qui minimise la distance entre le point $(u_{1\to2}, v_{1\to2})$ et le point image (u_2, v_2) dans le repère image 2 correspond à la distance entre le point 3D $(X, Y, Z)^T$ projeté sur l'axe $\overrightarrow{z}$ du repère caméra 1 et le centre de ce repère. Il est à noter cependant que, même si Hartley et Zisserman préconisent de rechercher cette distance z_{cam} spécifique, ils ne donnent ni la méthode pour y aboutir ni la technique de résolution directe.

La **figure 3.11** illustre le principe de cette méthode de triangulation. Deux positions 2D (u_1, v_1) et (u_2, v_2), images du point 3D $(X, Y, Z)^T$ dans les repères images 1 et 2, sont utilisées pour estimer la position 3D de ce point. Le point (u_1, v_1) est ainsi projeté dans l'espace, le long du rayon perspectif, pour différentes valeurs de z_{cam}, pour être ensuite projeté dans le repère image de la caméra 2 (lignes en *rouge*). On notera que la droite obtenue dans le repère image 2 correspond à la droite épipolaire, associée au point (u_1, v_1), sur l'image 2. Idéalement, la droite épipolaire passe par le point (u_2, v_2), mais, en pratique, il y a toujours un décalage, allant de quelques centièmes de pixels à quelques pixels, en fonction de la qualité du calibrage.

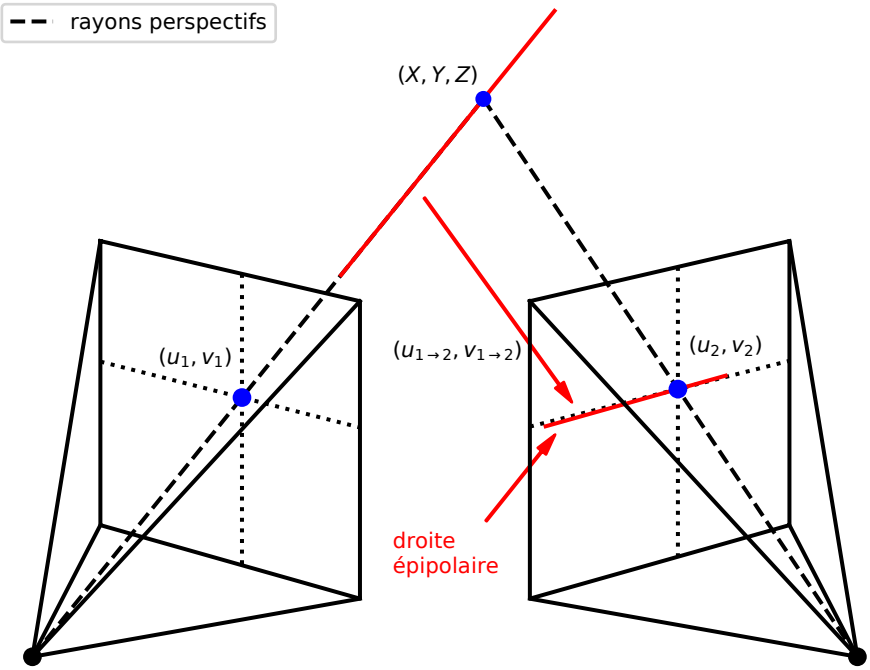


FIGURE 3.11 – Schéma de principe de la méthode de triangulation géométrique avec deux caméras

Les paragraphes suivants expliquent l'implémentation de la méthode.
Pour projeter le point (u_1, v_1) dans l'espace, l'équation (3.31) est inversée :

$$(X_1, Y_1, Z_1)^T = R_1^T K_1^{-1} \mathrm{x}_1 * z_{cam} - T_1 \tag{3.37}$$

avec $\mathrm{x}_1 = (u_1, v_1, 1)^T$.
Ce point de l'espace est ensuite projeté vers le repère image 2 (*cf.* équation (3.34)), ce qui donne :

$$k_{1\to 2} * (u_{1\to 2}, v_{1\to 2}, 1)^T = K_2 R_2 (R_1^T K_1^{-1} \mathrm{x} * z_{cam} + T2 - T1) \tag{3.38}$$

avec $(u_{1\to 2}, v_{1\to 2}, 1)^T$ la position projetée dans le repère image 2. On souhaite mesurer des vibrations de plusieurs centaines (voire milliers) de points de mesure. La position de chacun de ces points sera donc triangulée un grand nombre de fois par acquisition, par exemple 25 000 fois pour 2 secondes d'acquisition à 12 500 images par seconde. Il est alors préférable d'optimiser le processus d'estimation de z_{cam} afin de limiter le temps de calcul.

Comme mentionné ci-dessus, la valeur de z_{cam} recherchée est celle qui minimise la distance entre la droite épipolaire et le point (u_2, v_2). En re-écrivant l'équation (3.38), on obtient :

$$k_{1\to 2} * (u_{1\to 2}, v_{1\to 2}, 1)^T = z_{cam} \begin{pmatrix} krx_1 \\ krx_2 \\ krx_3 \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} \quad . \tag{3.39}$$

La fonction à minimiser est donc :

$$dist^2 = \left(u_2 - \frac{krx_1 z_{cam} + t_1}{krx_3 z_{cam} + t_3} \right)^2 + \left(v_2 - \frac{krx_2 z_{cam} + t_2}{krx_3 z_{cam} + t_3} \right)^2 \quad . \tag{3.40}$$

L'étude de la fonction *dist* est présentée dans le **tableau 3.1** dans le cas d'un système stéréoscopique tel que décrit sur la **figure 3.11** : le point de mesure est ainsi vu par les deux caméras simultanément. On a :

— Pour différentes valeurs de z_{cam}, induisant une valeur positive de $k_{1\to2}$, le point $k_{1\to2} * (u_{1\to2}, v_{1\to2}, 1)^T$ est alors situé devant la caméra 2 et décrit une droite lorsqu'il est projeté dans le repère image 2 :
  — en partant de $k_{1\to2} = 0$, jusqu'à la position optimale $k_{1\to2} = k^{0+}$, la distance *dist* diminue jusqu'à la valeur $dist^+_{min}$;
  — puis, jusqu'à $z_{cam} = \infty$, donc $k_{1\to2} = \infty$, la distance *dist* augmente ;
— Pour z_{cam} induisant $k_{1\to2} = 0$, *dist* est alors égale à $+\infty$;
— Pour des valeurs de z_{cam} qui induisent une valeur négative de $k_{1\to2}$, le point $k_{1\to2} * (u_{1\to2}, v_{1\to2}, 1)^T$ est alors situé derrière la caméra 2 et décrit une seconde demi-droite lorsqu'il est projeté dans le repère image 2.

En fonction de la position et l'orientation relative des caméras ainsi que de la position $(u_1, v_1, 1)^T$, la distance *dist* peut avoir un minimum pour $k_{1\to2} = k^{0-}$. Elle diminue jusqu'à la valeur $dist^-_{min}$, puis augmente à nouveau si k^{0-} n'est pas situé à l'infini.

TABLEAU 3.1 – Tableau de l'étude de la fonction *dist*

$k_{1\to2}$	(-∞)	k^{0-}	0	k^{0+}	$+\infty$
$dist^2$	($+\infty$)	$dist^-_{min}$	$+\infty$	$dist^+_{min}$	$+\infty$

Comme l'illustre le **tableau 3.2**, la dérivée de la fonction *dist* par rapport à z_{cam}, $\frac{\partial dist^2}{\partial z_{cam}}$, présente une discontinuité pour $k_{1\to2} = 0$ et admet potentiellement deux zéros, pour z_{cam} induisant $k_{1\to2} = k^{0-}$ et $k_{1\to2} = k^{0+}$.

TABLEAU 3.2 – Tableau de l'étude de la fonction $\frac{\partial dist^2}{\partial z_{cam}}$

$k_{1\to2}$	(-$\infty \to k^{0-}$)	k^{0-}	$k^{0-} \to 0^-$	0	$0^+ \to k^{0+}$	k^{0+}	$+\infty$
$\frac{\partial dist^2}{\partial z_{cam}}$	$-$	0	$+$	$\emptyset$	$-$	0	$+$

Cette dérivée peut s'écrire sous la forme :

$$\begin{aligned}\frac{\partial dist^2}{\partial z_{cam}} = \frac{2}{(krx_3 z_{cam} + t_3)^3}\Big(&\big(krx_3 t_1 - krx_1 t_3\big)\big((krx_3 z_{cam} + t_3)u_2 - (krx_1 z_{cam} + t_1)\big)\\ &+ \big(krx_3 t_2 - krx_2 t_3\big)\big((krx_3 z_{cam} + t_3)v_2 - (krx_2 z_{cam} + t_2)\big)\Big) \quad .\end{aligned} \tag{3.41}$$

L'astuce consiste alors à multiplier $\frac{\partial dist^2}{\partial z_{cam}}$ par $(krx_3 z_{cam} + t_3)^3$, c'est à dire $k^3_{1\to2}$, ce qui peut alors être écrit sous la forme :

$$\frac{\partial dist^2}{\partial z_{cam}} * k^3_{1\to2} = -2(z_{cam}\beta + \gamma) \tag{3.42}$$

avec

$$\begin{aligned}\beta = {}&(krx_2 krx_3 t_3 - krx_3^2 t_2)v_2 + (krx_1 krx_3 t_3 - krx_3^2 t_1)u_2 + (-krx_1^2 - krx_2^2)t_3\\ &+ krx_2 krx_3 t_2 + krx_1 krx_3 t_1\end{aligned} \tag{3.43}$$

$$\begin{aligned}\gamma = {}&krx_3 t_1^2 + krx_3 t_2^2 + (-krx_2 t_2 - krx_1 t_1)t_3 + (krx_1 t_3^2 - krx_3 t_1 t_3)u_2\\ &+ (krx_2 t_3^2 - krx_3 t_2 t_3)v_2 \quad .\end{aligned} \tag{3.44}$$

La fonction résultante est une droite dont l'étude est présentée dans le **tableau 3.3**. En partant de $k^3_{1\to2} = +\infty$, $\frac{\partial dist^2}{\partial z_{cam}}$ étant positive (*cf.* **tableau 3.2**), $\frac{\partial dist^2}{\partial z_{cam}} * k^3_{1\to2}$ l'est aussi *ipso facto*. De plus, $\frac{\partial dist^2}{\partial z_{cam}} = 0$ pour $k^3_{1\to2} = k^{0+}$, donc $\frac{\partial dist^2}{\partial z_{cam}} * k^3_{1\to2}$ l'est également. La fonction $\frac{\partial dist^2}{\partial z_{cam}} * k^3_{1\to2}$ est négative pour $k^3_{1\to2} < k^{0+}$.
Finalement, la distance z_{cam} minimisant la distance entre (u_2, v_2) et $(u_{1\to2}, v_{1\to2})$ est :

$$z_{cam} = -\gamma/\beta \quad . \tag{3.45}$$

TABLEAU 3.3 – Tableau de l'étude de la fonction $\frac{\partial dist^2}{\partial z_{cam}} * k^3_{1\to2}$

$k_{1\to2}$	(-$\infty \to k^{0+}$)	k^{0+}	$k^{0+} \to +\infty$
$\frac{\partial dist^2}{\partial z_{cam}} * k^3_{1\to2}$	$-$	0	$+$

Ainsi, la position 3D estimée du point $(X, Y, Z)^T$ est obtenue grace à l'équation (3.37) avec la distance z_{cam} calculée avec l'équation (3.45).

3.3.4 Passage vers le repère objet

Si les paramètres extrinsèques R_1, T_1, R_2 et T_2 des deux caméras sont exprimés dans le repère de la mire, positionnée de telle sorte à être alignée au plan moyen de l'objet à quelques degrés près, la forme triangulée sera alors exprimée dans ce repère, ce qui peut être gênant pour juger de la précision voire de la qualité de la mesure. Sur la **figure 3.12**, (à *gauche*), une forme présentant une déformation de 6% est affichée dans un repère décalé de seulement 2° par rapport au plan moyen. Cette rotation est clairement visible, et perturbe l'analyse de la forme. La même forme exprimée dans un repère parallèle au plan moyen (*cf.* **figure 3.12** à *droite*), est plus facilement identifiable.

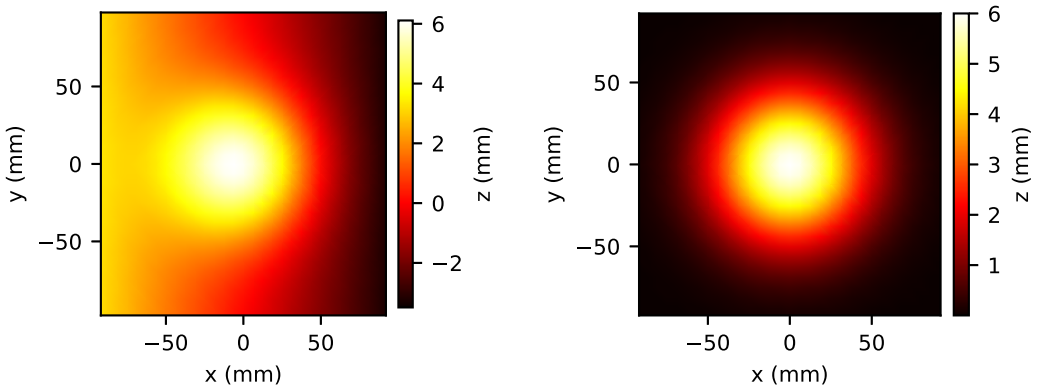


FIGURE 3.12 – Exemple de forme : (*gauche*) affichée dans un repère décalé de 2°par rapport au plan moyen, (*droite*) affichée dans un repère aligné avec le plan moyen

Sachant qu'il est impossible de parfaitement aligner la mire avec le plan moyen de la surface mesurée, comment peut-on alors calculer la rotation nécessaire pour afficher la forme mesurée de manière plus visible dans ce nouveau repère ? Et peut-on modifier les paramètres extrinsèques afin de trianguler des déplacements directement dans ce repère objet ? L'astuce est de calculer le plan moyen de la surface mesurée, afin d'obtenir la normale associée. Cette normale peut ensuite être utilisée pour calculer deux angles de rotation, autour des axes $\overrightarrow{y}$ et $\overrightarrow{x}$ successivement, comme présenté en **section 3.3.1** avec l'expression (3.28). Ces rotations, ainsi que le barycentre de la forme mesurée, sont alors utilisés afin de calculer de nouveaux paramètres extrinsèques permettant de trianguler des déformations dans ce nouveau repère objet, évitant ainsi de transposer les signaux de vibration mesurés d'un repère à un autre. Les paragraphes suivants expliquent la méthode utilisée.

Calcul du plan moyen

Un plan de l'espace est défini tel que :

$$ax + by + cz + d = 0 \quad . \tag{3.46}$$

À partir de n positions 3D $(X, Y, Z)^T$ d'une surface, les coefficients (a, b, c, d) peuvent être obtenus en utilisant la méthode de la pseudo-solution [106]. Ainsi, en fixant $d = 1$, on définit le système d'équations :

$$Au = B \tag{3.47}$$

avec $u = (a, b, c)^T$ le vecteur à déterminer, B une matrice $(n \times 1)$ remplie de 1 correspondant au paramètre d, et

$$A = [\mathbf{x}\,\mathbf{y}\,\mathbf{z}] \tag{3.48}$$

avec $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ trois matrices $(n \times 1)$ correspondant aux n positions 3D. Si la valeur moyenne d'un des trois vecteurs $\mathbf{x}$, $\mathbf{y}$ ou $\mathbf{z}$ n'est pas nulle, alors $(A^T A)^{-1}$, l'inverse de la matrice normale à A, peut être calculée afin de déterminer le vecteur u. On a :

$$A^T Au = A^T B \tag{3.49}$$

$$u = (A^T A)^{-1} A^T B \tag{3.50}$$

$$u = A^{-1} A^{-T} A^T B. \tag{3.51}$$

Il s'agit alors d'une approximation par la méthode des moindres carrés, en prenant un poids identique pour chaque point d'observation.

Calcul de la normale et passage du repère mire vers le repère objet

Le plan moyen $ax + by + cz + d = 0$ peut être calculé à partir de positions 3D d'une forme à l'aide de la pseudo-solution (3.50). Le vecteur normal à ce plan est simplement $\overrightarrow{v_n} = (a, b, c)^T$. Soient deux points $\mathrm{x}_1 = (x_1, y_1, z_1)^T$ et $\mathrm{x}_2 = (x_2, y_2, z_2)^T$ du plan moyen validant les équations suivantes :

$$ax_1 + by_1 + cz_1 + d = 0 \tag{3.52}$$

$$ax_2 + by_2 + cz_2 + d = 0 \quad . \tag{3.53}$$

En soustrayant les deux équations (3.52) et (3.53), on obtient :

$$a(x_2 - x_1) + b(y_2 - y_1) + c(z_2 - z_1) = 0 \quad . \tag{3.54}$$

Ceci correspond au produit vectoriel $\overrightarrow{v_n} \,.\, \overrightarrow{\mathrm{x}_1\mathrm{x}_2}$ et est nécessairement nul, quel que soit le vecteur du plan. Le vecteur $\overrightarrow{v_n}$, exprimé dans le repère de la mire, permet alors de calculer la rotation R_{pm} permettant d'aligner le repère de la mire au plan moyen de la surface mesurée. En effet, les angles θ_x et θ_y sont calculés à partir de l'expression (3.28) puisque l'angle θ_x est facilement réglable à une valeur inférieure à 90° en alignant la mire avec la surface mesurée. On obtient alors, avec l'expression (3.22) :

$$R_{pm} = R_x R_y \tag{3.55}$$

avec $R_y = R(\theta_y, (0, 1, 0)^T)$ et $R_x = R(\theta_x, R_y(1, 0, 0)^T)$. La position 3D x_{pm} dans le plan moyen est calculée avec l'équation suivante :

$$\mathrm{x}_{pm} = R_{pm}(\mathrm{x}_{mire} - \overline{\mathrm{x}_{mire}}) \tag{3.56}$$

avec $\overline{\mathrm{x}_{mire}}$ le barycentre de la surface dans le repère de la mire.

Paramètres extrinsèques correspondant au repère objet

Il s'agit maintenant, à partir de la matrice R_{pm} et du point $\overline{\mathrm{x}_{mire}}$, de modifier les paramètres extrinsèques qui permettent de passer du repère de la mire aux repères des deux caméras, à savoir les matrices R_1, R_2 et les vecteurs T_1 et T_2. Ceci permet de trianguler de nouvelles positions directement dans le repère lié à l'objet mesuré, *i.e.* à son plan moyen. Ainsi, en inversant l'équation (3.56), on obtient :

$$\mathrm{x}_{mire} = R_{pm}^T(\mathrm{x}_{pm} + R_{pm}\overline{\mathrm{x}_{mire}}) \quad . \tag{3.57}$$

En reprenant l'équation (3.32) et en remplaçant R, T, R_1, T_1 et x_{mire} par R_i, T_i, R_{pm}^T, $R_{pm}\overline{\mathrm{x}_{mire}}$ et x_{pm}, on obtient :

$$k_i * (u_i, v_i, 1)^T = K_i R_i (R_{pm}^T (\mathrm{x}_{pm} + R_{pm}\overline{\mathrm{x}_{mire}}) + T_i) \quad . \tag{3.58}$$

Les nouveaux paramètres extrinsèques R_1^{pm}, T_1^{pm}, R_2^{pm} et T_2^{pm} sont alors calculés à partir des équations (3.35) et (3.36) modifiées :

$$R_i^{pm} = R_i R_{pm}^T \tag{3.59}$$

$$T_i^{pm} = R_{pm}(\overline{\mathrm{x}_{mire}} + T_i) \quad . \tag{3.60}$$

3.4 Conclusion

Ce chapitre a présenté le fonctionnement d'un capteur stéréoscopique et des caméras qui le composent, en définissant les différents repères associés (à savoir le repère image, le repère caméra, le repère de la mire et le repère objet). Le lien existant entre le repère image et le repère caméra a été défini avec le modèle de caméra sténopé. Les relations géométriques et mathématiques permettant de passer d'un repère à un autre ont aussi été rappelées, avec les notions de matrice de rotation et de vecteur de translation. Les méthodes utilisées pour déterminer les paramètres intrinsèques et extrinsèques du système stéréoscopique ont également été présentées. Il a finalement été rappelé comment obtenir, à partir des positions 2D de points dans chaque image, des positions 3D dans l'espace par triangulation. La question de l'extraction de ces positions des points 2D dans chaque image se pose maintenant. Notons aussi que, dans la suite du manuscrit, les distorsions optiques des objectifs ne sont plus mentionnées car corrigées préalablement.

Chapitre 4

Traitement d'images

« *Une image vaut mille mots.* »
Confucius

Sommaire

4.1 Introduction **49**
4.2 Filtrage convolutif **50**
4.3 Appariement des points **51**
4.4 Détection de cibles circulaires **52**
4.4.1 Centroïde du niveau de gris ou moyenne pondérée 52
4.4.2 Détecteur de blobs de la bibliothèque OpenCV 53
4.4.3 Coniques et ellipses 54
4.4.4 Estimateur d'ellipses de Fitzgibbon 55
4.4.5 Estimateur d'ellipses duales de Ouellet 57
4.4.6 Détection initiale de cibles 61
4.4.7 Tri et appariement initial 61
4.5 Suivi de cibles **63**
4.6 Motif aléatoire **65**
4.6.1 Outils de corrélation d'images (DIC) 66
4.6.2 Appariement initial - mesure de forme 70
4.7 Suivi de subsets **73**
4.8 Conclusion **74**

4.1 Introduction

Comme vu dans le chapitre précédent, l'homme, tout comme d'autres animaux, est capable avec deux yeux et des processus cérébraux innés et performants d'estimer une forme à partir de deux images. Le capteur stéréoscopique calibré permet d'imiter l'acquisition de ces deux images et d'effectuer la triangulation. L'objectif de ce chapitre est d'expliquer les algorithmes de traitement d'images qui tentent de copier le fonctionnement du cerveau humain et permettent l'appariement des points de mesure dans les images qui serviront à trianguler les positions dans l'espace. Cette approche requiert initialement l'obtention des positions 2D des points dans chaque image.

Pour extraire ces positions dans les images, des outils de traitement d'images sont utilisés, en distinguant trois cas :

1. soit des cibles sont collées sur l'objet à mesurer ;
2. soit un motif aléatoire est appliqué sur l'objet à mesurer ;

3. soit aucun motif n'est utilisé : il s'agit de détecter des contours.

Dans cette thèse, les deux premiers cas sont utilisés afin d'effectuer une mesure de forme initiale. Ceci nécessite l'utilisation d'algorithmes adaptés : des détecteurs de cibles et des outils de corrélation d'images. Dans un deuxième temps, un suivi des points d'intérêt dans les séries temporelles d'images doit être effectué afin de triangulation les signaux vibratoires correspondant aux signaux détectés.

Ce chapitre est organisé de la façon suivante. La **section 4.2** reprend les principes mathématiques et algorithmiques du filtrage convolutif, *i.e.* une des techniques de base du traitement d'images. La **section 4.3** propose un état des lieux des méthodes visant à détecter et appairer les points d'intérêt. Sont ensuite présentés les différents détecteurs de cibles circulaires utilisés ainsi que les algorithmes d'appariement correspondants (**section 4.4**), avant d'aborder le suivi de cibles (**section 4.5**). Finalement, le cas correspondant à l'utilisation d'un motif aléatoire sera détaillé en **section 4.6**, notamment les outils de corrélation d'images choisis, l'algorithme d'appariement initial appliqué avec les techniques de rectification d'images, avant d'aborder le suivi des points d'intérêt définis (**section 4.7**).

4.2 Filtrage convolutif

Comme vu en **section 3.2.1**, l'acquisition d'une image par une caméra ou un appareil photo consiste à enregistrer le nombre de photons convertis en électrons sur chaque pixel de la surface photosensible. Le nombre de photons arrivant au capteur, durant le temps d'exposition, correspond à une loi gaussienne centrée autour d'une valeur moyenne liée à l'intensité lumineuse, ce qui engendre un bruit visible sur les images. En conséquence, différentes méthodes ont été développées afin de réduire ce bruit et d'améliorer la qualité des images enregistrées, ou encore de restaurer des images détériorées. Citons par exemple la technique du filtre convolutif linéaire [104] ou non linéaire, par exemple le filtre médian [71] et le filtre de Lee [87], de la décomposition SVD [105], de le filtrage de Wiener [172] ou encore de la décomposition en ondelettes [97].

Pour ce travail de recherche, le choix s'est porté sur la technique de base, largement utilisée dans de nombreux contextes, à savoir la technique du filtre convolutif. Recommandée comme pré-traitement avant d'effectuer l'appariement de points par corrélation d'images [151] par exemple, elle sert notamment à réduire le bruit des images et à calculer leur gradient, information requise dans les techniques de détection et d'appariement utilisées ultérieurement. La pertinence des autres méthodes, appliquées au contexte de cette recherche, reste à étudier afin d'évaluer leur apport potentiel sur la précision de la mesure finale. Néanmoins, la technique de base a montré qu'elle pouvait largement satisfaire aux besoins de ce travail, les précisions des mesures finales obtenues étant très satisfaisantes.

Le produit de convolution 2D est l'outil de base utilisé pour effectuer le filtrage convolutif. Il est calculé tel qu'une image I, convoluée avec un filtre H, ici à support borné, donne une image I'. Cette opération est notée :

$$I' = I \circledast H \quad . \tag{4.1}$$

Le niveau de gris à chaque pixel de la nouvelle image $I'(u, v)$ est calculé selon l'équation suivante :

$$I'(u, v) = \sum_{i=-R}^{R} \sum_{j=-R}^{R} I(u + i, v + j) * H(R + i, R + j) \tag{4.2}$$

avec $R \in \mathbb{N}$ et $(2R + 1 \times 2R + 1)$ la dimension de H. Notons que l'ordre dans lequel sont effectués différents produits de convolution ne change pas le résultat. Ainsi :

$$(I \circledast H_1) \circledast H_2 = I \circledast (H_1 \circledast H_2) = I \circledast H_2 \circledast H_1 \quad . \tag{4.3}$$

De plus, si H peut être séparé en deux filtres uni-directionnels H_x et H_y, l'utilisation de ces deux filtres est plus rapide que l'application d'un seul filtre 2D. Prenons à titre d'exemple un filtre H_x de

dimension (1×5), $H_x = [1, 0, 1, 2, 1]$, et un filtre H_y de dimension (3×1), $H_y = [1, 2, 3]^T$. On a alors :

$$H = H_x \circledast H_y = \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 2 & 0 & 2 & 4 & 2 \\ 3 & 0 & 3 & 6 & 3 \end{bmatrix} \quad . \tag{4.4}$$

Le filtre H résultant est de dimenson (3×5). Effectuer la convolution 2D de ce filtre avec une image $(n_{cols} \times n_{rows})$ correspond à un nombre d'opérations égal à $(15 + 14) * n_{cols} * n_{rows}$. Si l'on convolue successivement H_x puis H_y (ou inversement) avec la même image, le nombre d'opérations sera alors $((3 + 2) + (5 + 4)) * n_{cols} * n_{rows} = 14 * n_{cols} * n_{rows}$. Le temps de calcul est alors approximativement divisé par deux (rapport de 14/29).

Un des filtres convolutifs utilisés pour réduire le bruit est le filtre gaussien 2D [57], H_G, de dimension $(2R + 1 \times 2R + 1)$. On a :

$$H_G(u, v) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(u - R)^2 + (v - R)^2}{2\sigma^2}\right) \tag{4.5}$$

avec σ l'écart type du filtre. Ce filtre H_G est décomposé en deux filtres 1D, à savoir h_G et h_G^T, respectivement de largeur ou hauteur $(2R + 1)$. On a :

$$h_G(r) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(r - R)^2}{2\sigma^2}\right) \tag{4.6}$$

avec r un indice de pixels. Ces deux filtres sont convolués successivement avec l'image. La **figure 4.1** présente une illustration des filtres gaussiens 1D et 2D. Sur la *gauche*, la courbe noire correspond à la fonction définie par la formule (4.6) et les points *rouges* aux valeurs du vecteur du filtre 1D h_G. Ici, on choisit $\sigma^2 = 2R^2/9$, afin que la valeur des points aux extrémités du filtre soit au moins dix fois inférieure à la valeur centrale, et afin que la valeur du filtre soit négligeable si $r = R + 1$. Sur la **figure 4.1**, pour $R = 3$, la valeur aux extrémités est environ trente fois inférieure à la valeur centrale ; la valeur du filtre pour $r = 4$ est égale à 0,01 et est donc négligeable. Sur la *droite*, les valeurs du filtre 2D H_G, obtenues avec la formule (4.5) ou avec $h_G^T \circledast h_G$, sont affichées comme une image, en niveaux de gris.

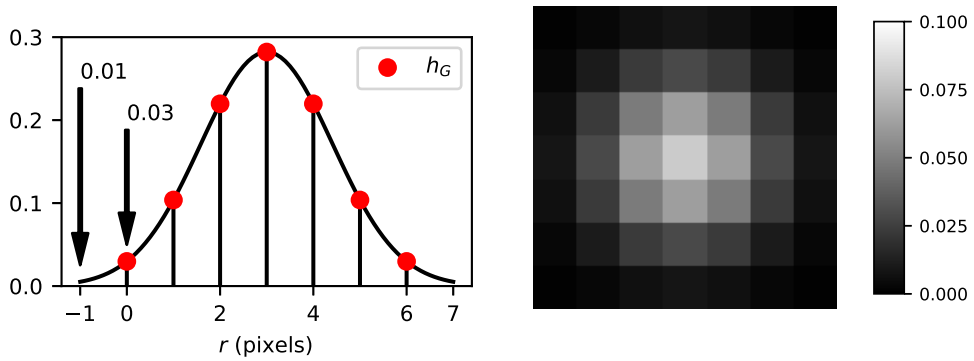


FIGURE 4.1 – Filtre gaussien ; (*gauche*) filtre 1D : valeurs du filtre en *rouge* et fonction continue en *noir* ; (*droite*) filtre 2D : valeurs affichées comme une image en niveaux de gris

4.3 Appariement des points

Afin de trianguler la position 3D d'un point dans l'espace avec un montage stéréoscopique à partir des données de calibrage (*cf.* **sections 3.2.3 et 3.3.2**), deux positions 2D sont nécessaires, soit une

par image. Dans un article [6], Baqersad présente les différentes avancées en matière de mesure de vibrations avec des caméras rapides et des méthodes basées sur la photogrammétrie. En ce qui concerne l'appariement des points, il décrit trois groupes de techniques effectuant :

— soit un appariement géométrique, qui associe des points correspondant à des cibles, croix ou points par exemple ;
— soit un appariement de contours ;
— soit un appariement photométrique, qui fournit un appariement de plusieurs centaines ou milliers de points.

Pour l'appariement géométrique, la détection de cibles circulaires est généralement appliquée pour les mesures d'objets très volumineux comme les pales d'éoliennes [115], d'hélicoptères [95], d'hélices [146], de ponts [111]. La détection d'angles avec les méthodes des coins de Harris [61] et Shi-Tomasi [139], ou la méthode SIFT (*Scale Invariant Feature Transform* en anglais) [92], peuvent aussi être une option. La détection de contours de l'objet [147] est régulièrement sélectionnée pour mesurer des vibrations parallèles au plan image d'une seule caméra. La technique de magnification de mouvement y est parfois associée [32, 178]. Finalement, l'appariement photométrique utilise la corrélation d'images [151] (DIC) afin d'obtenir un appariement d'un grand nombre de points. Pour la mesure de vibrations, les méthodes DIC sont le plus souvent choisies [10, 6].

Les sections suivantes présentent les techniques utilisées au cours de ce travail de recherche : différentes méthodes de détection de cibles, le principe de l'appariement par corrélation d'images, ainsi que des algorithmes de suivi des points de mesure dans des séries temporelles d'images.

4.4 Détection de cibles circulaires

Les cibles circulaires sont très utilisées pour les mires de calibrage et, de manière générale, pour mesurer le déplacement 3D d'objets tournants. De par leur symétrie de révolution, il n'y a, de fait, pas de changement de motif avec la rotation. De plus, différentes méthodes de détection de ces cibles sont simples d'implémentation. Le suivi de cibles peut ainsi être choisi comme une première étape pour élaborer des codes de mesure de vibrations par vision 3D en permettant de suivre l'évolution des positions 2D de cibles dans une série temporelle d'images, à partir de leurs positions 2D dans une image initiale. Les caméras utilisées pour générer des images étant généralement décalées par rapport à la surface de l'objet, l'image d'un disque prend la forme d'une ellipse. Il est alors nécessaire de pouvoir estimer le centre d'une ellipse. Parmi les différents méthodes de détection existantes, sont présentées dans ce manuscrit :

— le centroïde du niveau de gris ou moyenne pondérée ;
— le détecteur de blobs de la bibliothèque OpenCV ;
— l'estimateur d'ellipses de Fitzgibbon ;
— l'estimateur d'ellipses duales de Ouellet.

En effet, la première méthode est très simple et rapide, la deuxième est un standard de la bibliothèque OpenCV. La troisième et la quatrième sont les deux méthodes finalement utilisées dans cette recherche, en fonction des besoins, de par leur précision.

4.4.1 Centroïde du niveau de gris ou moyenne pondérée

Pour une sous-image donnée, ne contenant l'image que d'une seule ellipse, le centroïde en niveaux de gris est défini comme la moyenne des coordonnées de chaque pixel (u, v), pondérées par leurs niveaux de gris respectifs (*cf.* équation (4.2)). Soit $I(u, v)$ le niveau de gris du pixel de position (u, v) dans une image, la position du centroïde de l'ellipse est obtenue avec la formule suivante :

$$\begin{bmatrix} u_c \\ v_c \end{bmatrix} = \begin{bmatrix} \sum_{u,v} I(u,v) * u / \sum_{u,v} I(u,v) \\ \sum_{u,v} I(u,v) * v / \sum_{u,v} I(u,v) \end{bmatrix} . \tag{4.7}$$

En illustration, la **figure 4.2** présente une image de l'estimation du centroïde d'une ellipse qui apparait en noir sur fond blanc. La pondération ne s'effectue alors pas sur l'intensité $I(u, v)$ mais sur l'intensité

inverse selon l'équation :

$$I_{inv}(u, v) = 255 - I(u, v) \tag{4.8}$$

pour une image codée sous 8 bits de 0 à 255.

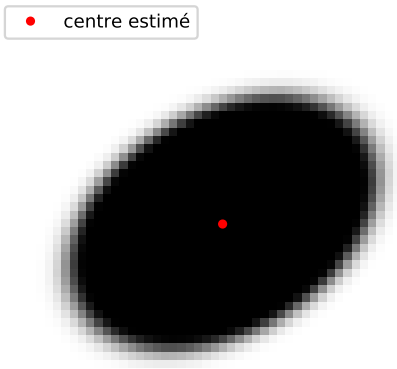


FIGURE 4.2 – Estimation du centroïde d'une ellipse

4.4.2 Détecteur de blobs de la bibliothèque OpenCV

Ce détecteur convertit une image source en images binaires en appliquant un seuillage avec plusieurs valeurs seuil. Ainsi, si $I(u, v)$, le niveau de gris du pixel de position (u, v) dans une image, est supérieur à la valeur seuil, alors $I(u, v)$ est réglé à 1, sinon il est réglé, de façon binaire, à 0. Sont ensuite extraites de chacune de ces images binaires les zones de pixels de valeur 1, grâce au détecteur de contours de Suzuki [154]. Leurs centres sont alors calculés. L'étape suivante consiste à regrouper les centres de plusieurs images binaires à partir de leurs coordonnées. À chaque groupe correspond un "blob", pour lequel le centre et le rayon sont estimés à partir des données du détecteur de contours. La **figure 4.3** présente un exemple réalisé avec la fonction implémentée dans la bibliothèque OpenCV [20]. Le centre et le rayon estimés du blob, ici une ellipse, sont tracés en *rouge*.

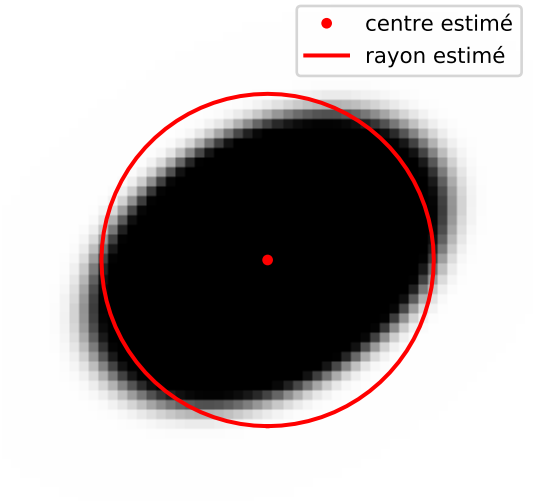


FIGURE 4.3 – Exemple de détection de blob

4.4.3 Coniques et ellipses

Les notions de conique et d'ellipse, ainsi que certains principes géométriques associés, sont présentés dans cette section car ils servent de base géométrique et mathématique pour les estimateurs d'ellipses et d'ellipses duales détaillés ci-après. De plus, ces notions sont utilisées pour le suivi de cibles et de subsets (*cf.* **sections** 4.5 et 4.7).

Conique

Une conique est définie tel qu'un point $\mathrm{x}_i = (x_i, y_i, 1)^T$ soit sur la conique $\mathcal{C}$ si :

$$\mathrm{x}_i^T \mathcal{C} \mathrm{x}_i = 0 \tag{4.9}$$

avec

$$\mathcal{C} = \begin{bmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{bmatrix} \tag{4.10}$$

puisque l'équation (4.9) correspond à la fonction polynomiale suivante :

$$\mathcal{D}_A(\mathbf{a}, \mathbf{x}) = \mathbf{a}^T \mathbf{x} = Ax^2 + Bxy + Cy^2 + Dx + Ey + F \tag{4.11}$$

avec $\mathbf{a} = (A, B, C, D, E, F)^T$ et $\mathbf{x} = (x^2, xy, y^2, x, y, 1)^T$. $\mathcal{D}_A(\mathbf{a}, \mathbf{x})$ est appelée la "distance algébrique" d'un point (x, y) à la conique $\mathcal{C}$. Si la distance algébrique est égale à zéro, alors la condition de l'équation (4.9) est validée et le point quelconque (x, y) se situe sur la conique $\mathcal{C}$. L'ajustement d'une conique peut être approché en minimisant la somme des distances algébriques $\mathbf{D}_A$ à la courbe de la conique pour les n points x_i, avec :

$$\mathbf{D}_A = \sum_{i=1}^{n} \mathcal{D}_A(\mathbf{a}, \mathbf{x_i})^2. \tag{4.12}$$

Afin d'éviter la solution triviale $\mathbf{a} = (0, 0, 0, 0, 0, 0)$, le vecteur $\mathbf{a}$ est contraint à l'aide d'une matrice de contrainte $\mathbf{C}$ de dimension (6×6). Habituellement, deux types de contraintes sont utilisés :

— une contrainte linéaire $\mathbf{Ca} = 1$;
— une contrainte quadratique $\mathbf{a^T Ca} = 1$.

En géométrie projective, le rôle des points et des lignes en coordonnées homogènes peut être interchangé, selon la "relation de dualité". Il existe alors une conique duale $\mathcal{C}^*$ telle que la ligne $l_i = (a_i, b_i, c_i)^T$ soit tangente à la conique $\mathcal{C}$ si l'on a :

$$l_i^T \mathcal{C}^* l_i = 0 \tag{4.13}$$

avec $\mathcal{C}^* = \mathcal{C}^{-1}$. Les paramètres de la conique duale sont notés :

$$\mathcal{C}^* = \begin{bmatrix} A^* & B^*/2 & D^*/2 \\ B^*/2 & C^* & E^*/2 \\ D^*/2 & E^*/2 & F^* \end{bmatrix} \quad . \tag{4.14}$$

Ellipse

Une ellipse est une conique respectant la condition $B^2 - 4AC < 0$. Dans le cadre du suivi de cibles, il est intéressant de savoir la décaler (*cf.* **section 4.5**). Cette section explique ainsi comment décaler une ellipse dans le repère image.

Soit une conique $\mathcal{C}$, correspondant à une ellipse. La distance algébrique entre un point $\mathrm{x} = (x, y, 1)^T$ et la conique $\mathcal{C}$ est donc égale à $\mathrm{x}^T \mathcal{C} \mathrm{x}$ (*cf.* équation (4.9)). Le point x peut être décalé à l'aide d'une homographie affine h [64] selon la formule suivante :

$$\begin{pmatrix} x + \delta_x \\ y + \delta_y \\ 1 \end{pmatrix} = h\mathrm{x} = \begin{bmatrix} 1 & 0 & \delta_x \\ 0 & 1 & \delta_y \\ 0 & 0 & 1 \end{bmatrix} \mathrm{x} \quad . \tag{4.15}$$

En décalant simultanément le point x et la conique $\mathcal{C}$, la distance algébrique reste inchangée. On a :

$$\mathcal{D}_A = \mathrm{x}^T \mathcal{C} \mathrm{x} = (\mathrm{x}^T h^T)(h^{-T} \mathcal{C} h^{-1})(h\mathrm{x}) \tag{4.16}$$

avec

$$h^{-1} = \begin{bmatrix} 1 & 0 & -\delta_x \\ 0 & 1 & -\delta_y \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{4.17}$$

L'ellipse décalée correspond alors au produit matriciel $h^{-T}\mathcal{C}h^{-1}$.

Notons aussi que, à l'instar d'une conique, l'inverse d'une ellipse est une ellipse duale. Une propriété intéressante de l'ellipse duale est que son centre, $\mathrm{x}_c = (x_c, y_c)$, est obtenu directement à partir de ses paramètres [114]. On a :

$$f^*(x_c,\, y_c,\, 1)^T = (D^*/2,\, E^*/2,\, F^*)^T \quad . \tag{4.18}$$

Modification de la taille d'une ellipse Il est aussi intéressant de modifier la taille d'une ellipse dans le cadre du suivi de cibles, en particulier de l'agrandir pour calculer un masque (*cf.* **section 4.5**). Ce paragraphe explique la technique utilisée.

Soit une conique $\mathcal{C}$ correspondant à une ellipse, centrée autour d'un point quelconque x_c (calculé avec l'équation (4.18)). En recentrant l'ellipse autour de l'origine $(0, 0)$, comme présenté ci-dessus, le sixième paramètre F du vecteur **a** associé à la conique est lié à l'opposé du carré de la taille de l'ellipse.

Afin de simplifier la démonstration, l'exemple pris pour illustration sera celui d'un cercle, c'est-à-dire un cas particulier d'ellipse. La distance algébrique (*cf.* équation (4.20)) à un cercle est :

$$\mathcal{D}_A = x^2 + y^2 - 2x_0x - 2y_0y + x_0^2 + y_0^2 - R^2 \tag{4.19}$$

avec (x_0, y_0) la position du centre du cercle et R son rayon. Si le cercle est centré autour de $(0, 0)$, l'équation devient :

$$\mathcal{D}_A = x^2 + y^2 - R^2 \quad . \tag{4.20}$$

Le vecteur **a** associé au cercle est alors $(1, 0, 1, 0, 0, -R^2)^T$. La valeur du sixième paramètre F est ainsi liée dans ce cas au rayon R et donc à la taille du cercle.

En élargissant au cas d'une ellipse quelconque, le paramètre F est lié à la taille de l'ellipse si celle-ci est centrée autour de l'origine du repère image. En conséquence, pour une ellipse de centre $\mathrm{x}_c = (x_c, y_c)$, sa taille peut être modifiée en la décalant de $\delta_x = -x_c$ et $\delta_y = -y_c$ comme vu ci-dessus, puis en multipliant F par un facteur α^2. Sa taille est alors augmentée ou réduite d'un facteur α. L'ellipse peut ensuite être recentrée autour de son centre d'origine, x_c.

En appliquant ce qui vient d'être exposé, si le sixième paramètre F du vecteur **a** est strictement négatif lorsque l'ellipse est centrée autour de $(0, 0)$, il est alors possible de vérifier si un point est dans l'ellipse ou en dehors en calculant la distance algébrique entre le point en question et la conique. En illustration, la **figure 4.4** affiche la carte de la distance algébrique à une ellipse centrée autour du point $(100, 200)$, avec $F = -36$ lorsque l'ellipse est centrée autour de $(0, 0)$. La distance algébrique est égale à zéro au niveau de l'ellipse, est inférieure à zéro à l'intérieur et supérieure à zéro en dehors.

4.4.4 Estimateur d'ellipses de Fitzgibbon

Fitzgibbon propose d'estimer une conique elliptique à partir de positions (x_i, y_i) dans un repère quelconque [44]. Afin de s'assurer que la conique estimée corresponde à une ellipse, le discriminant $B^2 - 4AC$ doit être négatif. Fitzgibbon choisit d'imposer la contrainte quadratique $4AC - B^2 = 1$.

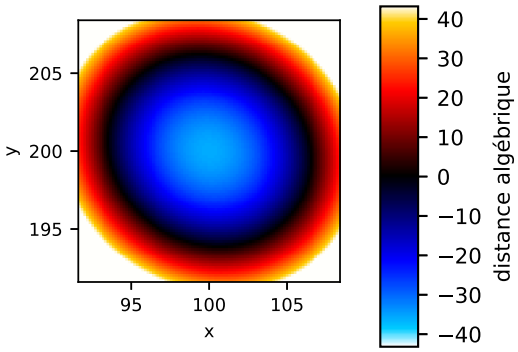


FIGURE 4.4 – Distance algébrique à une conique centrée autour du point $(100, 200)$

Pour ce faire, la matrice de contrainte $\mathbf{C}$ utilisée s'écrit :

$$\mathbf{a}^T \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{a} = 1 \quad . \tag{4.21}$$

Le problème d'ajustement contraint s'écrit alors sous la forme d'un système à deux équations :

$$\begin{aligned} 2\mathbf{D}^T\mathbf{D}\mathbf{a} - 2\lambda\mathbf{C}\mathbf{a} &= 0 \\ \mathbf{a}^T\mathbf{C}\mathbf{a} &= 1 \end{aligned} \tag{4.22}$$

où $\mathbf{D}$ est la "matrice de conception" de dimension $(n \times 6)$ et λ est le multiplicateur de Lagrange. Pour n points de coordonnées rangées dans deux vecteurs $\mathbf{x}$ et $\mathbf{y}$, la matrice conception $\mathbf{D}$ s'écrit :

$$\mathbf{D} = [\mathbf{x}^2, \mathbf{x}\mathbf{y}, \mathbf{y}^2, \mathbf{x}, \mathbf{y}, \mathbf{1}]^T \tag{4.23}$$

avec $\mathbf{1}$ un vecteur de dimension $(n \times 1)$ rempli de 1. Ce système peut être ré-écrit comme :

$$\mathbf{S}\mathbf{a} = \lambda\mathbf{C}\mathbf{a} \tag{4.24}$$

$$\mathbf{a}^T\mathbf{C}\mathbf{a} = 1 \tag{4.25}$$

avec $\mathbf{S}$ la "matrice de diffusion" égale à $\mathbf{D}^T\mathbf{D}$. Ce système est résolu en considérant ses vecteurs propres (*cf.* équation (4.24)). Si un couple valeur propre / vecteur propre $(\lambda_i, \mathbf{u}_i)$ résout l'équation (4.24), alors $(\lambda_i, \mu_i\mathbf{u}_i)$ le résout également. Le coefficient μ_i est défini tel que $\mu_i^2\mathbf{u}_i^T\mathbf{C}\mathbf{u}_i = 1$, soit :

$$\mu_i = \sqrt{\frac{1}{\mathbf{u}_i^T\mathbf{C}\mathbf{u}_i}} = \sqrt{\frac{\lambda_i}{\mathbf{u}_i^T\mathbf{S}\mathbf{u}_i}} \quad . \tag{4.26}$$

Pour plus de détails sur la démonstration, *cf.* [44].

Finalement, parmi les vecteurs $\mathbf{a}_i = \mu_i\mathbf{u}_i$ résolvant le système (4.22), une seule solution appartient à $\mathbb{R}$. En effet, le coefficient μ_i appartient à $\mathbb{R}$ si $\lambda_i > 0$. Or, Fitzgibbon a démontré que la contrainte $4AC - B^2 = 1$ impose une unique valeur propre λ_i strictement positive [44], puisque les valeurs propres de la matrice de contrainte $\mathbf{C}$ sont $(-2, -1, 2, 0, 0, 0)$ (*cf.*(4.21)). En conséquence, le système (4.24) (4.25) admet une seule et unique valeur propre strictement positive.

Récapitulatif de l'algorithme d'estimation d'ellipses de Fitzgibbon :

1. estimer n positions du contour d'une ellipse et les stocker dans deux vecteurs $\mathbf{x}$ et $\mathbf{y}$;

2. construire la matrice $\mathbf{D} = [\mathbf{x}^2, \mathbf{x}.\mathbf{y}, \mathbf{y}^2, \mathbf{x}, \mathbf{y}, \mathbf{1}]^T$ (de dimensions $n \times 6$) ;
3. définir la matrice $\mathbf{S} = \mathbf{D}^T\mathbf{D}$;
4. construire la matrice de contrainte $\mathbf{C}$ selon l'equation (4.21) ;
5. résoudre le système à valeur propre à partir de la matrice $\mathbf{S}^{-1}\mathbf{C}$;
6. sélectionner le vecteur propre correspondant à la valeur propre strictement positive et non infinie.

Pour illustration, la **figure 4.5** (*gauche*) présente un exemple d'estimation d'ellipse à partir de huit points de son contour localisés grâce à une valeur seuil de niveau de gris de l'image. En recherchant des points dans une direction donnée, par exemple horizontale, la position des points peut être estimée avec une précision sub-pixélique en pondérant les coordonnées des pixels par l'inverse de la différence entre le niveau de gris de chaque pixel et la valeur seuil. En effet, si l'on cherche une position (u_{seuil}, v_{seuil}) correspondant à un niveau de gris seuil I_{seuil}, sur une ligne horizontale, si $I(u+1, v) <= I_{seuil}$ [1], alors la position estimée est :

$$\begin{pmatrix} u_{seuil} \\ v_{seuil} \end{pmatrix} = \begin{pmatrix} (c_{u+1}(u+1) + c_u u)/(c_{u+1} + c_u) \\ v \end{pmatrix} \tag{4.27}$$

avec $c_{u+1} = 1/||I_{seuil} - I(u+1, v)||$ et $c_u = 1/||I_{seuil} - I(u, v)||$. La **figure 4.5** (*droite*) montre une estimation d'ellipse avec un nombre de points supérieur. Ceci permet, lorsque les images sont bruitées, d'améliorer la précision de l'estimation.

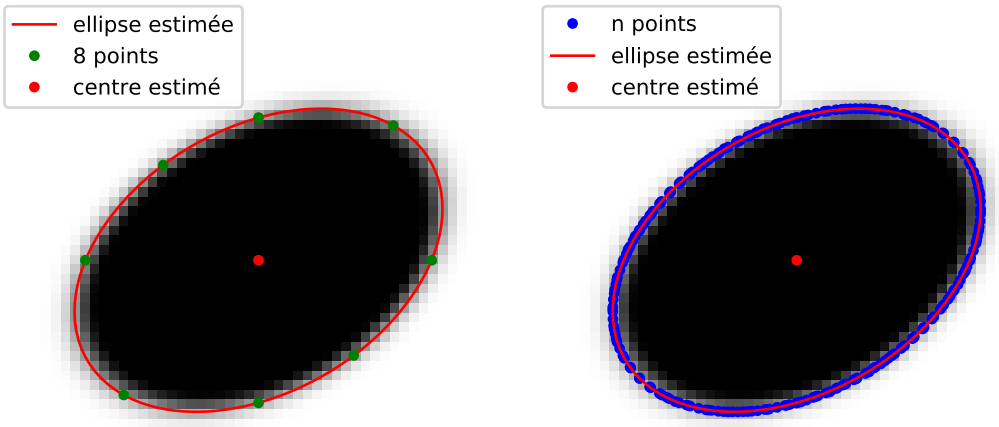


FIGURE 4.5 – Ajustement d'une ellipse

4.4.5 Estimateur d'ellipses duales de Ouellet

Une méthode de localisation de cibles circulaires, utilisant le gradient d'une image, a été développée par Forstner [47], évitant ainsi de chercher les coordonnées du contour d'une conique ou d'une ellipse. Le point correspondant à la convergence des lignes colinéaires (en *vert*) aux gradients de l'image d'une cible est ainsi estimé (*cf.* **figure 4.6** (*gauche*)). Néanmoins, cette méthode ne donne des résultats précis que pour des images de disques, puisque les lignes perpendiculaires à une ellipse ne convergent pas, comme le montre la **figure 4.6** (*droite*).

1. Si l'intérieur de l'ellipse correspondant à la cible dans l'image a un niveau de gris inférieur à celui du fond, les valeurs des niveaux de gris de l'image sont au préalable inversées selon l'équation (4.8).

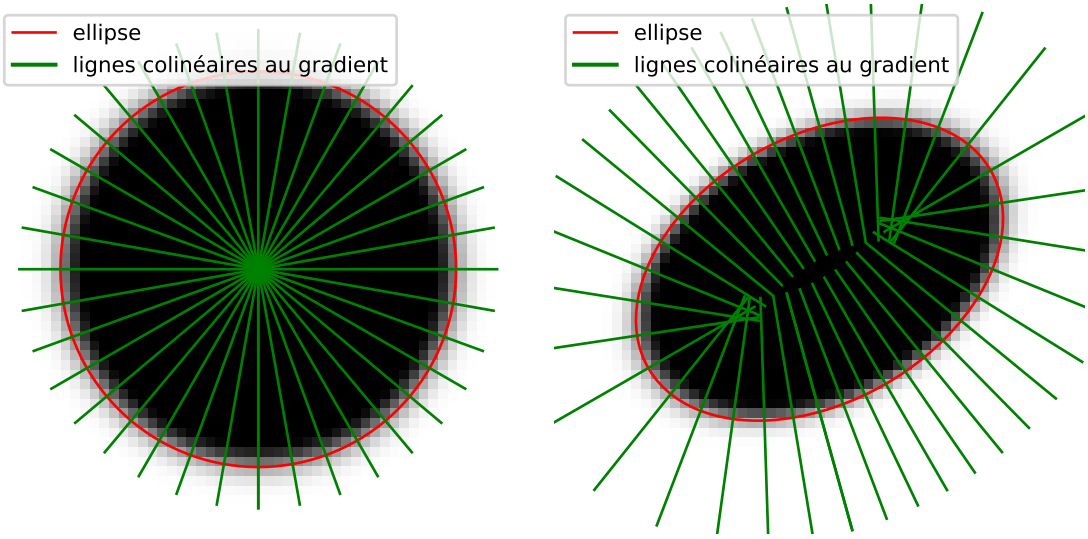


FIGURE 4.6 – (gauche) Convergence des lignes colinéaires aux gradients d'un cercle ; (droite) non-convergence des lignes colinéaires aux gradients d'une ellipse

Dans un article [114], Ouellet apporte une amélioration à l'opérateur de Forstner. Le but est d'estimer une conique duale, correspondant à une ellipse, en s'intéressant aux lignes perpendiculaires au gradient de l'image. En effet, celles-ci sont alors tangentes à l'ellipse imagée, comme représenté en *rouge* sur la **figure 4.7** (notons qu'un cercle est un cas particulier d'ellipse). Comme rappelé en **section 4.4.3**, l'équation d'une conique duale (*cf.* (4.13)) n'est pas liée à la position d'un point, mais à la tangentialité d'une droite l_i à la conique en question. Les lignes perpendiculaires au gradient permettent donc d'estimer une ellipse duale, sans passer par la phase d'estimation de points de contour.

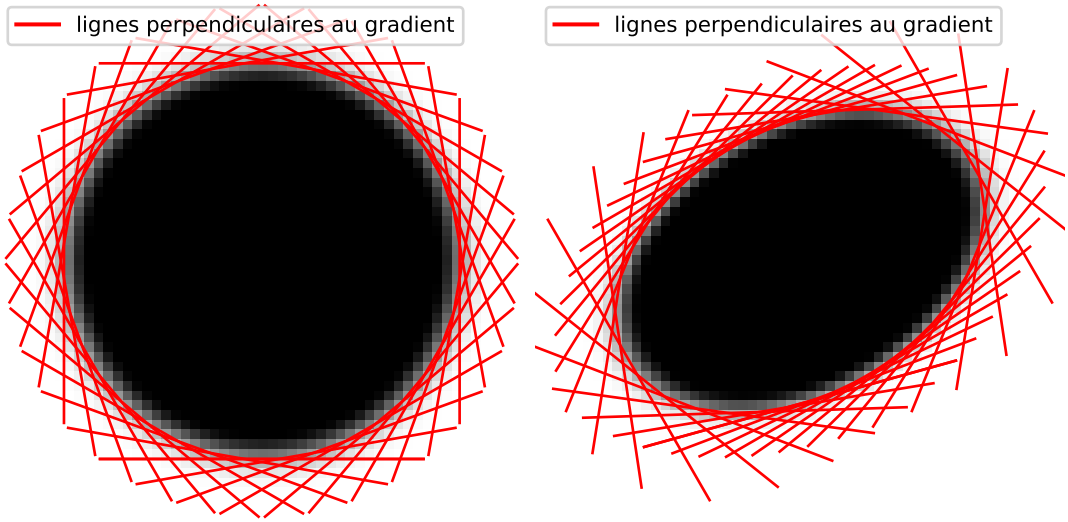


FIGURE 4.7 – Tangentialité des lignes perpendiculaires au gradient de l'image d'une ellipse (un cercle étant un cas particulier d'ellipse)

Conique duale Pour rappel, une ligne l_i est tangente à la conique duale si $l_i^T \mathcal{C}^* l_i = 0$. Étant donné un ensemble de lignes l_i, le vecteur de paramètres $\Theta = (A^*, B^*, C^*, D^*, E^*, F^*)^T$ peut alors être

estimé par la technique des moindres carrés. L'estimation est obtenue en trouvant Θ qui minimise le fonction $\Phi(\Theta)$ suivante :

$$\Phi(\Theta) = \sum_{i \in R} \omega_i (l_i^T \mathcal{C}^*(\Theta) l_i)^2 \tag{4.28}$$

avec R l'ensemble des lignes utilisées pour l'estimation de la conique duale, $\mathcal{C}^*(\Theta)$ la matrice conique duale calculée avec les paramètres du vecteur Θ et ω_i un facteur de pondération. Rappelons que un point $\mathbf{x}$ est sur la droite $l = (a, b, c)^T$ si $l_{\mathbf{x}} = ax + by + c = 0$. Comme l_i est exprimé en coordonnées homogènes, $k * l_i$ et l_i correspondent à la même droite $l_i = (a_i, b_i, c_i)^T$. Le facteur d'échelle peut alors être arbitrairement fixé à $||a_i, b_i|| = 1$. Le système d'équations dérivé de l'équation (4.28) peut s'écrire sous la forme :

$$\left[\sum_{i \in R} \omega^2 K_i K_i^T \right] \Theta = \mathbf{0} \tag{4.29}$$

avec K_i composé des coefficients de la ligne l_i, tel que $K_i = [a_i^2, a_i b_i, b_i^2, a_i c_i, b_i c_i, c_i^2]^T$ et $\mathbf{0}$ un vecteur rempli de zéros. Ce système peut être résolu en utilisant la décomposition SVD, en posant la contrainte $||\Theta|| = 1$. Pour plus de précision sur la démonstration, *cf.* [62].

De la conique duale à l'ellipse duale Comme vu précédemment, la conique duale $\mathcal{C}^*$ correspond à l'inverse mathématique de la matrice conique $\mathcal{C}$. On a : $\mathcal{C}^* = \mathcal{C}^{-1}$. Le paramètre F^* de la conique $\mathcal{C}^*$ peut donc être défini en fonction des paramètres de $\mathcal{C}$, selon l'équation suivante :

$$F^* = \frac{1}{4|\mathcal{C}|}(4AC - B^2) = \frac{1}{|\mathcal{C}|} \begin{vmatrix} A & B/2 \\ B/2 & C \end{vmatrix}. \tag{4.30}$$

Le facteur $1/4|\mathcal{C}|$ peut être ignoré puisque $\mathcal{C}^*$ est déterminée à un facteur près.

On obtient : $F^* = (4AC - B^2)$. Pour rappel, si $4AC - B^2 > 0$, la conique est alors une ellipse. La contrainte pour obtenir une ellipse duale est donc $F^* > 0$. Ouellet choisit d'imposer $F^* = 1$. L'équation régissant le système à 5 inconnues $\Theta' = (A'^*, B'^*, C'^*, D'^*, E'^*)$ est ainsi de la forme :

$$\left[\sum_{i \in R} \omega_i^2 K_i' K_i'^T \right] \Theta' = \sum_{i \in R} -\omega_i^2 K_i' c_i^2 \tag{4.31}$$

avec K_i' correspondant aux cinq premières valeurs de K_i. Ce système est ensuite résolu grace à la décomposition SVD.

Notons que, afin d'optimiser le conditionnement du système d'équations et d'améliorer l'estimation, les lignes sont modifiées tel que le système de coordonnées soit décalé vers le centre approximatif de l'ellipse [62]. Ouellet choisit de définir la ligne l_i à partir du gradient $\nabla I_i = (I_{ui}, I_{vi})$ du pixel $\mathbf{x}_i = (u_i, v_i)$ tel que :

$$l_i = \left[\frac{I_{ui}}{||\nabla I_i||}, \frac{I_{vi}}{||\nabla I_i||}, -\frac{\nabla I_i}{||\nabla I_i||} \mathbf{x}_i \right] \quad . \tag{4.32}$$

Ce gradient ∇I_i est calculé à l'aide du filtre gaussien dérivatif de dimension (5×5), dont seules les plus hautes valeurs sont utilisées. Le coefficient de pondération ω_i^2 est défini tel que :

$$\omega_i^2 = ||\nabla I_i||^2 \quad . \tag{4.33}$$

Là encore, pour plus de détails sur la démonstration, *cf.* [114].

Récapitulatif de l'algorithme d'estimation d'ellipses duales de Ouellet :

1. calculer la dérivée de l'image selon les axes $\overrightarrow{u}$ et $\overrightarrow{v}$ avec le filtre gaussien dérivatif (5×5) ;
2. calculer le gradient et les trois coefficients des lignes l_i avec les plus fortes valeurs du gradient, (Ouellet propose 30% du gradient maximal) ;
3. centrer le repère autour de $(0, 0)$ en modifiant les coefficients des droites ;
4. construire les deux matrices correspondant au système (4.31) ;
5. résoudre le système à l'aide de la décomposition SVD (*cf.* [62]) pour obtenir le centre de l'ellipse duale à partir des coefficients de $\mathcal{C}^*$.

Un exemple d'estimation de conique duale est présenté sur la **figure 4.8**. Les flèches *oranges* correspondent au gradients utilisés pour le calcul de l'ellipse duale et sont bien perpendiculaires à l'ellipse estimée (en *rouge*). L'ellipse résultante se situe là où le gradient est le plus fort, comme l'illustre la **figure 4.9**, qui présente le tracé du gradient en niveaux de gris (les valeurs hautes sont en blanc) et l'ellipse estimée.

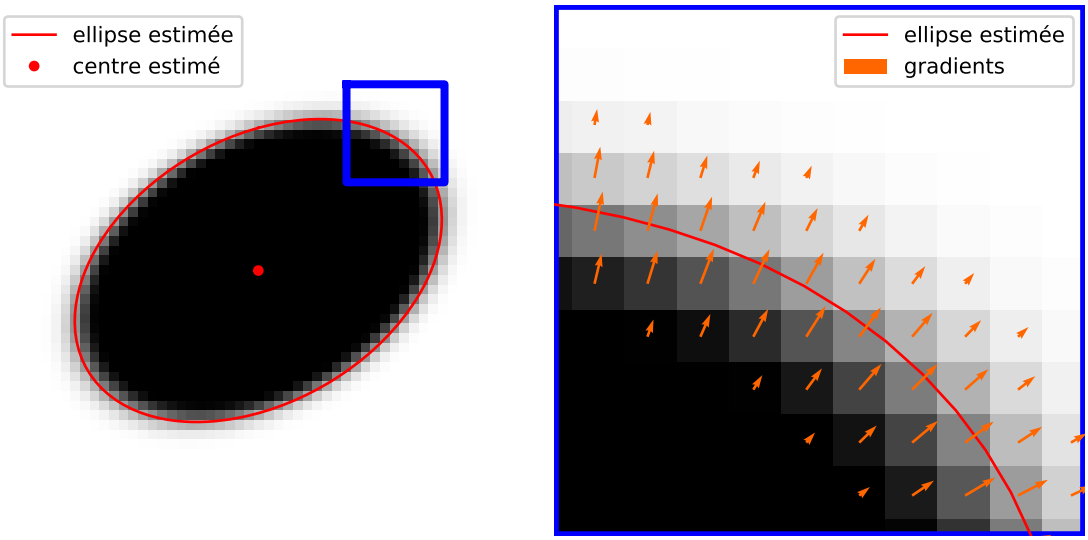


FIGURE 4.8 – (*gauche*) Ellipse estimée, (*droite*) zoom et gradients.

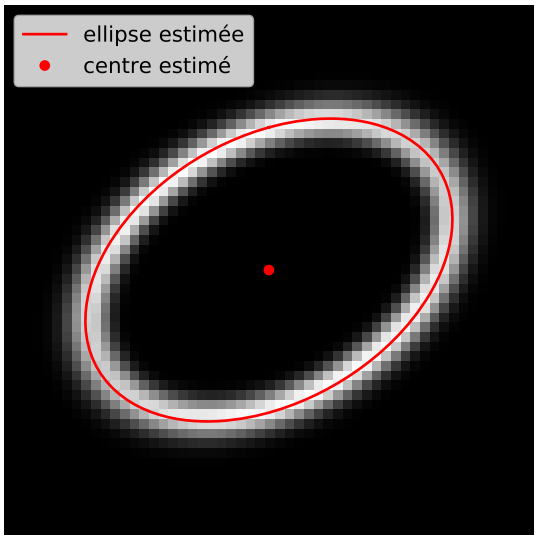


FIGURE 4.9 – Gradients de l'image et ellipse estimée, les pixels blancs correspondent à un gradient élevé

En bref :
- le centroïde est très simple à implémenter ;
- le détecteur de blobs est fourni par la bibliothèque OpenCV ;
- l'estimateur d'ellipses de Fitzgibbon est précis et permet de détecter des ellipses relativement proches, mais nécessite une étape de recherche des points de contours ;
- l'estimateur d'ellipses duales, quant à lui, est aussi très précis et est calculé à partir du gradient des images, mais il peut être perturbé s'il n'y a pas suffisamment d'espace entre le contour de la cible et un élément adjacent ;

— la précision associée à chaque algorithme sera discutée dans le cadre d'une étude de sensibilité en **section 6.3.2**.

4.4.6 Détection initiale de cibles

Que ce soit pour une mire de calibrage ou des points de mesure sur un objet, plusieurs cibles sont utilisées et donc visibles simultanément sur une image. Les détecteurs présentés ci-dessus permettent d'estimer la position d'une cible. L'enjeu est maintenant de détecter plusieurs cibles dans une image. Différents algorithmes peuvent être envisagés. Celui conçu pour ce travail s'appuie sur le détecteur d'ellipses duales. En prenant pour exemple l'image affichée sur la **figure 4.10**, l'algorithme suivant est utilisé pour détecter des ellipses dans l'image :

1. utiliser le produit de convolution pour calculer le gradient de l'image à chaque pixel (*cf.* **figure 4.11**) avec le filtre gaussien dérivatif ;
2. masquer la matrice de gradients obtenue afin de ne garder que les valeurs du gradient supérieures à un seuil, 1/3 de la valeur maximum par exemple (*cf.* **figure 4.12**) ;
3. regrouper les valeurs supérieures à zéro qui sont adjacentes et calculer les ellipses correspondantes (*cf.* **figure 4.13**).

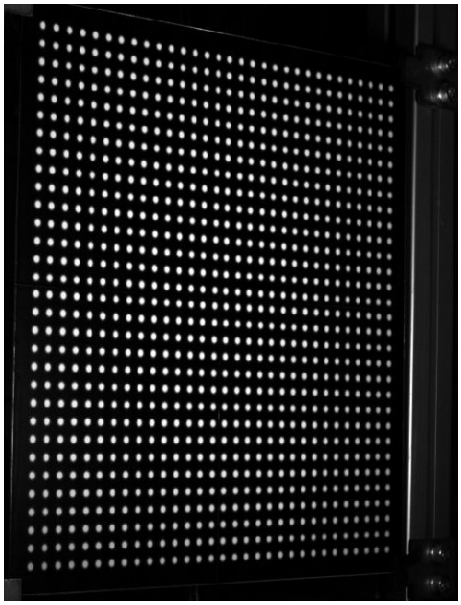


FIGURE 4.10 – Exemple d'image d'une plaque avec un quadrillage de cibles

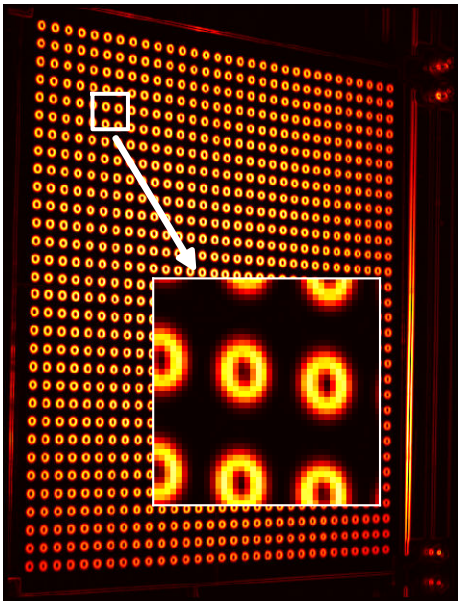


FIGURE 4.11 – Gradient de l'image calculé à chaque pixel

4.4.7 Tri et appariement initial

Les positions estimées des centres des cibles détectées peuvent alors être utilisées pour calibrer une caméra ou un système stéréoscopique à l'aide d'une mire de calibrage, ou encore pour mesurer des positions avec des cibles collées sur un objet, le tout sans avoir besoin de connaître l'espacement *a priori* entre les cibles. Dans tous les cas, les positions obtenues doivent être triées. S'il y a peu de points, le tri peut être effectué manuellement avec une entrée graphique, en utilisant la fonction *ginput* de Matlab ou du module *pyplot* du package *Matplotlib* de Python par exemple. Cependant, si les cibles forment un motif, un tri automatique ou semi-automatique peut être opéré. Ce tri permet d'appairer les positions détectées dans une image avec les coordonnées théoriques d'une mire (*cf.* **section 3.2.5**), ou bien d'appairer des points de mesure entre deux images d'un capteur stéréoscopique. Dans le travail exposé dans ce manuscrit, deux motifs sont utilisés : un quadrillage et un motif en spirale. Les sections suivantes expliquent la démarche à suivre dans ces deux cas.

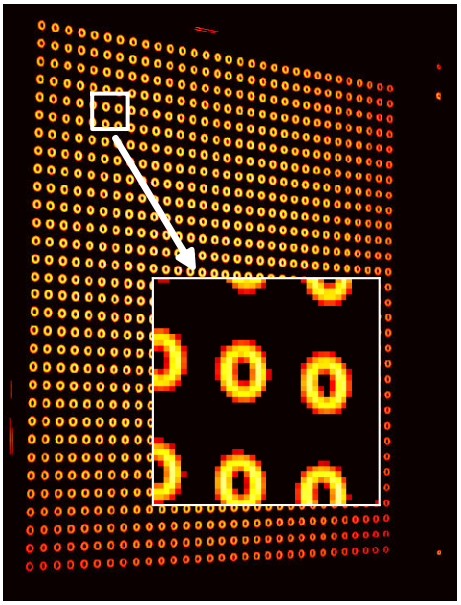


FIGURE 4.12 – Gradient masqué : seules les plus hautes valeurs du gradient sont conservées

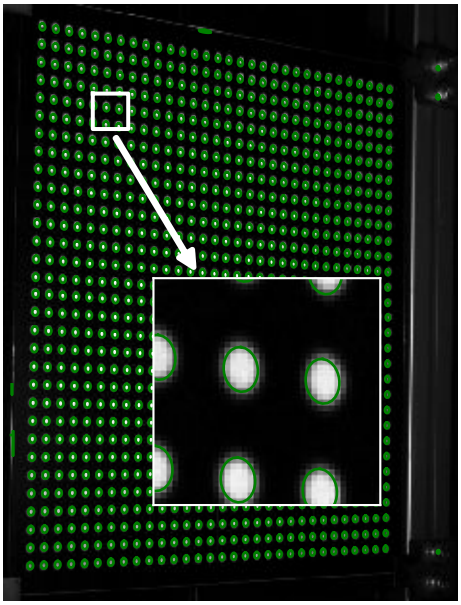


FIGURE 4.13 – Ellipses détectées dans l'image

Tri d'un quadrillage de cibles

Pour un quadrillage de cibles, celles-ci sont triées dans une matrice correspondant au quadrillage. Comme l'illustre la **figure 4.14**, les quatre coins sont d'abord sélectionnés avec une entrée graphique. Ceci permet de définir quatre vecteurs (*flèches blanches*) reliant ces points, et entourant le motif. En connaissant le nombre de points selon les deux axes du quadrillage, les positions détectées peuvent être triées en matrice, en se déplaçant suivant les vecteurs calculés et en cherchant les points de proche en proche. Le tri est alors vérifié en traçant les droites reliant les points de chaque ligne puis de chaque colonne de la matrice. La **figure 4.14** illustre un tri effectué pour les cibles visibles sur les deux images d'un capteur stéréoscopique. Le tri étant effectué de façon identique pour les deux images, chaque point de l'image de gauche peut finalement être facilement appairé au point correspondant de l'image de droite.

Tri d'une spirale de cibles

Les différents essais de calibrage effectués durant ce travail ont abouti à la conception d'une mire en spirale, permettant de simplifier le calibrage d'un système stéréoscopique dans le cas où l'objet mesuré est circulaire (ceci est présenté en **section 5.3**). Un exemple d'une telle mire est affiché sur la **figure 4.15**. Un grand disque encercle des cibles circulaires (ici au nombre de 340), disposées selon une spirale centrée autour d'un point central. La distance entre deux cibles consécutives est constante sur toute la spirale.

Le disque extérieur est détecté (en *bleu* sur la **figure 4.16** *gauche*) et permet de sélectionner les ellipses (en *vert*) se situant à l'intérieur. Les distorsions de perspective ne permettent pas d'effectuer un tri automatique direct. Les positions (u, v) sont alors centrées puis normées, c'est-à-dire divisées par la largeur et la hauteur du motif dans le repère image, comme présenté sur la **figure 4.16** *droite*. Les positions sont ensuite triées de proche en proche. Les deux points les plus proches du point $(0, 0)$ correspondent aux premiers points de la spirale, et ainsi de suite. La vérification du tri est effectuée en traçant la ligne rejoignant différents points (*cf.* **figure 4.17** *gauche*) et en calculant tout d'abord les coordonnées polaires de chaque point dans un repère centré autour du point central. Le tri est alors correct si toutes les différences angulaires entre deux points sont négatives (*cf.* **figure 4.17** *droite*).

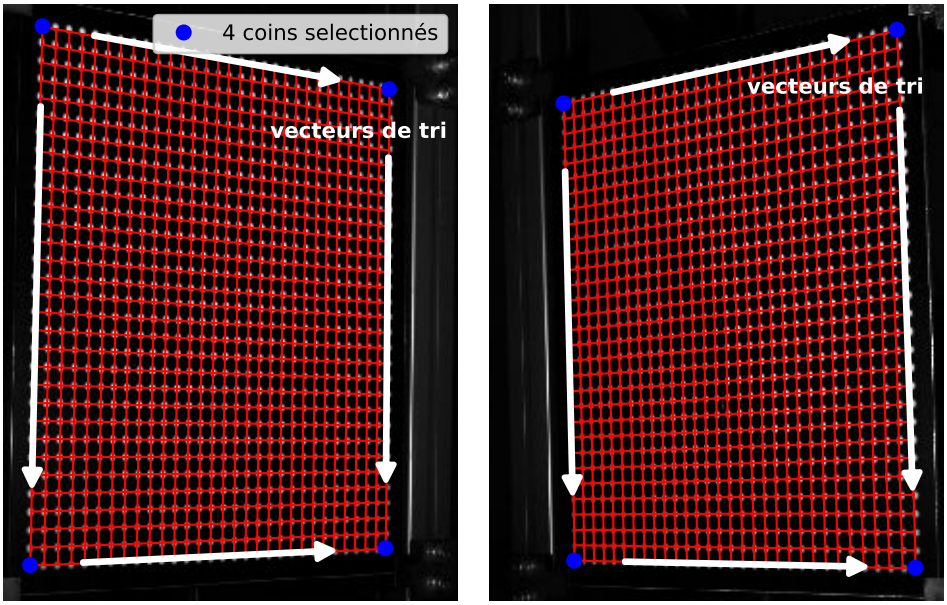


FIGURE 4.14 – Les centres des cibles sont reliés selon un quadrillage afin de vérifier la validité du tri effectué

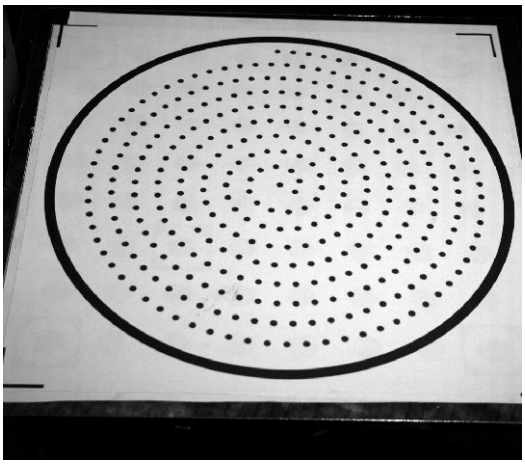


FIGURE 4.15 – Exemple d'image d'une mire en spirale vue de côté

4.5 Suivi de cibles

Les sections précédentes ont présenté l'approche utilisée pour détecter des cibles circulaires. Afin de mesurer des vibrations à l'aide de ces cibles et d'un capteur stéréoscopique, il est nécessaire de connaitre les positions de leurs images dans la ou les séries temporelles d'images correspondant aux deux points de vue. Afin de minimiser le temps de calcul, le protocole de détection initiale et de tri ne doit pas être réitéré à chaque image. La solution choisie est alors de suivre les cibles d'image en image.

Le quadrillage de cibles utilisé est volontairement relativement serré afin de maximiser le nombre de points de mesure. Cependant, comme l'illustre la **figure 4.18**, plusieurs cibles sont visibles, au moins partiellement, dans chaque sous-image centrée autour de l'image d'une cible, ce qui peut perturber l'estimation de la position du centre de la cible centrale.

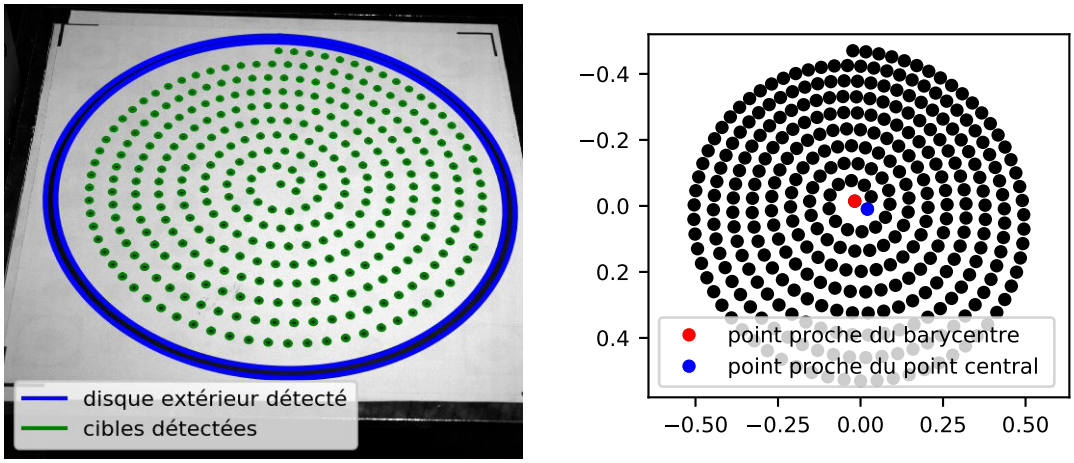


FIGURE 4.16 – Ellipses détectées pour une mire en spirale ; (*gauche*) dans l'image ; (*droite*) positions normalisées des cibles

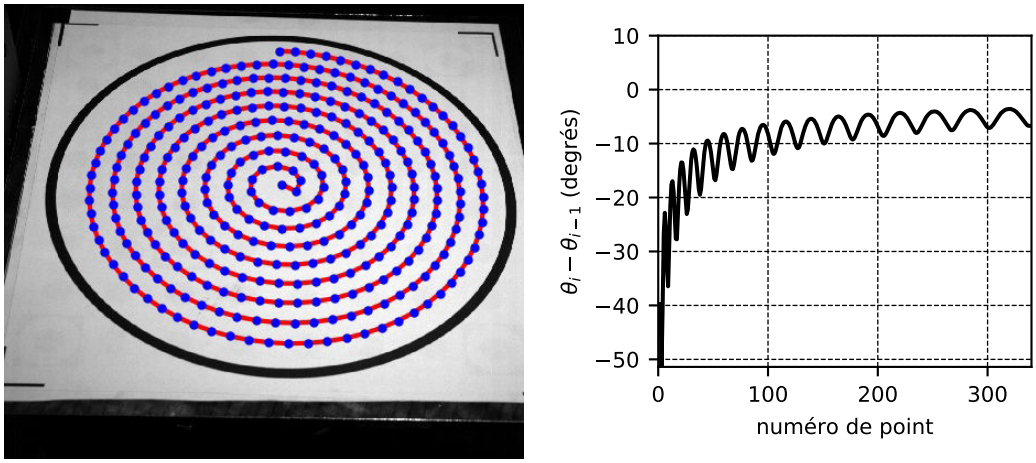


FIGURE 4.17 – (*gauche*) Positions triées des cibles détectées pour une mire en spirale ; (*droite*) différence angulaire entre deux positions une fois le tri effectué

La solution est de masquer la sous-image autour de la cible centrale à l'aide d'une "ellipse-masque", calculée à partir de l'ellipse détectée et estimée dans l'image initiale. Cette ellipse-masque est centrée autour du centre estimé à l'image précédente. La dernière position détectée peut être utilisée comme centre de l'ellipse-masque car le décalage du centre de la cible dans l'image est de l'ordre du centième de pixel. En effet, le déplacement maximum mesuré entre deux positions d'une même cible au sein d'une séquence complète d'images est de l'ordre du pixel. Ceci est dû principalement aux composantes basses fréquences, inférieures à 100 Hz. La cadence d'acquisition des images étant de l'ordre de 10 kHz, le décalage observé entre deux images successives est ainsi de l'ordre de 100/10 000 soit de l'ordre du centième de pixel. Finalement, cette approche permet de suivre chaque cible dans la séquence temporelle en suivant le déplacement de son centre dans les sous-images masquées correspondantes. Le paragraphe suivant explique le processus en détail.

Tout d'abord, dans l'image initiale (*cf.* **figure 4.19** (*gauche*)), sont calculées les ellipses associées aux images des cibles, ici en *vert* (*cf.* **section 4.4.6**). Ces ellipses sont ensuite agrandies d'un facteur α afin d'obtenir des ellipses-masques, ici en *rouge*. Les pixels externes à l'ellipse-masque sont alors masqués dans chaque sous-image en calculant leur distance algébrique à l'ellipse-masque, qui est nécessairement positive (*cf.* **figure 4.19** (*droite*)). Les ellipses-masques sont ensuite décalées d'image en image, afin

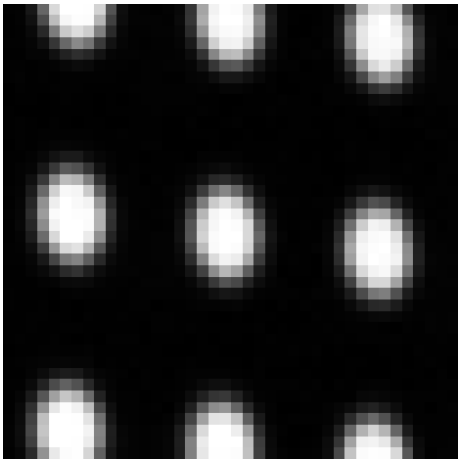


FIGURE 4.18 – Sous-image centrée autour de l'image d'une cible ; plusieurs cibles sont visibles et peuvent perturber l'estimation du centre de la cible centrale

qu'elles soient centrées autour de la position centrale estimée des images des cibles qui leur sont associées. Pour rappel, les notions mathématiques sous-jacentes sont expliquées en **section 4.4.3**. Finalement, la précision de l'estimation des positions des centres des ellipses dépend du motif des

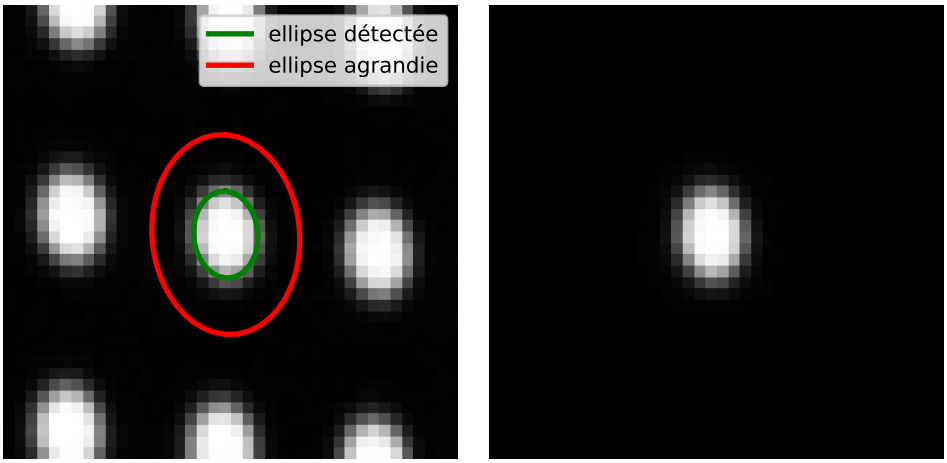


FIGURE 4.19 – Traitement d'images appliqué au suivi d'une cible ; (*gauche*) image initiale avec l'ellipse détectée et agrandie, (*droite*) image masquée à l'aide de l'ellipse agrandie

cibles, de la taille de leurs images et du détecteur utilisé. Cet aspect est présenté en **section 6.3.2**.

4.6 Motif aléatoire

Comme mentionné précédemment, l'utilisation de cibles est très pratique pour le calibrage d'un capteur stéréoscopique ainsi que pour la mesure de vibrations, notamment d'objets rotatifs. Cependant, le nombre de points de mesure est alors limité au nombre de cibles collées sur l'objet. *A contrario*, si une surface présente un motif aléatoire, les outils de corrélation d'images permettent d'effectuer des mesures 3D avec des points de mesure définissables sur toute la surface visible dans les images. La mesure est alors dite "plein champ". Cette section présente les outils utilisés pour appairer des points de mesure, avec des images d'une surface relativement plane présentant un motif aléatoire, acquises avec un capteur stéréoscopique.

4.6.1 Outils de corrélation d'images (DIC)

Coefficient de corrélation

Le premier outil utilisé est le calcul du coefficient de corrélation. Celui-ci permet d'obtenir une valeur chiffrée de la ressemblance de deux fonctions f et g, à une, deux ou n dimensions. Soient deux vecteurs $\mathbf{f}$ et $\mathbf{g}$, avec f_i et g_i les éléments numéro i de ces vecteurs, l'estimation du coefficient de corrélation C_{correl} est obtenue avec la formule discrétisée :

$$C_{correl} = \frac{\sum_{i=1}^{N} \left(f_i - \bar{f}\right)\left(g_i - \bar{g}\right)}{\sqrt{\sum_{i=1}^{N} \left(f_i - \bar{f}\right)^2}\sqrt{\sum_{i=1}^{N} \left(g_i - \bar{g}\right)^2}} \tag{4.34}$$

avec $\bar{f} = 1/N \sum_{i=1}^{N} f_i$, $\bar{g} = 1/N \sum_{i=1}^{N} g_i$ et N le nombre d'éléments du vecteur. Plus les deux fonctions se ressemblent, plus la valeur de C_{correl} est proche de 1 ; moins elles se ressemblent, plus cette valeur est proche de 0. Si les fonctions se ressemblent mais sont de signes opposés, la valeur de C_{correl} est alors négative. Notons aussi que la valeur de C_{correl} est insensible à une différence d'échelle : les coefficients de corrélation estimés entre $\mathbf{f}$ et $\mathbf{g}$ et entre $10 * \mathbf{f}$ et $\mathbf{g}$ sont identiques (*cf.* (4.34)).

En illustration, deux fonctions à une dimension, représentées par des vecteurs, sont affichées sur la **figure 4.20**. Le vecteur $\mathbf{f}$ correspond à un cosinus d'amplitude 1 et le vecteur $\mathbf{g}$ à un cosinus d'amplitude 0,6 auquel est ajouté un bruit d'écart-type égal à 0,09. Le coefficient de corrélation obtenu est égal à 98%.

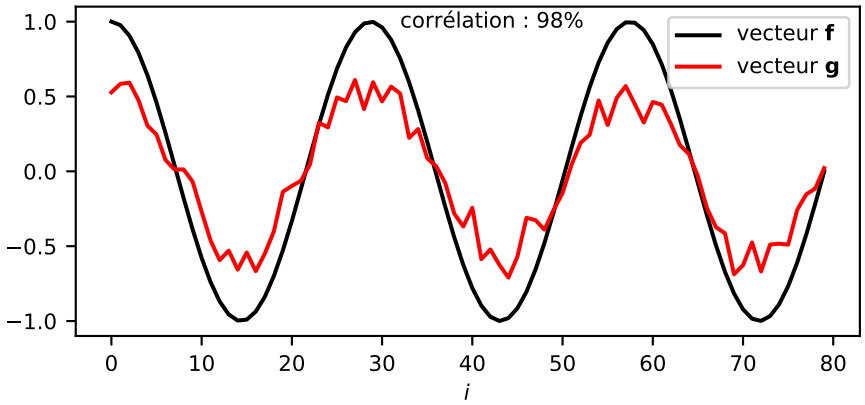


FIGURE 4.20 – Exemple de corrélation 1D de deux vecteurs

Avec deux images représentant chacune une fonction 2D, soient I_1 et I_2, le calcul est habituellement réalisé avec deux matrices :

$$C_{correl} = \frac{\sum_{i=1}^{N_u} \sum_{j=1}^{N_v} \left(I_1(u_i, v_j) - \bar{I_1}\right)\left(I_2(u_i, v_j) - \bar{I_2}\right)}{\sqrt{\sum_{i=1}^{N_u} \sum_{j=1}^{N_v} \left(I_1(u_i, v_j) - \bar{I_1}\right)^2}\sqrt{\sum_{i=1}^{N_u} \sum_{j=1}^{N_v} \left(I_2(u_i, v_j) - \bar{I_2}\right)^2}} \quad . \tag{4.35}$$

Pour plus de souplesse, dans ce manuscrit, les valeurs des niveaux de gris de chaque pixel (u_i, v_i) des images I_1 et I_2 sont stockées dans deux vecteurs de longueur N. Ceci facilite l'implémentation du calcul du coefficient de corrélation pour des sous-images d'une forme quelconque avec la formule (4.34). La **figure 4.21** illustre une corrélation 2D calculée pour les pixels à l'intérieur du cercle *rouge*, dans les deux images, la seconde image étant égale à la première image à laquelle a été ajouté un bruit d'écart-type égal à 0,09. Le coefficient de corrélation estimé est, ici, égal à 97%.

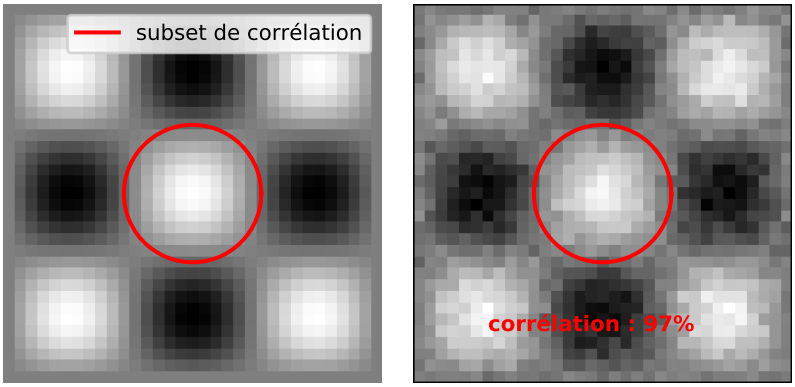


FIGURE 4.21 – Exemple de corrélation 2D pour des pixels inscrits dans un cercle

Subsets et recherche de motif

Le coefficient de corrélation permet de calculer la ressemblance entre deux sous-images ou deux zones de pixels appelées aussi "fenêtres" ou "*subsets*" en anglais. Pour simplifier l'exposé, c'est ce dernier terme qui a été retenu pour ce manuscrit. La plus haute valeur obtenue pour ce calcul de corrélation, entre un motif recherché et différents subsets d'une image cible, donne la position du subset présentant la plus grande ressemblance avec le subset de référence. Pour illustration, le motif recherché est celui qui correspond aux niveaux de gris des pixels dans le cercle *rouge* de l'image de *gauche* de la **figure 4.21**. Sur l'image cible, affichée sur la **figure 4.22** à *gauche*, le cercle *rouge* est déplacé selon les flèches bleues afin de calculer différentes valeurs de corrélation, stockées dans une matrice de coefficients (**figure 4.22** *droite*). On parle alors de corrélation 2D. La position du centre du motif recherché correspond à la position de la valeur maximale dans la matrice obtenue, dans ce cas $\approx 0{,}97$ au centre de la matrice.

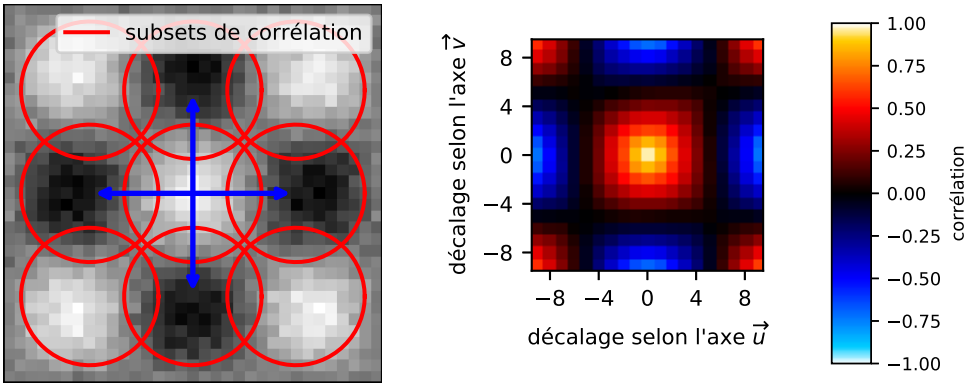


FIGURE 4.22 – (*gauche*) Image cible dans laquelle le motif est recherché ; le cercle de corrélation est déplacé dans l'image pour calculer une matrice de coefficients de corrélation (*droite*)

Corrélation sub-pixélique

La corrélation 2D permet ainsi d'estimer la position d'un subset mais les valeurs (u, v) obtenues pour son centre appartiennent à $\mathbb{N}$, ce qui limite la précision de la reconstruction par triangulation, puisqu'aucune valeur sub-pixélique ne peut alors être obtenue. Toutefois, différentes méthodes existent pour calculer un décalage sub-pixélique entre deux subsets. On peut citer, par exemple, la méthode de recherche *coarse-fine* [153, 183], l'interpolation ou l'ajustement de la courbe des coefficients de corrélation [72, 168], la méthode itérative Newton-Raphson [23, 134, 93], les méthodes basées sur le

gradient [26, 187, 184], la transformée de Fourier 2D [31, 112], les algorithmes génétiques [121] ou encore l'utilisation de réseaux de neurones [122, 123].

Pan *et al.* ont réalisé une comparaison de trois méthodes largement utilisées [17], à savoir l'ajustement de la courbe des coefficients de corrélation, la méthode Newton-Raphson et la méthode basée sur le gradient développée par Davis et Freeman. Les critères d'évaluation appliqués aux données obtenues par traitement d'images simulées sont la précision et le temps de calcul relatif. Les images d'un mouchetis sont simulées, sans bruit ajouté, avec des déplacements imposés dans le repère image variant de 0 à 1 pixel, et servent de données d'entrée aux calculs d'erreur. Les résultats avec des subsets de taille (31 × 31) sont reportés dans les **tableaux** 4.1 et 4.2. Premièrement, la méthode Newton-Raphson s'avère plus précise, mais nécessite un temps de calcul environ vingt-huit fois supérieur aux deux autres méthodes. Deuxièmement, pour un même temps de calcul, la méthode basée sur le gradient est plus précise que la méthode d'ajustement de la courbe des coefficients. Afin de limiter le temps de calcul, pour la mesure de vibrations, la méthode basée sur le gradient de Davis et Freeman a donc été utilisée pour les travaux présentés dans ce manuscrit. Notons que ces auteurs indiquent que la précision attendue est inférieure à 1/50 pixel si le bruit des images est 300 fois inférieur au signal de l'image.

Tableau 4.1 – Écarts-types obtenus par Pan *et al.* [17] à partir d'images simulées avec des subsets de taille (31 × 31)

méthode	Newton-Raphson	gradient	ajustement de courbe
erreur systématique (pixel)	≈ 0,0005	≈ 0,0008	≈ 0,002
écart-type moyen (pixel)	≈ 0,001	≈ 0,0015	≈ 0,014

Tableau 4.2 – Temps de calculs relatifs obtenus par Pan *et al.* [17] pour des subsets de taille (31 × 31)

méthode	Newton-Raphson	gradient	ajustement de courbe
temps de calcul relatif	28,42	1	1,03

Cette section s'appuie sur l'article de Pan *et al.* [17] et présente de façon plus détaillée la méthode basée sur le gradient de Davis et Freeman[2]. Pour de petites déformations, lorsque le subset est suffisamment petit, le subset dans l'image déformée $g(u', v')$ peut être considéré comme non-déformé par rapport au subset dans l'image de référence $f(u, v)$ [26]. On a donc :

$$f(u, v) = g(u', v') \tag{4.36}$$

avec $u' = u+\Delta u+\delta u$ et $v' = v+\Delta v+\delta v$, Δu et Δv correspondant aux valeurs entières du déplacement et δu et δv aux valeurs sub-pixéliques. En négligeant les termes d'ordre élevé, le développement limité de $g(u', v')$ au premier ordre, selon la formule de Taylor, est :

$$\begin{aligned} g(u + \Delta u + \delta u, v + \Delta v + \delta v) = g(u + \Delta u, v + \Delta v) + \delta u * g_u(u + \Delta u, v + \Delta v) \\ + \delta v * g_v(u + \Delta u, v + \Delta v) \end{aligned} \tag{4.37}$$

avec g_u et g_v les dérivées du premier ordre de l'intensité (c'est-à-dire le gradient) de l'image déformée selon les axes $\overrightarrow{u}$ et $\overrightarrow{v}$ respectivement. Ces dérivées sont calculées avec le filtre convolutif 1D $[1/12, -8/12, 0, 8/12, -1/12]$ (*cf.* **section 4.2**). La solution de l'équation (4.37) est obtenue avec la méthode des moindres carrés. On a :

$$\begin{bmatrix} \delta u \\ \delta v \end{bmatrix} = \begin{bmatrix} \sum_i g_{ui}^2 & \sum_i g_{ui} g_{vi} \\ \sum_i g_{ui} g_{vi} & \sum_i g_{vi}^2 \end{bmatrix}^{-1} \times \begin{bmatrix} \sum_i (f_i - g_i) * g_{ui} \\ \sum_i (f_i - g_i) * g_{vi} \end{bmatrix} \tag{4.38}$$

avec f_i, g_i les valeurs des niveaux de gris des images f et g, g_{ui} et g_{vi} les valeurs des gradients g_u et g_v aux pixels (u_i, v_i) choisis.

2. Notons que cette méthode était déjà utilisée par Lucas et Kanade en 1981 [94].

Estimation de la précision

Afin d'estimer la précision de l'estimation des valeurs sub-pixéliques du déplacement, δ_u et δ_v, une simulation est réalisée avec le logiciel Blender [18]. Un motif aléatoire est appliqué sur une plaque en utilisant une texture de Musgrave [107]. Cent images sont simulées en décalant la plaque devant la caméra. Un décalage est alors généré, égal à un centième de pixel entre deux images selon l'axe $\overrightarrow{u}$ et nul selon l'axe $\overrightarrow{v}$. De plus, un bruit photonique (*cf.* **section 3.2.1**) est ajouté, en prenant en compte les caractéristiques de la caméra Photron SA-X2 (*cf* **section 5.2**). La **figure 4.23** présente l'image initiale ainsi que les positions des centres des subsets (en *rouge*) utilisées pour réaliser le suivi. Les subsets correspondent aux pixels situés dans un cercle de corrélation (en *vert*) de rayon égal à neuf pixels, centré autour de chaque point de mesure.

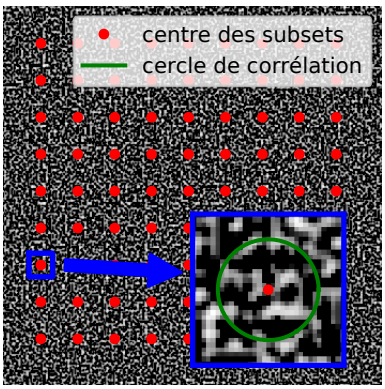


FIGURE 4.23 – Image initiale et points de mesure pour le calcul d'erreur lié à la corrélation sub-pixélique

La **figure 4.24** présente le déplacement mesuré en *noir*, pour le subset affiché sur la **figure 4.23**, selon les axes $\overrightarrow{u}$ (à *gauche*) et $\overrightarrow{v}$ (à *droite*). L'erreur pour le signal δ_u estimé est affichée en *rouge* sur la figure de *gauche*. Sur la figure de *droite*, le déplacement imposé étant nul selon l'axe $\overrightarrow{v}$, le décalage mesuré affiché en *noir* correspond directement à l'erreur de l'estimation sur δ_v. On constate que l'erreur absolue est inférieure à trois centièmes de pixels dans les deux cas.

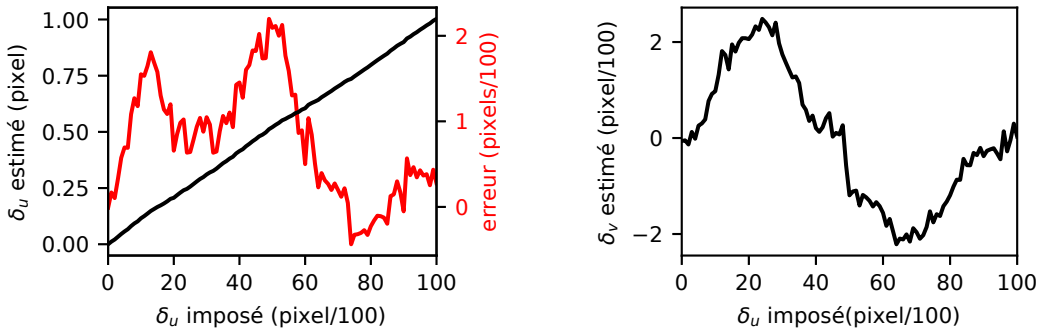


FIGURE 4.24 – Décalage mesuré pour le point de mesure affiché sur la **figure 4.23** ; (*gauche*) pour l'axe $\overrightarrow{u}$ et (*droite*) pour l'axe $\overrightarrow{v}$

Les écarts-types de l'erreur observée pour tous les points de mesure pour tous les subsets sont calculés afin d'obtenir une valeur globale. Les résultats sont affichés dans le **tableau 4.3**. Ils montrent un écart-type inférieur à deux centièmes de pixels et une erreur maximale inférieure au dixième de pixels. Notons que ces valeurs sont obtenues à partir de simulations et ne donnent qu'un ordre de grandeur de l'erreur de mesure.

Tableau 4.3 – Résultats des calculs d'erreur pour la corrélation sub-pixélique

	σ (pixels/100)	erreur maximale (pixels/100)
axe $\overrightarrow{u}$	1,9	8,4
axe $\overrightarrow{v}$	1,6	4,9

4.6.2 Appariement initial - mesure de forme

Recherche de subsets à l'aide de la droite épipolaire

Avant d'effectuer une mesure de vibrations, il est nécessaire de réaliser une mesure de forme initiale. Cette section présente la méthode utilisée pour mesurer la forme d'une surface présentant un motif aléatoire avec un capteur stéréoscopique. Comme précisé ci-dessus, pour chaque point de mesure, une position 2D par image du capteur est requise afin de trianguler la position 3D correspondante. Les outils de corrélation d'images (*cf.* **section 4.6.1**) sont utilisés pour appairer un subset de l'image cible à chaque subset défini dans l'image de référence, en recherchant des "points homologues" à l'aide de la géométrie épipolaire (*cf.* **section 3.3**). Avec un capteur stéréoscopique calibré, chaque point homologue à un point défini dans l'image de référence est proche de la droite épipolaire (*cf.* **section 3.3.3**). Afin d'éviter d'effectuer une corrélation 2D sur toute l'image cible, la recherche est effectuée aux abords de cette droite. La position retenue correspond à la plus haute valeur du coefficient de corrélation. Pour illustration, la **figure 4.25** présente deux subsets appairés pour un objet plan. Le subset de référence est encadré en *bleu*, la droite épipolaire est tracée en *vert*, et le subset appairé est encadré en *rouge*. Notons que le coefficient de corrélation obtenu ici est de 57%, alors que la valeur minimale préconisée est de 70% pour initialiser les méthodes itératives du type Newton-Raphson par exemple [151]. Ceci est dû aux distorsions de perspective, qui sont très différentes entre les deux points de vue puisque l'angle entre les caméras est ici de 80°. Afin d'améliorer l'appariement initial, les images sont alors "rectifiées".

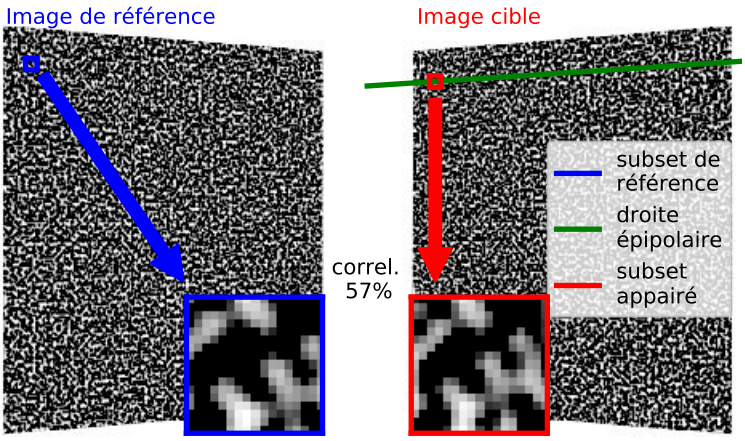


Figure 4.25 – Exemple d'appariement d'un subset à l'aide de la droite épipolaire

Rectification d'images

Cette section présente ainsi une méthode de rectification basée sur la géométrie épipolaire [43] (et non sur la rectification épipolaire [103]). La **figure 4.26** en illustre le principe. Les deux caméras observent une surface (en *vert*), à partir de deux points de vue différents correspondant à des orientations différentes. Les plans images (en *bleu*) ne sont donc pas parallèles, ce qui amène des différences de perspective. Deux matrices de pixels (en *rouge*), dont on choisit les dimensions, sont virtuellement positionnées parallèlement dans l'espace. La position 3D de ces pixels est projetée sur les images initiales afin de calculer une interpolation de niveaux de gris et obtenir des images rectifiées. Ces images

sont filtrées afin de réduire le bruit.

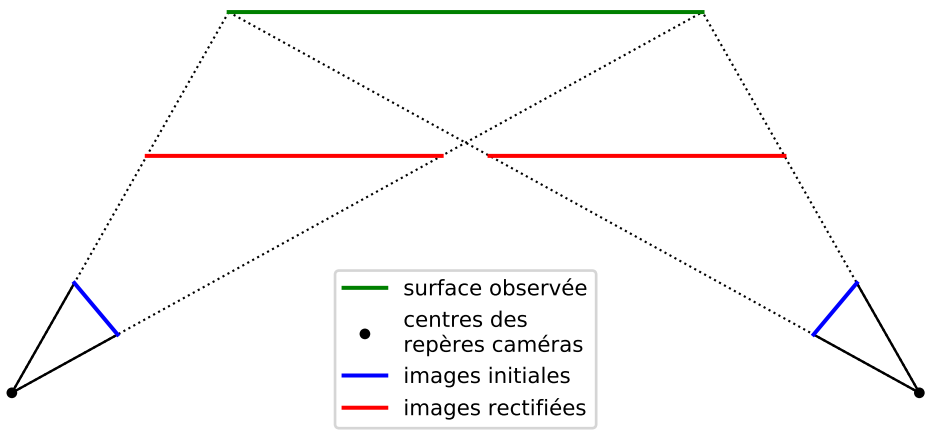


FIGURE 4.26 – Principe de la rectification épipolaire

Notons qu'il est possible d'orienter les caméras afin d'aligner les axes $\overrightarrow{x}$ des repères caméras, pour que les droites épipolaires soient horizontales ; on parle alors d'images normales [156]. Cette option n'a pas été choisie pour ce travail : les repères images sont alignés avec le plan moyen de la mire afin de simplifier le positionnement des points de mesure dans l'image de référence. Les images rectifiées obtenues à partir des images de la **figure 4.25** sont affichées sur la **figure 4.27**. Le coefficient de corrélation entre les subsets appairés, qui était de 57% (*cf.* **figure 4.25**), atteint désormais la valeur de 98%. Dans le cas d'un objet relativement plan, avec des différences importantes de perspective entre les deux points de vue, l'appariement est ainsi plus précis avec les images rectifiées qu'avec les images initiales. Toutefois, les appariement réalisés avec les méthodes de rectifications telles que celle présentée sont généralement utlisées comme initialisation pour des méthodes plus précises, comme la stéréo-corrélation fine améliorée [49].

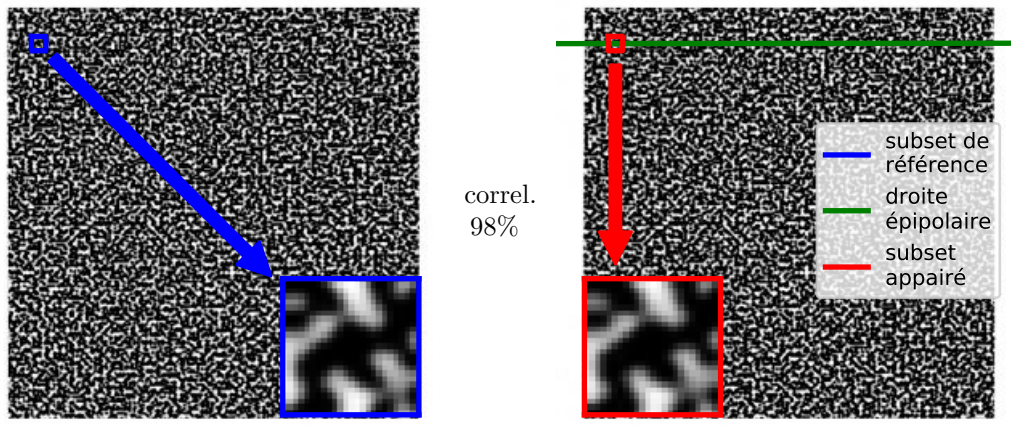


FIGURE 4.27 – Images de la **figure 4.25** rectifiées ; le coefficient de corrélation obtenu est de 98%

Implémentation de la rectification d'images

Comme vu ci-dessus, la position 3D des pixels de la matrice virtuelle choisie (en *rouge*) sur la **figure 4.26** est projetée sur les images initiales afin de calculer une interpolation de niveaux de gris

et d'obtenir des images rectifiées. Cette section décrit le protocole utilisé pour calculer les niveaux de gris de chaque pixel.

La position 3D $(X_c, Y_c, Z_c)^T$ d'un point au centre de l'objet observé est triangulé dans le repère de la mire (*cf.* **section 3.2.5**) en sélectionnant manuellement les positions 2D correspondantes dans les deux images, en s'aidant de la droite épipolaire, comme l'illustre la **figure 4.28**.

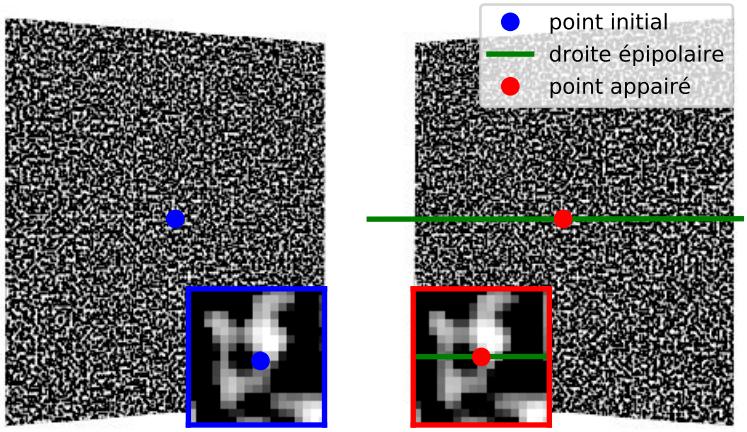


FIGURE 4.28 – Triangulation du point central de l'objet $(X_c, Y_c, Z_c)^T$

Une nouvelle image $(n_{rows} \times n_{cols})$ est créée, dont les coordonnées homogènes des pixels sont $(u_r, v_r, 1)^T$ avec u_r et v_r allant de 0 à $(n_{cols} - 1)$ et $(n_{rows} - 1)$ respectivement.

À partir de la distance du point $(X_c, Y_c, Z_c)^T$ au centre du repère caméra i, et de la taille approximative de l'objet observé (largeur ou hauteur), une nouvelle focale f_{ri} est calculée pour chaque caméra. On a alors :

$$f_{ri} = n_{cols}/width_{obj} * (t_{i3} - Z_c) \tag{4.39}$$

avec $width_{obj}$ la largeur donnée de l'objet et t_3 la troisième valeur du vecteur de translation T_i du repère caméra i.

La position 3D $(X_c, Y_c, Z_c)^T$ est alors projetée sur les plans unitaires [3] des repères caméras rectifiés, qui sont alignés avec le repère mire et centrés autour des centres des repères caméras initiaux. On a :

$$(x_{ci}, y_{ci}, 1)^T = \big((X_c + t_{i1})/(Z_c + t_{i3}), (Y_c + t_{i2})/(Z_c + t_{i3}), 1\big)^T \tag{4.40}$$

avec $(X_c, Y_c, Z_c)^T$ la position 3D centrale triangulée, et $T_i = (t_{i1}, t_{i2}, t_{i3})^T$ le vecteur de translation associé au repère caméra i.

Une nouvelle matrice caméra K_{ri} est calculée par caméra pour que le point $(u_{ci}, v_{ci}) = f_{ri}(x_{ci}, y_{ci})$ soit situé au centre des images rectifiées, selon la formule :

$$K_{ri} = \begin{bmatrix} f_{ri} & 0 & n_{cols}/2 + 0.5 - u_{ci} \\ 0 & f_{ri} & n_{rows}/2 + 0.5 - v_{ci} \\ 0 & 0 & 1 \end{bmatrix} \tag{4.41}$$

avec f_{ri} la nouvelle focale, (n_{rows}, n_{cols}) les dimensions de la nouvelle image, et (u_{ci}, v_{ci}) la position du point central triangulé dans la nouvelle image.

Les positions 3D des pixels $(u_r, v_r, 1)^T$ sont projetées sur les plans images initiaux avec l'équation suivante :

$$k_i * (u_{ri}, v_{ri}, 1)^T = K_i * R_i * K_{ri}^{-1} * (u_r, v_r, 1)^T \tag{4.42}$$

3. Le plan unitaire correspond au plan parallèle au plan $(\vec{x}, \vec{y})$, situé à une distance unitaire du centre du repère caméra.

avec K_i et R_i la matrice caméra et la matrice de rotation du repère caméra i initial, (u_r, v_r) et (u_{ri}, v_{ri}) la position de chaque pixel, respectivement dans le repère image rectifié et dans le repère image initial après reprojection.

Les positions 2D (u_{ri}, v_{ri}) obtenues sont modifiées afin de prendre en compte les distortions optiques des objectifs (*cf.* **section 3.2.4**) puis utilisées pour effectuer une interpolation des niveaux de gris [4] dans les images initiales. Les images rectifiées sont ensuite filtrées avec un filtre gaussien afin de limiter le bruit.

Au final, les paramètres extrinsèques associés à chaque image rectifiée sont : la matrice caméra K_{ri}, le vecteur de translation initial T_i et une matrice de rotation égale à la matrice identité Id. Rappelons que les distortions optiques sont corrigées dans le processus. Un appariement dense, *i.e.* d'un bon nombre de points, est obtenu de manière automatique si l'objet est suffisamment plan, comme l'illustre la **figure 4.29**. Les positions des points de référence (en *bleu*) sont appairées aux positions des points homologues (en *rouge*) dans l'image cible. La corrélation sub-pixélique peut alors être calculée, *cf. supra*.

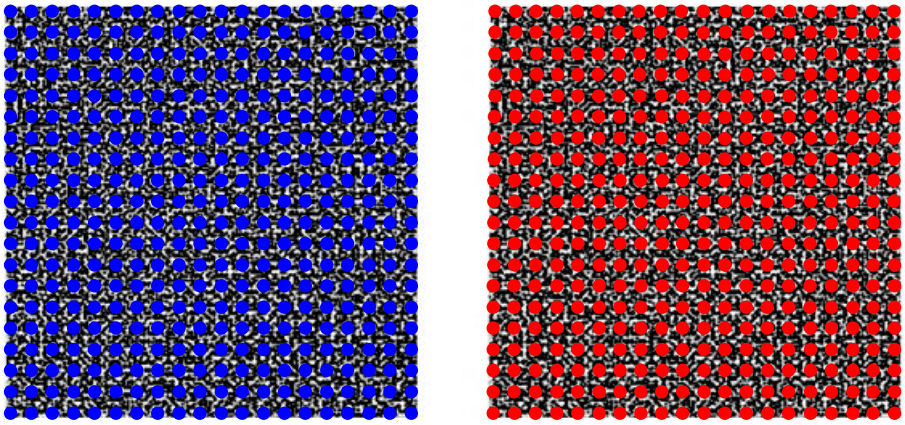


FIGURE 4.29 – Exemple d'appariement effectué avec des images rectifiées

4.7 Suivi de subsets

Avec le suivi de cibles détaillé en **section 4.5**, la position de chaque point de mesure doit être estimée dans toutes les images de la ou des séries temporelles d'images en suivant le déplacement des images des cibles, afin de mesurer des vibrations. De même, pour le suivi de subsets, la position de chaque point de mesure doit être estimée mais cette fois-ci en suivant le déplacement d'image en image de zones de pixels. Cette section décrit la méthode utilisée.

La section précédente a présenté le principe de l'appariement effectué à l'aide de la rectification d'images. Cet appariement étant effectué, il s'agit maintenant de suivre les subsets d'image en image. Pour ce suivi, le calcul d'une paire d'images rectifiées pour chaque image de la série temporelle serait chronophage (de une à plusieurs secondes par image, selon la puissance du processeur utilisé). Dans sa thèse, Garcia [49] montre par ailleurs que travailler avec des images brutes, non rectifiées, apporte une meilleure précision avec les méthodes de corrélation d'images utilisant la déformation de subsets et la méthode Newton-Raphson. Même si, pour ce travail de recherche, le choix s'est porté sur une autre méthode, à savoir celle basée sur le gradient de Davis et Freeman, l'option consistant à travailler sur des images brutes a été appliquée pour le suivi des subsets. Ceux-ci ont été tout d'abord définis dans

4. Sous Python, la fonction "*map_coordinates*" du module *ndimage* de *scipy* est utilisé pour effectuer une interpolation de type *spline* d'ordre 4.

les images initiales non rectifiées. Dans un deuxième temps, un suivi par corrélation et corrélation sub-pixélique a été effectué, avec la méthode du gradient, pour chaque image, sans chercher à ré-appairer les subsets issus des deux points de vue. Les déplacements observés dans le cadre des expérimentations présentées ici sont, comparés à l'image initiale, de l'ordre de un ou deux pixels maximum. Rappelons que, les déformations des images étant relativement petites, la méthode basée sur le gradient s'avère bien adaptée et suffisante : les structures choisies comme objets d'étude ne sont pas très flexibles et l'utilisation de méthodes estimant la déformation des subsets n'est donc pas nécessaire.

Comme mentionné précédemment, avec un système stéréoscopique, des différences de perspective peuvent être visibles dans les images si l'angle entre les caméras est important. Le suivi de subsets carrés ou circulaires équivaudrait à suivre deux surfaces différentes sur l'objet mesuré. L'option choisie ici, illustrée sur la **figure 4.30**, est de définir, dans les images rectifiées (à *gauche*), huit points (en *bleu*) situés sur un cercle (en *vert*) autour de chaque point de mesure (en *rouge*).

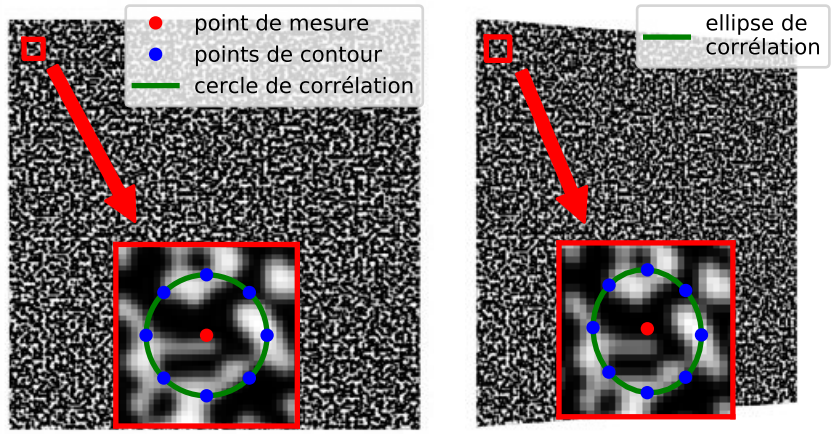


FIGURE 4.30 – Définition du subset de mesure ; (*gauche*) dans l'image rectifiée, (*droite*) dans l'image initiale après reprojection

Ces points sont ensuite projetés dans le repère image initial (à *droite*), avec l'équation (4.42) en tenant compte des distortions optiques, afin de calculer une ellipse (en *vert*) à l'aide de l'estimateur d'ellipses (*cf.* **section 4.4.4**). Les positions et niveaux de gris des pixels à l'intérieur de cette ellipse (*cf.* **section 4.4.3**) sont alors utilisés pour effectuer le suivi par corrélation avec les méthodes présentées en **section 4.6.1**.

4.8 Conclusion

Ce chapitre a présenté les méthodes utilisées pour obtenir les positions 2D des points de mesure nécessaires pour trianguler dans l'espace et le temps les positions 3D correspondantes. Deux approches différentes, adaptées au contexte de cette recherche, ont été retenues, chacune avec un motif différent (cibles ou mouchetis) et des atouts différents. Les cibles sont particulièrement intéressantes pour calibrer un système stéréoscopique mais demeurent plus limitées pour une mesure plein champ. *A contrario*, le mouchetis s'avère particulièrement adapté à la mesure plein champ, avec plus de points de mesure possibles, mais reste limité à la mesure de déplacements relativement petits si l'on ne souhaite pas ré-implémenter les algorithmes complexes d'estimation de déformation des subsets. Les suivis de cibles ou de subsets correspondants ont été présentés, ainsi que les techniques numériques de traitement d'images associées, à savoir les différents détecteurs de cibles circulaires et les outils de corrélation d'images utilisés. Le diagramme synoptique de la démarche (**figure 4.31** ci-dessous) clot ce chapitre, ainsi que la mise en perspective méthodologique du travail présenté dans ce manuscrit.

La partie suivante présente les résultats expérimentaux obtenus, en commençant par la description du protocole utilisé.

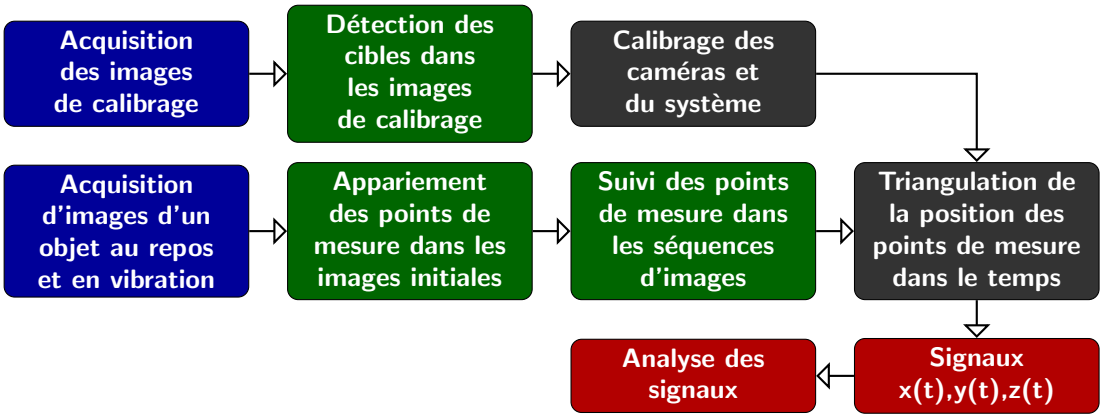


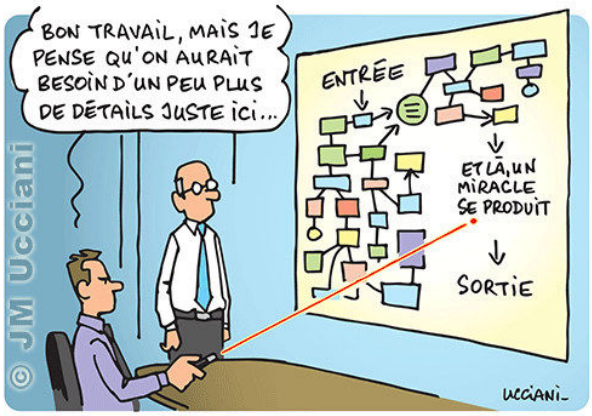
FIGURE 4.31 – Diagramme synoptique général de la démarche adoptée

Deuxième partie

Montages, méthodes et applications

Chapitre 5

Protocole de mesure



« *La science est le capitaine, et la pratique, ce sont les soldats.* »
Léonard de Vinci

Sommaire

5.1 Introduction **79**
5.2 Caméras ultra-rapides **79**
5.3 Mire de calibrage **80**
5.4 Protocole de mesure **82**
5.5 Traitements et logiciels **83**
5.6 Conclusion **85**

5.1 Introduction

La démarche présentée dans la suite de ce manuscrit s'articule autour de l'utilisation d'une seule caméra ultra-rapide, et d'autres caméras, plus ou moins rapides. Suite aux études préliminaires mettant en évidence la pertinence du système pseudo stéréoscopique avec un jeu de quatre miroirs, la partie expérimentale de ce travail s'est, en premier lieu, appuyée sur ce montage, avant de le faire évoluer, comme les chapitres suivants l'exposeront. La question a donc été de cerner les potentialités et les limites du matériel initialement à disposition, ainsi que d'obtenir des données expérimentales permettant de cadrer l'utilisation de ce montage pour la mesure de vibrations. À partir des différentes notions théoriques présentées dans la partie précédente, un protocole général a été défini pour les différentes campagnes de mesures à effectuer, en organisant les étapes incontournables et les opérations associées.

Ce chapitre donne tout d'abord quelques informations sur les caméras rapides utilisées (**section 5.2**), avant de présenter en **section 5.3** les mires de calibrage créées. Le protocole de mesure est ensuite détaillé en **section 5.4**. Finalement, sont abordées les questions du temps de traitement ainsi que des logiciels et/ou programmes développés (**section 5.5**).

5.2 Caméras ultra-rapides

La première caméra rapide utilisée est la Phantom v5.1 du LAUM. Elle a été utilisée pour réaliser une première mesure avec le montage pseudo stéréoscopique et un jeu de quatre miroirs, qui a servi

de mesure-test afin de concevoir le protocole général. Les résultats obtenus ont été présentés au CFA 2016 au Mans [39], mais ne sont pas rapportés dans ce manuscrit. Cette caméra a aussi été utilisée en seconde caméra pour la méthode à caméras asynchrones présentée au **chapitre 10**. Ses caractéristiques sont les suivantes :

— la cadence d'acquisition maximale est de 1200 images par seconde pour une définition de 1020×1024 pixels ;
— les photosites sont carrés et mesurent 12 µm, avec un convertisseur analogique-numérique 8-bits et une sensibilité ISO de 2400 ;
— les valeurs de niveaux de gris de tous les pixels d'un image sont lues simultanément (avec le système *global shutter*).

Les deux caméras ultra-rapides successivement utilisées par la suite sont du modèle Photron SA-X2, et proviennent soit de l'IRT Jules Verne, soit du LAUM. Pour les mesures présentées dans le **chapitre 6**, la caméra de l'IRT Jules Verne a été utilisée ; les chapitres suivants présentent les résultats obtenus avec celle du LAUM. La seule différence notable entre ces deux caméras est leur date de fabrication : la caméra de l'IRT Jules Vernes est plus ancienne et ne possède pas l'option *quiet fan* permettant d'arrêter momentanément les ventilateurs de la caméra. Les autres caractéristiques sont identiques :

— la cadence d'acquisition maximale est de 12 500 images par seconde pour une définition de (1024×1024) pixels ;
— les photosites sont carrés et mesurent 20 µm, avec un convertisseur analogique-numérique 12-bit et une sensibilité ISO de 25 000 ;
— les images de sortie sont codées sous 8 ou 16-bit ;
— les valeurs de niveaux de gris de tous les pixels d'un image sont lues simultanément (avec le système *global shutter*) ;
— le nombre d'électrons correspondant à la charge maximale d'un photosite est de 16 000.

Notons aussi que la caméra rapide du LAUM, avec ses accessoires, a coûté $\approx 60\,000$ €, alors que celle de l'IRT Jules Verne valait $\approx 100\,000$ € quelques années plus tôt. Dans un futur relativement proche, on peut s'attendre à ce que les cadences d'acquisition et les résolutions des caméras rapides augmentent en même temps que leurs prix diminuent, notamment avec le lancement sur le marché grand public de caméras aux performances de plus en plus poussées. En 2016, la caméra GoPro Hero 5 Black, par exemple, affichait une cadence d'acquisition de 240 images par seconde pour des images en couleur de 1280×720 pixels, pour ≈ 400 €, alors que, en 2018, l'appareil photo Sony RX100 IV permet de filmer une vidéo de 1000 images couleurs par seconde avec une résolution de 1920×1080, pour ≈ 900 €.

5.3 Mire de calibrage

Le calibrage est une étape clef de la mesure de vibrations menée avec les méthodes de vision présentées. L'idéal est d'estimer les paramètres du système de mesure avec une précision suffisante, soit des distances épipolaires de quelques dixièmes de pixel, tout en facilitant le processus d'acquisition des images nécessaires et le traitement associé. Plusieurs mires de calibrage ont été testées afin de trouver le compromis optimal entre précision et facilité d'utilisation.

Rappelons que la précision de l'estimation du calibrage dépend de la précision de la détermination des coordonnées des cibles dans le repère 3D associé. Il n'est cependant pas forcément simple de fabriquer des mires de calibrage précises, sachant qu'il est de plus assez coûteux de mesurer précisément la position des cibles correspondantes afin d'avoir des données de référence [86]. Une solution consiste à utiliser, parmi les outils de la bibliothèque OpenCV [20], les méthodes de calibrage proposées (*cf.* **sections 3.2.3** et **3.3.2**), estimant la position des cibles en même temps que les paramètres intrinsèques d'une caméra, et/ou extrinsèques d'un système stéréoscopique. La mire de calibrage peut alors être réalisée à partir d'un papier imprimé sur une imprimante classique.

La première mire de calibrage utilisée, pour les expérimentations du **chapitre 6**, est la mire asymétrique recommandée pour les procédures de calibrage avec la librairie OpenCV. Elle est constituée de 44 disques de 13 mm de diamètre, agencés selon un quadrillage asymétrique, permettant une détection automatique à l'aide des outils proposés par OpenCV. Le papier utilisé est un papier plastique, et l'impression a été réalisée sur une imprimante laser. Une image de la mire posée devant une plaque, observée avec le montage pseudo stéréoscopique avec jeu de quatre miroirs, est affichée sur la **figure 5.1**. Plusieurs inconvénients sont à noter :

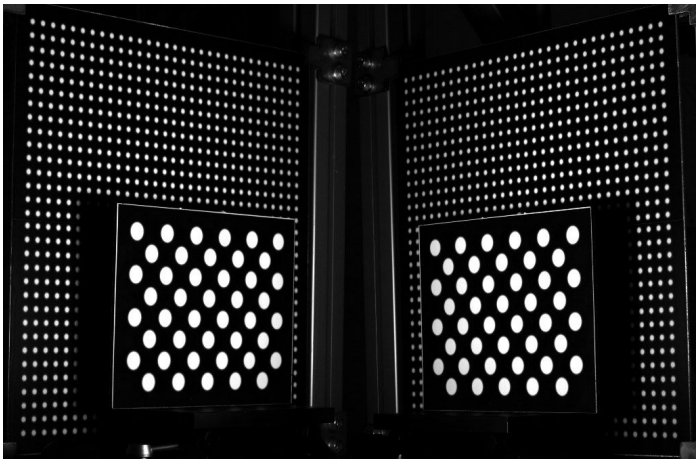


FIGURE 5.1 – Photographie de la mire de calibrage en face de la plaque

— la plaque en aluminium a une épaisseur de 2 mm, et a donc nécessairement été légèrement pliée lors de la découpe ;
— le motif a une dimension fixe, 200 mm × 140 mm au total, ce qui induit un temps certain pour prendre suffisamment de clichés afin d'obtenir des points de mesure sur toute l'image, selon la taille de l'objet mesuré ;
— le motif présente 44 cibles assez grosses (13 mm de diamètre), ce qui implique un écart, de l'ordre de 0,1 pixel dans les cas présentés **chapitre 6**, entre le centre de l'image d'un point et la projection du centre du point sur l'image [36] ;
— le motif point des disques s'avère assez sensible aux erreurs d'impression, de par la discontinuité associée, et ne peut être parfaitement imagé (*cf.* **section 6.3.2**).

De plus, l'algorithme de détection utilisé par OpenCV est le détecteur de blobs, alors que, comme le montre le **chapitre 6**, un motif point-Gaussien couplé au détecteur d'ellipses duales [114] améliore la précision des mesures. En conséquence, des codes originaux ont été développés afin de créer des images de mire avec un motif et un agencement adaptés au contexte, et ainsi d'optimiser l'adéquation des mires utilisées et des objets d'étude choisis.

Deux agencements ont finalement été retenus lors des expérimentations présentées dans ce manuscrit : un quadrillage de cibles (*cf.* **figure 5.2** *gauche*) et une disposition en spirale (*cf.* **figure 5.2** *droite*). Dans les deux cas il s'agit d'un motif point-Gaussien (*cf.* **section 6.3.2**).

Le quadrillage de cibles est principalement utilisé pour le calibrage des caméras, puisqu'il permet d'obtenir facilement des images de cibles dans les coins de l'image, mais aussi pour le calibrage de systèmes stéréoscopiques avec des objets volumineux (*cf.* **chapitre 10**). Le tri des cibles est effectué à partir de quatre points sélectionnés avec une interface graphique, comme présenté en **section 4.4.7**.

L'agencement en spirale est, quant à lui, particulièrement pratique pour calibrer les systèmes (pseudo) stéréoscopiques lorsque l'objet à mesurer présente une géométrie circulaire. Cette disposition permet un tri réalisé de manière complètement automatique (*cf.* **section 4.4.7**).

Pour les mires utilisées dans ce travail, l'image de la mire est tout d'abord imprimée sur un papier auto-collant non plastifié avec une imprimante laser ou jet d'encre, au format A4 ou A3, en fonction

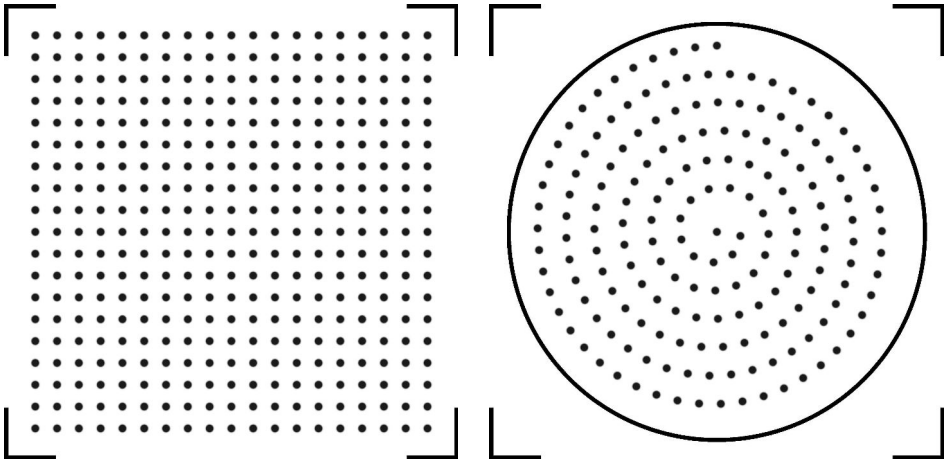


FIGURE 5.2 – Exemples d'images utilisées pour la fabrication de mires ; (*gauche*) quadrillage de cibles, (*droite*) disposition en spirale

des besoins. Le papier est ensuite posé à plat sur une table, avant de poser dessus une plaque en aluminium de 8 mm d'épaisseur, suffisamment épaisse pour éviter tout pliage lors de sa découpe. Ce protocole permet de limiter très fortement la déformation du papier lors du collage. Le motif étant connu, la principale incertitude sur la position des cibles vient de l'impression. On peut ainsi vraisemblablement s'attendre à des incertitudes de l'ordre de 0,1 mm, mais cette valeur n'a pas été vérifiée lors de ce travail. Notons aussi qu'à l'impression le motif point-Gaussien ne présente pas de discontinuité (contrairement au disque du motif point) mais des bords qui suivent un profil gaussien : on peut alors supposer que les erreurs d'impression auront, dans ce cas, moins d'impact sur la précision des calibrages.

5.4 Protocole de mesure

Les différentes méthodes de vision présentées dans ce manuscrit s'appuient sur un protocole commun : même si certains calculs diffèrent, le processus général reste identique. Le diagramme synoptique global est présenté sur la **figure 5.3**. On peut distinguer les étapes suivantes :

1. (**noir**) Réglage du montage, acquisition des images initiales et enregistrement de la ou des vidéos d'un objet en vibrations ;
2. (**vert**) Détection et appariement des points d'intérêt dans les images initiales puis suivi d'image en image dans la vidéo ;
3. (**bleu**) Triangulation des positions 3D des points avec les données de calibrage (de la caméra et du système) et de détection ;
4. (**rouge**) Analyse avec les outils de traitement du signal afin de trouver les fréquences de résonance et les déformées modales associées.

Réglage du montage et acquisition des images : la première étape vise à obtenir les images nécessaires pour mesurer les vibrations d'un objet. La ou les caméras, les éventuels miroirs et l'éclairage doivent être positionnés de telle manière que l'objet soit imagé sur le plus grand nombre de pixels possibles avec une illumination suffisante. Une fois ce réglage effectué, les images nécessaires au traitement doivent être enregistrées, à savoir :

— les images de la mire de calibrage pour estimer les paramètres intrinsèques de chaque caméra ;
— les images de la mire de calibrage pour estimer les paramètres extrinsèques du système (pseudo) stéréoscopique ;
— la ou les séquences temporelles d'images de l'objet en vibration.

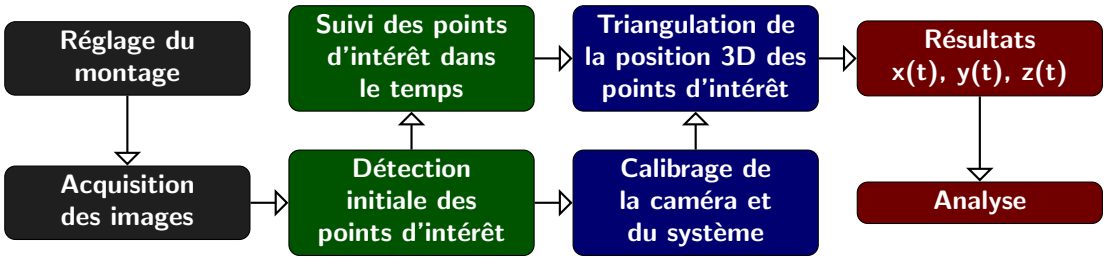


FIGURE 5.3 – Diagramme synoptique du protocole de mesure

Notons que l'ordre d'acquisition des images peut varier selon le montage, comme en feront état les chapitres suivants. De plus, avec un montage à deux caméras ou bien avec le montage sans miroir (*cf.* **chapitre 9**), on obtient un point de vue par image dans les séries temporelles, alors qu'avec les montages pseudo stéréoscopiques (*cf.* **chapitre 6**), deux points de vue sont présents par image.

Détection et suivi des points d'intérêt : une fois les images enregistrées, les positions 2D des points d'intérêt doivent être estimées dans toutes les images afin de calculer dans un deuxième temps, par triangulation, les positions 3D correspondantes. Les outils présentés au **chapitre 4** sont alors utilisés.

— Les positions des images des cibles sont estimées dans les images correspondantes de la mire, afin de réaliser le calibrage de la ou des caméras et du système (pseudo) stéréoscopique.
— La position 2D de chaque point de mesure doit être connue dans les images de l'objet au repos, acquises avec les deux caméras (virtuelles) du système (pseudo) stéréoscopique, soit en détectant et triant la position 2D des cibles, soit avec les outils de corrélation d'images. Les cibles ou les subsets sont ensuite suivis, d'image en image, dans la ou les séries temporelles d'images, avant d'effectuer la triangulation.

Triangulation : la ou les caméras sont d'abord calibrées, avant de calibrer le système (pseudo) stéréoscopique à partir des positions 2D détectées dans les images de la mire. La triangulation des positions 3D, dans le temps, des points de mesure, peut ensuite être effectuée avec les positions 2D des points de mesure dans toutes les images de la ou des séries d'images. Il en résulte trois signaux de positions, selon les trois axes de l'espace, pour chaque point de mesure.

Analyse fréquentielle : les outils de traitement du signal et d'analyse modale (*cf.* **section 1.3**) sont ensuite utilisés pour extraire différentes informations des signaux de déplacement 3D des points de mesure, notamment :

— la précision de la mesure ;
— les fréquences de résonance ;
— les déformées modales.

Rappelons que la mesure de vibrations et l'analyse modale sont utilisées afin de comprendre des phénomènes vibratoires et *in fine* d'améliorer le confort et la sécurité. Dans le cadre du travail présenté dans ce manuscrit, le travail s'arrête volontairement à l'étape de l'affichage et de la lecture de cartes de déformées modales [1].

5.5 Traitements et logiciels

Un des enjeux clefs de cette partie du travail a été de fluidifier la phase de mesures et le traitement associé, afin d'obtenir idéalement un retour d'information rapide pour vérifier la validité de la mesure *in situ*. Un des objectifs a donc été de simplifier le protocole général, sachant que la partie traitement était nécessairement réalisée à l'aide d'un ordinateur, de logiciels ou de programmes adaptés. Cependant, comme rappelé en *Avant-propos*, la mesure de vibrations par vision 3D est à l'intersection de

1. La mesure des vibrations d'un haut-parleur a toutefois été utilisée pour une application électro-acoustique ,*cf.* **Annexe B**, afin d'en déterminer les paramètres électro-mécaniques.

plusieurs champs disciplinaires, notamment la photogrammétrie / vision 3D / vision stéréoscopique, le traitement d'images, l'analyse vibratoire, ou encore l'informatique ou l'optique. Il a alors fallu trouver une solution pour faciliter l'utilisation des méthodes proposées par un public n'ayant pas forcément de connaissances poussées dans chaque domaine. Dans ce contexte, l'utilisation de divers logiciels pré-existants n'a pas semblé forcément pertinente, d'autant que la rapidité de traitement est centrale voire même vitale si l'on souhaite obtenir une vérification rapide de la mesure.

En effet, comme rappelé en **section 2.6.3**, les logiciels de photogrammétrie et de corrélation d'images étaient initialement conçus pour traiter relativement peu d'images. Depuis l'amélioration des caméras ultra-rapides, ces logiciels ont dû s'adapter à la demande de chercheurs et d'industriels qui souhaitent réaliser des mesures avec des milliers d'images [9]. Toutefois, les codes ne sont généralement pas accessibles et un certain degré de figement freine l'apparition de nouvelles méthodes de traitement. Il a alors semblé indispensable de concevoir des codes adaptés afin de maitriser la chaine complète et d'accélérer le plus possible le temps de traitement. La question du language de programmation ou du logiciel utilisé s'est alors posée. Un choix encore standard dans la communauté vibro-acoustique et de traitement du signal est le logiciel Matlab, qui s'avère relativement onéreux puisque chaque *toolbox* est payante. De plus, Matlab est optimisé pour le calcul matriciel mais présente un temps de traitement des boucles assez long. Une autre possibilité est de coder en language C++, qui permet d'obtenir des codes rapides si les algorithmes sont optimisés. Néanmoins, la difficulté de réalisation des parallélisations et le problème de compatibilité d'un ordinateur à un autre (sans parler des systèmes d'exploitation) ont fait écarter cette option. Finalement, le choix s'est porté sur une troisième option : le language de programmation Python. En effet, Python offre un bon compromis entre *open source*, compatibilité entre les différents systèmes d'exploitation, rapidité de traitement et facilité d'utilisation. Lors de l'écriture des codes, une attention particulière a été portée aux différentes étapes du processus de traitement, afin de gérer les points difficiles. La création quasi-totale des codes a grandement facilité la création des nouvelles méthodes présentées dans ce manuscrit : la méthode IRIs (*cf.* **chapitre 7**), la méthode sans miroir (*cf.* **chapitre 9**) et la méthode à caméras asynchrones (*cf.* **chapitre 10**).

Gestion de la grande quantité d'images

Les images fournies par les caméras rapides peuvent être stockées sous différents formats : une vidéo au format .avi non compressée, des images au format .tif, des fichiers .tif multi-pages, etc. Des tests ont aussi été réalisés avec des caméras du commerce (*cf.* **Annexe B**), fournissant des vidéos au format .mp4 par exemple. Afin d'homogénéiser le traitement, toutes les images obtenues, quel que soit le format initial, doivent être converties dans un format spécifique. De plus, la gestion de milliers de fichiers ralentit logiquement la copie d'un disque dur à l'autre, augmente la taille des données et nécessite autant d'ouvertures et de fermetures de fichiers qu'il y a d'images. Cet aspect est un des éléments clefs à prendre en compte pour diminuer le temps de traitement, sachant que, en plus, la conversion des séquences d'images en films .avi non compressés n'est pas évidente, à cause des codecs notamment, et ne permet pas facilement une cadence de lecture élevée. Le choix a alors été de stocker les images en format binaire, par lot de 100, 200 ou 1000 par exemple, en fonction du choix de l'utilisateur, ce qui permet une gestion relativement simple des données.

Traitement des images et temps de calcul

La détection et surtout le suivi des points d'intérêt dans les séries temporelles d'images est la partie la plus critique en termes de temps de traitement. Afin de limiter temporellement cette étape, les codes ont été parallélisés : chaque processus lit chaque fichier binaire d'images, puis suit le déplacement d'une partie des points d'intérêt dans chaque image. L'ouverture multiple de chaque fichier binaire d'images pénalise peu le temps de traitement, puisqu'en utilisant un disque SSD (*Solid State Drive*), l'opération ne nécessite guère plus d'une seconde pour un fichier de 1000 images. De plus, les données ne sont écrites qu'à la fin du processus de suivi sous un fichier binaire unique. Pour exemple, suivre 3200 subsets elliptiques de 205 pixels en moyenne, sur 400 images prend ≈ 10 secondes avec un processeur Intel® Xeon CPU E3-1505M v6, intégrant huit cœurs à 3 GHz. Si l'on suit ces 3200 subsets

sur 23 000 images, le fichier de détection résultant est de $\approx$ 600 Mo pour un temps de traitement d'environ 10 minutes, ce qui permet de vérifier assez rapidement la validité des mesures *in situ*.

OpenCV

La bibliothèque OpenCV, qui propose notamment des fonctions de calibrage de caméras et de systèmes stéréoscopiques, est implémentée sous C++ et Python. Ainsi, pour plus de simplicité, les codes développés utilisent le package d'OpenCV correspondant codé sous Python. Ce package est aussi très rapide pour ouvrir des images et effectuer des filtrages convolutifs, tout en permettant de créer des vidéos à partir de figures ou d'images réalisées sous Python.

Triangulation, traitement du signal et analyse modale

Afin de maximiser la souplesse et l'évolution des codes, les différentes fonctions de triangulation ont aussi été codées, afin notamment de simplifier les calculs de la méthode sans miroir (*cf.* **chapitre 9**). Finalement, des outils de traitement du signal et d'analyse modale ont aussi été conçus dans l'optique de gérer un grand nombre de signaux.

5.6 Conclusion

Ce chapitre introductif à l'exposé des résultats expérimentaux a tenté de clarifier le protocole général suivi tout en donnant les caractéristiques techniques des caméras utilisées et en présentant les mires de calibrage conçues en adéquation avec les objets d'étude choisis, à savoir une plaque, un haut-parleur, un tambourin et un capot de voiture. De plus, des codes spécifiques ont été développés pour presque toute la chaîne de traitement (à l'exception du calibrage dont les fonctions proviennent de la bibliothèque OpenCV), afin d'homogénéiser les données et de fluidifier le processus général de la mesure, c'est-à-dire de l'acquisition des images jusqu'à l'affichage des déformées modales, en passant par la détection et le suivi des points de mesure, le calibrage, la triangulation et le traitement des données. Finalement, ce choix permet d'obtenir un retour d'information rapide et ainsi d'évaluer directement, sur place, la validité des campagnes de mesure effectuées. Les chapitres suivants abordent les résultats obtenus pour les différents montages et objets d'étude, en commençant par le montage pseudo stéréoscopique appliqué à la mesure de vibrations d'une plaque.

Chapitre 6

Pseudo stéréo-vision pour la mesure de vibrations

« Un miroir est une surface polie, faite pour réfléchir, mais parfois bien impolie quand elle vous fait réfléchir. »
Gérard de Rohan Chabot

Sommaire

6.1 Introduction **87**
6.2 Système pseudo stéréoscopique avec jeu de quatre miroirs **88**
6.2.1 État des lieux des utilisations du montage 88
6.2.2 Principe et dimensionnement 89
6.2.3 Calibrage 92
6.3 Étude de sensibilité **93**
6.3.1 Approches et critères d'évaluation 93
6.3.2 Étude de sensibilité liée à la détection 94
6.3.3 Étude de sensibilité liée à l'agencement du système pseudo stéréoscopique . 97
6.4 Application à une plaque **100**
6.4.1 Matériel utilisé 100
6.4.2 Acquisition des images, calibrage et bruit de fond 102
6.4.3 Mesure de vibrations - évaluation de la précision 102
6.4.4 En bref... 105
6.5 Application à un haut-parleur **107**
6.5.1 Description du montage 107
6.5.2 Données de calibrage 108
6.5.3 Mesure de vibrations - évaluation de la précision 109
6.5.4 Bilan 115
6.6 Conclusion **116**

6.1 Introduction

Comme mentionné précédemment, la stéréo-vision permet d'avoir une mesure synchrone d'un grand nombre de points de mesure. Elle est applicable à des phénomènes non répétables, non linéaires et non stationnaires. Couplée à des caméras ultra-rapides, elle permet d'effectuer des mesures de déplacements et ainsi de vibrations (*cf.* **chapitre 3**). La méthode de stéréo-vision requiert deux points de vue de l'objet d'étude, et donc *a priori* l'utilisation de deux caméras avec la phase de synchronisation associée. Les deux acquisitions doivent en effet être lancées au même instant grâce

aux déclencheurs des caméras qui activent l'enregistrement des images. Pour une synchronisation précise, la paire de caméras doit posséder un système d'entrée et de sortie de synchronisation. Dans le cas contraire, les deux caméras ne démarrent pas leur cycle de pré-acquisition au même moment : on peut observer un décalage allant jusqu'à un demi-temps d'échantillonnage. Si les caméras ultra-rapides offrent généralement cette option, ce n'est pas nécessairement le cas des caméras industrielles qui peuvent actuellement enregistrer quelques milliers d'images par seconde. Notons qu'il existe des logiciels, payants, qui permettent globalement de régler ce problème, tel que VIC-Snap software de Correlated solutions, par exemple.

Un autre inconvénient de la méthode de stéréo-vision vient du coût des deux caméras : la caméra ultra-rapide utilisée pour cette thèse, la Photron SA-X2 (avec une cadence d'acquisition de 12 500 images par seconde avec (1024 × 1024) pixels et ses accessoires) avoisine 60 000 €, soit 120 000 € pour le système à deux caméras. Afin de réduire les coûts, d'autres alternatives ont vu le jour. Citons, par exemple, le montage avec bi-prismes [53, 73, 52] qui reste limité à de petits angles entre les caméras, d'autant que les bi-prismes induisent d'importantes distortions optiques. On peut aussi mentionner le montage avec caméra RGB-D (Red, Green, Blue, Depth) [144, 81, 37]. La caméra RGB fournit une information sur la distance en comparant un motif aléatoire connu projeté par un émetteur infra-rouge et son image reçue sur un capteur infra-rouge. Cette technique est actuellement limitée à trente images par seconde environ, avec un précision relativement faible [109]. À la suite de l'étude de faisabilité [142], le travail de Biggio [16], préliminaire à cette thèse, a permis d'adapter à la mesure de vibrations un montage utilisé dans le cadre de la robotique principalement [74, 108, 40, 165] et n'utilisant qu'une seule caméra rapide avec un jeu de miroirs : le "montage pseudo stéréoscopique à caméra unique et jeu de quatre miroirs", considéré comme le plus polyvalent de ce type de montages [116] [1]. Biggio ayant réussi à mesurer une forme et une déformation statique sur une plaque, les premières étapes du travail présenté dans ce manuscrit se sont articulées autour de la mesure de vibrations avec ce même montage, ne nécessitant qu'une seule caméra ultra-rapide. Le fil conducteur de cette phase de recherche a alors été d'optimiser sa configuration, et les algorithmes associés, afin d'obtenir la plus grande précision possible pour la mesure de vibrations effectuée avec cet équipement.

Ce chapitre s'articule de la façon suivante. La **section 6.2** décrit le système pseudo stéréoscopique à caméra unique et jeu de quatre miroirs. Après un état des lieux de ses utilisations actuelles (**section 6.2.1**), son principe est expliqué, ainsi que le dimensionnement choisi (**section 6.2.2**). Les résultats de l'étude de sensibilité menée afin d'optimiser certains paramètres du montage sont présentés en **section 6.3**. Les données correspondantes ont été appliquées, dans un premier temps, à la mesure de vibrations d'une plaque (**section 6.4**). Dans un deuxième temps, le protocole a été adapté à la mesure de vibrations d'un haut-parleur (**section 6.5**). Dans les deux cas, les résultats sont comparés avec ceux obtenus par deux méthodes de référence.

6.2 Système pseudo stéréoscopique à caméra unique et jeu de quatre miroirs

6.2.1 État des lieux des utilisations du montage

Dans la littérature, la première utilisation du capteur pseudo stéréoscopique avec une seule caméra ultra-rapide, couplé à un jeu de quatre miroirs, est attribuée à Inaba *et al.* [74] en 1993. Cette étude de faisabilité, et la validation expérimentale associée, visait des applications en robotique, notamment pour le suivi de déplacements 3D par un capteur stéréoscopique intégré à un robot. En 2005, Seal *et al.* [136] ont utilisé un système similaire pour la mesure de profondeur jusqu'à une distance de deux mètres. En 2006, Zhu *et al.* ont mené des analyses détaillées du modèle optique et proposé un protocole de calibrage du système [188]. Le champ de vision associé est étudié et optimisé en 2008 par Wang *et al.* [165]. Le capteur a été appliqué avec succès à un éventail de sujets allant de la mesure de la géométrie 3D de la surface de flammes [2] à la photographie stéréo de banderoles en mouvement [110],

1. Notons qu'un montage avec une caméra unique et six miroirs a été réalisé pour générer quatre points de vue [126].

en passant par la reconstruction de trajectoires de bulles dans un fluide [177], par exemple, mais aussi au suivi d'objets en temps réel [69].

Les premiers travaux mentionnant l'association du capteur pseudo stéréoscopique et des techniques de corrélation d'images sont attribués à Pankow *et al.* [117], en 2010, pour la mesure de grands déplacements hors-plan avec de hautes cadences d'acquisition d'images. Ces dernières années, le système a été de plus en plus utilisé pour la mesure 3D de formes et de déformations, dans des environnements à températures très élevées, hors laboratoire [90] ou encore dans le cadre d'analyse modale de poutres en porte-à-faux [28]. On le trouve aussi par exemple dans des études portant sur la mesure de bandes de cisaillement adiabatiques [171], de vibration de plaques partiellement encastrées [182], avec une caméra ultra-rapide monochrome ou même couleur pour mobiliser toute la matrice de pixels [181].

6.2.2 Principe et dimensionnement

Comme expliqué précédemment, le système pseudo stéréoscopique utilise une seule caméra. Le but est d'obtenir, grâce à l'utilisation de miroirs, deux points de vue d'un même objet sur la matrice de pixels de cette unique caméra. Afin d'optimiser le processus, le capteur doit être soigneusement dimensionné afin de maximiser la taille des images et de prévoir les éléments opto-mécaniques correspondants. Il s'agit ainsi de comprendre le fonctionnement du système avant de calculer, à partir de sa géométrie, différents paramètres essentiels en fonction de l'objet d'étude et de la configuration angulaire choisie, notamment la taille des miroirs, la distance entre les miroirs latéraux ou encore la distance entre la caméra et l'objet d'étude.

Les objectifs de cette section sont de présenter :
- le principe du capteur pseudo stéréoscopique avec une caméra et un jeu de quatre miroirs ;
- les étapes de calcul des paramètres géométriques requis pour faciliter le réglage du montage et prévoir les éléments opto-mécaniques nécessaires.

Principe

Cette solution propose de générer deux caméras virtuelles à partir d'une caméra réelle, d'où le terme pseudo stéréoscopique. Deux miroirs, ou un miroir à angle droit, sont placés devant l'objectif de la caméra afin de séparer le point de vue en deux. Les deux points de vue virtuels sont ensuite ré-orientés vers l'objet d'étude à l'aide de deux autres miroirs latéraux (*cf.* **figure 6.1**). Le montage est réalisé sur une table optique et s'utilise logiquement en laboratoire. Pour une potentielle utilisation hors laboratoire, une conception mécanique spécifique, entre autres, est requise [90]. Ce montage présente néanmoins un nombre de pixels disponibles divisé par deux dans la largeur (1024×512 pour une définition initiale de 1024×1024). De plus, les miroirs peuvent ne pas être parfaitement plans et induisent des distortions dans les images.

Dimensionnement

L'étape de dimensionnement vise à calculer la taille optimale des miroirs afin de faciliter le réglage du montage tout en maximisant la taille des images obtenues. À partir des caractéristiques de la caméra et de l'objectif, ainsi que de la taille de la zone à mesurer, il est ainsi possible de calculer les différents paramètres liés à l'agencement des miroirs, à savoir (*cf.* **figure 6.2**) :
- la base B ou distance entre les centres optiques des deux caméras virtuelles ;
- la distance D entre le centre des caméras virtuelles et la surface à mesurer ;
- la distance D' entre le centre optique de la caméra réelle et la surface à mesurer ;
- la distance n entre la caméra et le miroir central ;
- la distance m entre les miroirs centraux et chaque miroir latéral ;
- l'angle δ entre les miroirs latéraux et la droite à 45° ;
- les angles Fov_h et Fov_v du champ de vision horizontal et vertical de la caméra.

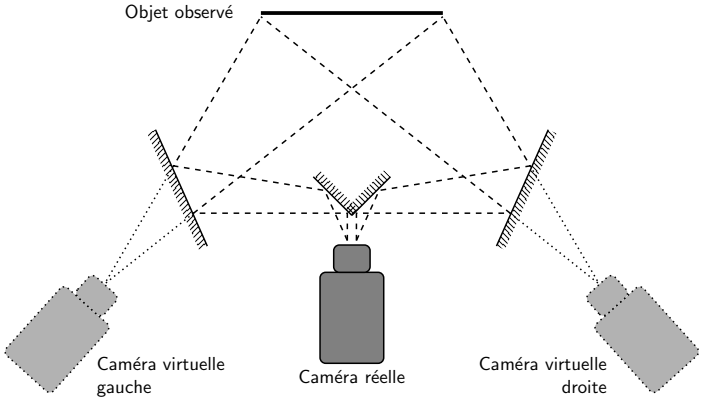


FIGURE 6.1 – Schéma du montage pseudo stéréoscopique avec quatre miroirs

La sensibilité aux déplacements hors-plan de l'objet étant fortement liée à l'angle entre les caméras et la surface [151], on cherche à calculer les distances m et D', à l'aide de la **figure 6.2**, pour les paramètres donnés n, Fov_h, Fov_v et δ.

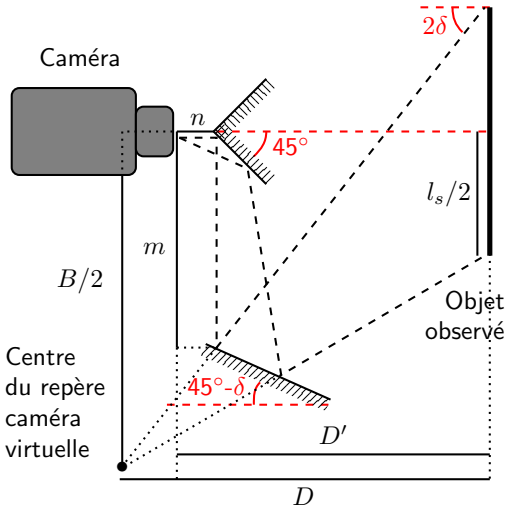


FIGURE 6.2 – Schéma utilisé pour le dimensionnement

Agencement des miroirs Afin d'optimiser l'utilisation de la matrice de pixels, les distances D, D', B et m sont calculées pour un angle de rotation 2δ entre les repères caméras et la surface mesurée. Les champs de vision Fov_h et Fov_v sont calculés tels que :

$$Fov_h = 2 * \arctan(l_c/2f) \tag{6.1}$$

et

$$Fov_v = 2 * \arctan(h_c/2f) \tag{6.2}$$

avec l_c et h_c respectivement la largeur et la hauteur de la matrice de pixels, et f la longueur focale. La hauteur de la surface mesurée et la taille de son image sont liées par :

$$h_s * \frac{\cos(2\delta - Fov_h/2)}{D\cos(Fov_h/2)} = 2\tan(Fov_v/2) \tag{6.3}$$

avec h_s la hauteur de la surface mesurée. La distance D, calculée de manière à ce que la hauteur de l'image de la surface observée soit imagée sur toute la hauteur du capteur pour le rayon perspectif en bordure du champ de vision, est donc :

$$D = h_s * \frac{\cos(2\delta - Fov_h/2)}{2\tan(Fov_v/2)\cos(Fov_h/2)} \quad . \tag{6.4}$$

La largeur de la base est ensuite calculée telle que le bord de la surface observée se projette au bord de la matrice de pixels :

$$B = 2\big(l_s/2 + D * \tan(2\delta - Fov_h/2)\big) \tag{6.5}$$

avec l_s la largeur de la surface mesurée. Pour la distance m, on note d'abord que :

$$B/2 = m + (m+n) * \sin(2\delta) \tag{6.6}$$

et donc que :

$$m = \frac{B/2 - n * \sin(2\delta)}{1 + \sin(2\delta)} \quad . \tag{6.7}$$

Finalement, pour le calcul de D' on obtient :

$$D' = D - (n+m) * \cos(2\delta) + n \quad . \tag{6.8}$$

Taille des miroirs Afin de faciliter la formulation des équations concernant la taille des miroirs, la représentation est simplifiée en un schéma optique "déplié" [99], tel que représenté sur la **figure 6.3**. Les dimensions requises des miroirs sont ainsi facilement calculables à l'aide de la "loi des sinus". Les

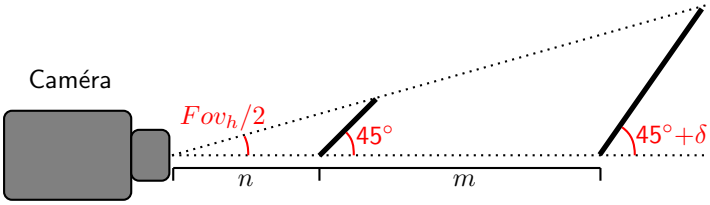


FIGURE 6.3 – Schéma optique déplié

largeurs requises l_{m1} et l_{m2}, pour le miroir central et les miroirs latéraux respectivement, sont :

$$l_{m1} = n * \frac{\sin(Fov_h/2)}{\sin(\pi/4 - Fov_h/2)} \tag{6.9}$$

et

$$l_{m2} = (n+m) * \frac{\sin(Fov_h/2)}{\sin(\pi/4 + \delta - Fov_h/2)} \quad . \tag{6.10}$$

Selon le modèle sténopé (*cf.* **section 3.2.2**), les hauteurs h_{m1} et h_{m2} correspondantes sont :

$$h_{m1} = (n + \frac{lm1}{\sqrt{2}}) * 2\tan(Fov_v/2) \tag{6.11}$$

et

$$h_{m2} = (n + m + l_{m2} * \cos(\pi/4 + \delta)) * 2\tan(Fov_v/2) \quad . \tag{6.12}$$

Application numérique Une application numérique est réalisée en prenant comme paramètres :
- la distance n égale à 60 mm ;
- la matrice de pixel fixée à 672 × 1024, avec une taille de pixel de 20 µm (correspondant à celle de la Photron SA-X2 à disposition) ;
- une focale de 50 mm ;

— une surface observée de 330 mm × 330 mm ;
— un angle entre chaque caméra virtuelle et la surface allant de 30°à 55°.

Les dimensions requises ($h_{m1} \times l_{m1}$) obtenues pour les miroirs centraux sont 20,3 mm × 21,9 mm. La **figure 6.4** montre les résultats pour les miroirs latéraux, avec l_{m2} en *noir* et h_{m2} en *rouge*. Pour 2δ égal à un minimum de 30°, ($h_{m2} \times l_{m2}$) est environ égal à 125 mm × 110 mm. Pour 2δ égal à un maximum de 55°, ($h_{m2} \times l_{m2}$) est environ égal à 135 mm × 170 mm.

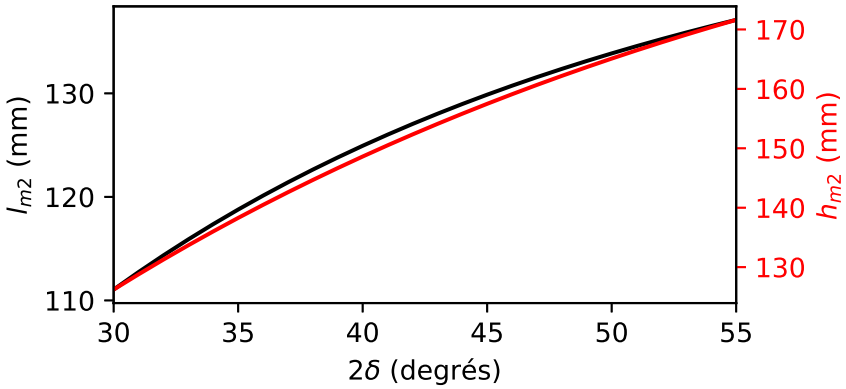


FIGURE 6.4 – Dimensions requises l_{m2} et h_{m2} pour les miroirs latéraux

Les miroirs choisis, plaqués argent de chez Thorlabs, sont :
— un prisme central *knife edge* (*cf.* **figure 6.5**) avec deux côtés en miroir, d'une surface de 25 mm × 25 mm ;
— deux miroirs latéraux carrés de 200 mm × 200 mm.

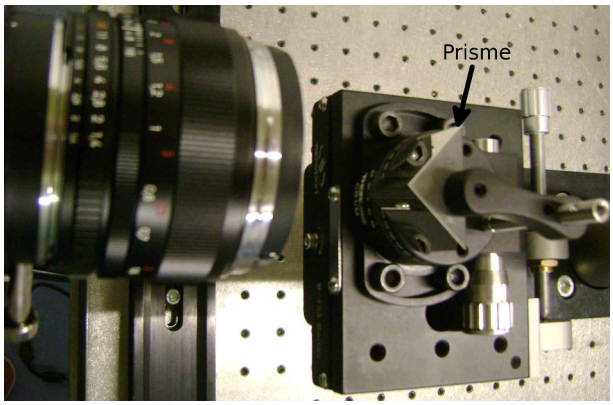


FIGURE 6.5 – Photographie du prisme central *knife edge*

Pour l'agencement du système (*cf.* **figure 6.6**), la largeur du système, liée à m en *noir*, augmente avec l'angle 2δ, alors que la distance D' en *rouge*, entre la caméra réelle et la surface observée, diminue. En fonction de l'angle choisi, la distance entre les miroirs latéraux peut ainsi dépasser 1 m.

6.2.3 Calibrage

Comme expliqué en **section 5.4**, il est préférable de calibrer la caméra indépendamment du système pseudo stéréoscopique, or la présence des miroirs centraux juste devant l'objectif entrave le processus. Plusieurs solutions sont envisageables :

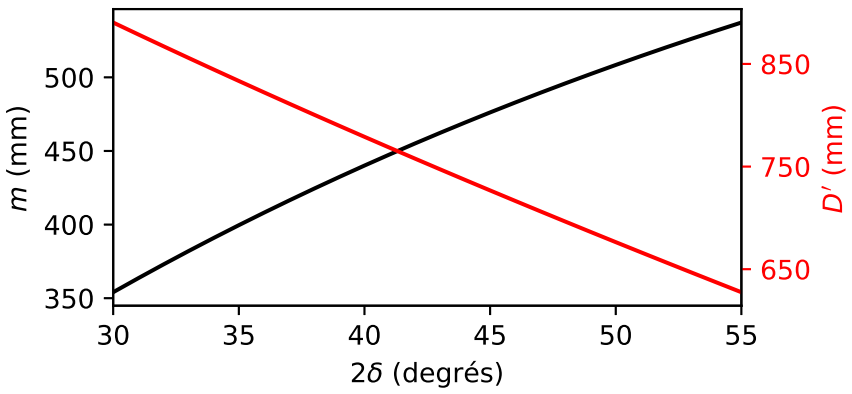


FIGURE 6.6 – Distances m et D' en fonction de l'angle entre les caméras virtuelles et une surface observée de 330 mm × 330 mm

1. Calibrer les deux caméras virtuelles indépendamment, à partir d'images de la mire (vue à travers les miroirs), ce qui induit deux jeux différents de paramètres intrinsèques [180].
2. Calibrer la caméra à partir de clichés de la mire (vue à travers les deux miroirs), en masquant la moitié de l'image à chaque cliché [186]. Les images de la mire seront alors petites et ne pourront pas forcément couvrir toute la surface de la matrice de pixels.
3. Déplacer les miroirs avant ou après la mesure de vibrations, afin de calibrer la caméra indépendamment du système pseudo stéréoscopique et de fixer les paramètres intrinsèques des caméras virtuelles.

Cette troisième solution est celle qui a été retenue. Le système pseudo stéréoscopique est d'abord mis en place, en positionnant les miroirs et en réglant la netteté, afin d'obtenir les images désirées de la surface observée pour les deux caméras virtuelles. Il est ensuite nécessaire de réaliser les clichés utilisés pour le calibrage de la caméra et du système pseudo stéréoscopique ainsi que l'enregistrement de la ou des vidéos de mesure de vibrations. Deux conditions sont à respecter :

— les miroirs ne doivent pas être déplacés entre l'enregistrement des vidéos et l'acquisition des clichés de la mire à travers les miroirs ;
— la caméra doit être déplacée le moins possible afin d'éviter de modifier les paramètres de l'objectif.

Une solution consiste, une fois le montage réglé, à déplacer le miroir central ou bien le support de la caméra sur la table optique. On peut ainsi prendre les clichés de la mire pour le calibrage de la caméra, puis replacer le miroir central ou la caméra et refixer le tout sur la table. La position des miroirs peut être légèrement modifiée mais l'objectif de la caméra ne doit pas être touché. Les clichés de la mire vue à travers les miroirs sont ensuite réalisés pour calibrer le système avant d'enregistrer la ou les vidéos de l'objet en vibration.

Une autre solution consiste à prendre les clichés de calibrage de la caméra après l'enregistrement des vidéos. C'est cette dernière solution qui a été retenue dans ce travail.

6.3 Étude de la sensibilité de la précision de la mesure à certains paramètres

6.3.1 Approches et critères d'évaluation

Dans la littérature, il semble y avoir globalement deux approches afin d'obtenir la précision de la mesure [131] : soit théorique, avec des modèles et des simulations [50, 45, 176, 135, 89, 4, 98], soit expérimentale [85, 30, 1, 58].

Les approches théoriques permettent de comprendre les limites de la précision d'un système stéréoscopique et l'impact des erreurs d'estimation des différents paramètres sur la précision de la reconstruction 3D. Toutefois, elles ne donnent pas une valeur précise de l'incertitude expérimentale. *A contrario*, les différentes approches expérimentales comparent généralement des valeurs mesurées à des données connues *a priori*, par exemple des différences de distances mesurées au distance-mètre. Cependant, ces données connues sont toujours imprécises. Une autre méthode consiste à mesurer un même objet dans différentes positions et à comparer les différentes formes mesurées [131, 58].

Dans le cadre de ce travail, l'objectif, à ce stade, est d'optimiser la mesure de vibrations effectuée avec un système stéréoscopique et une méthode de suivi de cibles. Une démarche en deux temps a ainsi été choisie :

— tout d'abord, une approche théorique, sous la forme d'une étude de sensibilité de la précision de la mesure à certains paramètres ;
— ensuite, une approche expérimentale, sous la forme de comparaison systématique des résultats avec ceux obtenus avec des méthodes de mesure de référence.

Pour la première étape, le logiciel Blender [18] a été choisi. Il s'agit d'un logiciel de simulation et d'animation 3D, intégrant un calculateur de tracé de rayons. Il a déjà été utilisé pour des essais de mesure de déformations quasi-statiques avec des outils de corrélation d'images, notamment afin d'étudier la possibilité de déplacer la caméra si une mire reste fixe autour de l'objet testé, lors de mesures effectuées sur de longues périodes [96].

Ce logiciel, indépendant des codes développés pour cette thèse, a permis de réaliser des simulations afin d'effectuer une étude de sensibilité et d'obtenir des valeurs d'incertitudes approximatives. L'objectif n'est pas de quantifier précisément l'incertitude par la simulation, mais d'étudier l'influence de certains paramètres sur la précision de la mesure de petits déplacements. Les paramètres étudiés sont :

— le motif des cibles ;
— les algorithmes de détection à disposition ;
— l'agencement du système : l'angle entre les caméras virtuelles et la surface mesurée.

Dans un premier temps, une simulation simple d'une plaque devant une caméra est effectuée pour déterminer le ou les meilleurs motifs et algorithmes de détection dans ce contexte. Dans un deuxième temps, une simulation d'un système pseudo stéréoscopique est réalisée afin de comprendre l'influence de l'agencement du système sur la précision de la mesure. Notons que Blender a plusieurs solveurs, dont un permettant de réaliser des simulations précises avec un tracé de rayons, le *cycle render*. C'est ce solveur qui est utilisé pour les simulations présentées dans ce manuscrit.

Critère d'évaluation

Le signal de déplacement étant défini, il est possible de calculer une erreur, égale à la différence entre le signal mesuré et le signal simulé. En choisissant plusieurs centaines de points de mesure sur plusieurs images, l'écart-type de l'erreur, noté σ, semble un critère pertinent pour évaluer l'influence des différents paramètres étudiés. A chaque simulation est calculé un écart-type, associé à un signal d'erreur, pour chaque point de mesure. Les deux valeurs retenues sont la valeur moyenne et la valeur maximale de cet écart-type.

6.3.2 Étude de sensibilité liée à la détection

Une première étape consiste à étudier l'impact du motif des cibles ainsi que des détecteurs présentés en **section 4.4**. Comme mentionné précédemment, un des objectifs est la maximisation du nombre de cibles sur la surface de mesure, ce qui induit que chaque cible soit imagée avec relativement peu de pixels, environ 10×10. Trois motifs sont alors testés : un motif "point", un motif "Gaussien" et un motif "point-Gaussien". La fonction qui permet de calculer le profil de chaque motif est :

$$f(x, r_1, r_2) = \begin{cases} 1, & \text{si } |x| \leq r_1. \\ exp\left(-\frac{(|x|-r_1)^2}{(2r_2^2/9)}\right), & \text{sinon.} \end{cases} \tag{6.13}$$

avec r_1 le rayon de valeur continue et r_2 la distance de décroissance. La **figure 6.7** affiche les trois motifs correspondant aux valeurs de r_1 et r_2 données dans le **tableau 6.1**. Le motif point en *rouge* présente une discontinuité qui ne peut être parfaitement imagée, contrairement aux motifs Gaussien et point-Gaussien dont les bords sont adoucis.

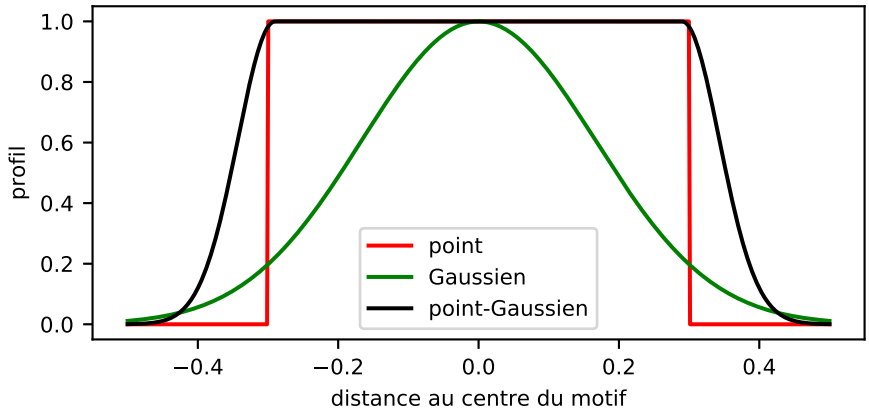


FIGURE 6.7 – Profils des motifs utilisés

TABLEAU 6.1 – Rayons relatifs à la taille de la cible pour les trois motifs

	point	Gaussien	point-Gaussien
r_1	0,3	0	0,2
r_2	0	0,5	0,2

Pour la simulation réalisée, une plaque présentant une image constituée de 31 × 31 cibles est positionnée en face d'une caméra. La taille des images est de 360 × 360 pixels. Cent images de la plaque sont simulées avec un déplacement égal à un centième de pixel entre deux images, le long de l'axe $\overrightarrow{u}$ du repère image, qui sera l'axe principal de déplacement lors des expérimentations. Les niveaux de gris vont de 4 à 173. Pour plus de réalisme, un bruit photonique (*cf.* **section 3.2.1**) est ajouté, en prenant en compte les caractéristiques de la caméra Photron SA-X2 (*cf.* **section 5.2**).

La **figure 6.8** présente une partie des images initiales obtenues pour les trois motifs. L'ellipse initialement détectée pour une cible est tracée en *vert* et l'ellipse-masque en *rouge*. Des irrégularités dans les images obtenues avec le motif point sont visibles à *gauche*, de par le repliement spectral. Elles sont visiblement moins importantes pour les motifs Gaussien, au *centre*, et point-Gaussien à *droite*. Notons finalement que le motif Gaussien permet d'obtenir une image présentant moins d'énergie qu'avec les deux autres motifs.

Le centre des cibles est ensuite suivi d'image en image tel que présenté en **section 4.5**. Pour chaque cible, à chaque image, une sous-image contenant la cible est copiée de l'image initiale et masquée, avec l'ellipse-masque centrée autour de la position détectée à l'étape précédente. La sous-image masquée est filtrée à l'aide d'un filtre Gaussien de largeur égale à cinq pixels (*cf.* **section 4.2**). Chaque algorithme de détection est ensuite utilisé pour détecter le centre du motif dans la sous-image. Pour plus de simplicité, la même sous-image, masquée avec l'ellipse centrée autour de la position détectée par l'estimateur d'ellipse duale à l'image précédente, est utilisée pour tous les algorithmes. Finalement, l'écart-type de l'erreur observée pour chaque point de mesure est calculé. Les valeurs moyenne σ_{moy} et maximale σ_{max} sont affichées par détecteur et par motif dans le **tableau 6.2** pour les axes $\overrightarrow{u}$ et $\overrightarrow{v}$ du repère image.

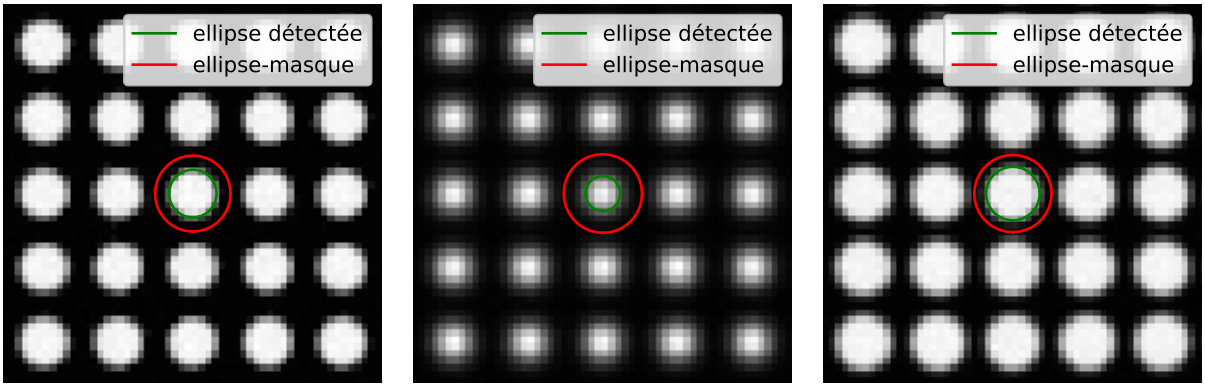


FIGURE 6.8 – Images initiales et ellipse-masque d'une cible ; (*gauche*) motif point, (*centre*) motif Gaussien et (*droite*) motif point-Gaussien

TABLEAU 6.2 – Résultats des calculs d'erreur pour les différents détecteurs de cibles, exprimés en centième de pixels et affichés au format σ_{moy} / (σ_{max})

Motif	axe	centroïde	détect. blobs	estim. ellipses	estim. ell. duales
Point	$\vec{u}$	5,4 / (6,4)	16,8 / (17,3)	1,0 / (1,7)	0,8 / (1,6)
	$\vec{v}$	1,2 / (9,9)	2,5 / (30,0)	0,6 / (1,3)	0,6 / (1,3)
Gaussien	$\vec{u}$	11,3 / (11,7)	22,1 / (23,0)	0,8 / (1,0)	1,3 / (1,5)
	$\vec{v}$	3,2 / (20,0)	6,1 / (38,6)	0,8 / (1,3)	0,9 / (2,2)
Point Gaussien	$\vec{u}$	4,9 / (5,6)	16,2 / (16,5)	0,5 / (0,9)	0,4 / (0,6)
	$\vec{v}$	0,7 / (9,1)	1,7 / (28,3)	0,5 / (0,7)	0,4 / (0,6)

En ce qui concerne les motifs, le classement du moins précis au plus précis donne l'ordre suivant : Gaussien, point et point-Gaussien, avec des différences moins marquées pour ces deux derniers lorsqu'ils sont utilisés avec l'estimateur d'ellipses. En effet, le motif Gaussien est plus petit, renvoyant moins de lumière que les deux autres. Quant au motif point-Gaussien, qui limite les irrégularités dans l'image visibles sur la **figure 6.8**, il permet, comme on s'y attendait, d'améliorer la précision par rapport au motif point.

En ce qui concerne les algorithmes, le classement selon le même critère donne l'ordre suivant : détecteur de blobs, centroïde, estimateur d'ellipses et estimateur d'ellipses duales. En effet, l'algorithme de détecteur de blobs calcule différentes images binaires, à partir de l'image initiale, pour différentes valeurs seuils de niveaux de gris. Ceci semble incompatible avec une résolution aussi faible pour imager les cibles. Pour les trois motifs et suivant les deux axes, on note que les résultats les plus précis sont obtenus grâce aux détecteurs d'ellipses et d'ellipses duales. De plus, les valeurs du temps de calcul relatif, obtenues pour trois images, *i.e.* $3 * (31 \times 31) = 2883$ détections, sont reportées dans le **tableau 6.3**. Le centroïde est visiblement le détecteur le plus rapide. Il est aussi à noter que l'estimateur d'ellipses duales est environ trois fois plus rapide que l'estimateur d'ellipses, qui nécessite la recherche des points du contour de l'ellipse.

Finalement, les paramètres retenus pour la simulation suivante, ciblant l'étude de l'impact de l'agencement du système pseudo stéréoscopique sur la précision, sont :

— le motif point-Gaussien ;
— les estimateurs d'ellipses et d'ellipses duales.

TABLEAU 6.3 – Temps de calcul relatif des différents algorithmes

Détecteur	centroïde	détect. blobs	estim. ellipses	estim. ell. duales
Temps de calcul relatif	1	15,5	60,5	22,4

6.3.3 Étude de sensibilité liée à l'agencement du système pseudo stéréoscopique

L'objectif de cette section est d'étudier l'influence de l'agencement angulaire des miroirs afin de déterminer le positionnement qui maximise la précision de la mesure. Un système pseudo stéréoscopique est simulé dans différentes configurations, ici avec différents angles entre les caméras virtuelles et la surface mesurée.

La **figure 6.9** affiche une capture d'écran de la visualisation 3D de la simulation réalisée. Une caméra est virtuellement placée devant un système pseudo stéréoscopique, composé d'un cube réfléchissant (simulant le miroir central) et de deux surfaces réfléchissantes localisées sur les côtés (simulant les miroirs latéraux). Le tout est orienté de manière à observer une plaque de dimensions 330 mm × 330 mm, présentant une image composée de 31 × 31 cibles selon un quadrillage régulier avec un espacement de 10 mm entre chaque cible. De plus, une lampe est positionnée en hauteur entre la caméra et la plaque, afin de simuler le spot.

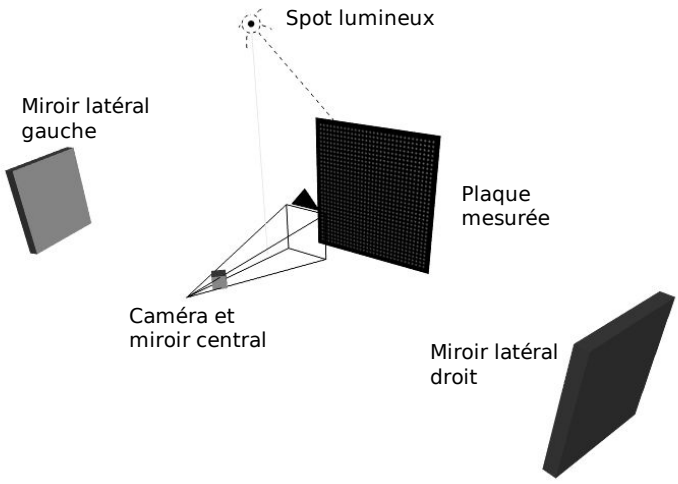


FIGURE 6.9 – Capture d'écran de la simulation réalisée sous Blender

Plus précisément, la surface des miroirs est définie comme plane et parfaitement réfléchissante avec un matériau *glossy BSDF*. Le miroir central est un cube de 25 mm de côté et les miroirs latéraux sont des plaques de 200 mm × 200 mm. Pour la plaque étudiée, une surface *diffuse BSDF* est utilisée. Une image, créée préalablement, permet de générer le quadrillage des cibles. Le motif utilisé est point-Gaussien.

Pour la caméra, les paramètres sont réglés selon les caractéristiques du modèle Photron SA-X2. La taille de la matrice de pixels est fixée à 1024 × 1024 avec une largeur de pixel égale à 20 µm. L'objectif est simulé avec une longueur focale de 50 mm, la distorsion optique est réglée à 0,2% et la durée d'ouverture du diaphragme à 25% du temps d'échantillonnage. La défocalisation liée à l'objectif n'est pas prise en compte.

En augmentant l'angle 2δ, les distortions de perspective augmentent et, conjointement, l'image de la plaque sur les plans images virtuels s'affine. La hauteur de l'image peut alors être augmentée, de manière à maximiser l'utilisation de la matrice de pixels de la caméra réelle : comme le montre la **figure 6.10**, l'image obtenue pour un angle $2\delta = 50°$ (*droite*) présente une surface plus importante que l'image obtenue pour un angle de 30° (*gauche*).

En ce qui concerne les vidéos, la fréquence d'acquisition est normalisée, *i.e.* réglée à une image par seconde. La plaque se déplace, sans déformation, selon un signal $s(t)$, composé de trois sinus, tels

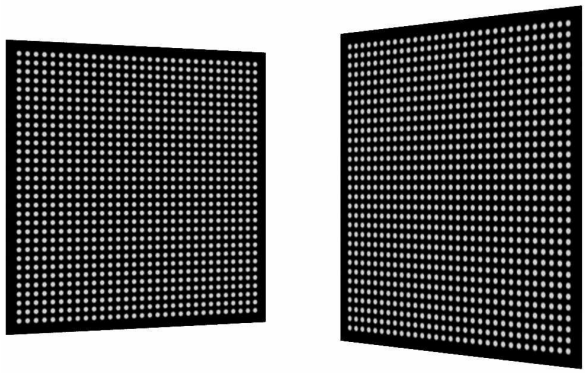


FIGURE 6.10 – Images obtenues pour un angle 2δ égal à 30° (*gauche*) et 50° (*droite*)

que décrits dans le **tableau 6.4**. Les fréquences de vibration normalisées par la cadence d'acquisition sont de 1/50, 1/10 à 1/4 et produisent un nombre entier de périodes pour 100 points temporels. Le vecteur de déplacement est colinéaire à la normale de la plaque, et correspond à l'axe $\overrightarrow{z}$ du repère de triangulation.

Composante	fréquence (normalisée)	amplitude (µm)
1	1/50	120
2	1/10	70
3	1/4	100

TABLEAU 6.4 – Composantes du signal de déplacement de la plaque

Pour chaque configuration angulaire, des images de la mire vues à travers les miroirs ainsi que cent images de la plaque en vibration, avec 961 points de mesure par image, sont simulées, avec un bruit photonique ajouté, en prenant toujours les caractéristiques de la caméra Photron SA-X2. La position de chacune des 961 cibles est initialement détectée dans les premières images virtuelles et suivie d'image en image, tel que présenté en **section 4.5**. Après avoir calibré la caméra et les systèmes correspondants aux angles 2δ allant de 30° à 55° par pas de 5°, et une fois les positions 2D des cibles estimées dans toutes les images, les signaux de vibration sont triangulés. Un signal d'erreur et un écart-type sont alors calculés pour le signal de déplacement selon l'axe $\overrightarrow{z}$ de chaque point de mesure. Les valeurs moyennes et maximales de l'écart-type en fonction de l'angle 2δ sont affichées sur la **figure 6.11**. On constate que les résultats obtenus avec l'estimateur d'ellipses duales sont plus précis que ceux obtenus avec l'estimateur d'ellipses. De plus, l'écart-type moyen diminue avec l'angle et atteint environ deux microns pour $2\delta = 50°$. Notons que pour des angles supérieurs à 55°, les images des cibles sont très étroites et donc difficilement détectables.

Si l'augmentation de l'angle 2δ améliore la mesure de déplacements selon l'axe $\overrightarrow{z}$, un biais est nécessairement induit, puisque la projection des centres des cibles est décalée par rapport aux centres des ellipses observées dans les images [36]. La forme triangulée dans le plan moyen de la surface, pour un angle $2\delta = 45°$, est affichée sur la **figure 6.12** et semble être courbe, alors qu'elle devrait être plane. L'écart-type calculé est de 27 µm.

Si l'on compare cette valeur avec l'écart-type moyen correspondant selon l'axe $\overrightarrow{z}$ pour la mesure de déplacement, σ_z = 2,3 µm, un facteur d'environ 12 est obtenu. La mesure de vibrations, dans cette configuration, est donc plus précise que la mesure de forme. De plus, si l'on passe dans le domaine fréquentiel, l'erreur sur la mesure d'amplitude est moins importante que l'erreur dans le domaine temporel [76]. En effet, si l'on prend pour exemple un signal sinusoïdal d'amplitude A, de fréquence f_0, $sig(t)$ = A $\sin(2\pi f_0$ t), l'amplitude peut être calculée à l'aide de la transformée de

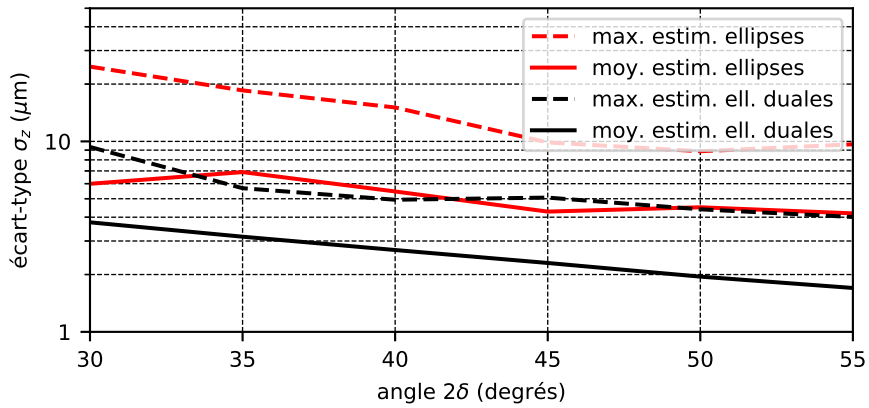


FIGURE 6.11 – Écart-type du signal de triangulation selon l'axe $\vec{z}$ en fonction de l'angle 2δ ; les lignes continues correspondent à la valeur moyenne et les lignes en pointillé à la valeur maximale

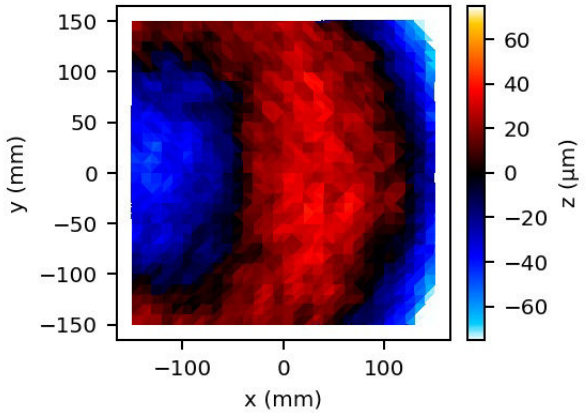


FIGURE 6.12 – Forme triangulée dans le repère du plan moyen de l'objet, à partir des images de la plaque avec $2\delta = 45°$

Fourier (*cf.* **section 1.3.1**) :

$$A(f) = \frac{1}{T}\int_{-\infty}^{+\infty} sig(t)exp(-2j\pi ft)dt \quad , \tag{6.14}$$

avec T la durée de la fenêtre d'échantillonnage afin d'obtenir un valeur d'amplitude consistante. Pour une fréquence d'échantillonnage égale à $10 * f_0$ et une amplitude A = 10 µm, si l'on ajoute un bruit d'écart-type égal à 2,3 µm et que l'on calcule A pour 1 000 signaux de 50 périodes, l'écart-type de l'erreur sur l'amplitude calculé est d'environ 0,15 µm, c'est à dire quinze fois inférieur à l'écart-type dans le domaine temporel. Ceci induit un facteur d'environ $12 * 15 = 180$ par rapport à l'erreur calculée pour la mesure de forme : la mesure de vibrations, dans cette configuration et dans le domaine fréquentiel, est ainsi ≈ 180 fois plus précise que la mesure de forme.

Rappelons que ces données sont obtenues par le biais de simulations : les valeurs sont nécessairement approximatives, probablement sous-estimées. Néanmoins, dans le cadre de cette étude de sensibilité, l'information recherchée n'est pas tant la valeur précise de l'erreur que l'impact des paramètres et donc la précision relative des différentes configurations. Les paramètres pris en compte sont le motif des cibles, le détecteur associé et l'agencement du système, à savoir l'angle entre la surface mesurée

et les caméras virtuelles. On retiendra que :
- le motif point-Gaussien améliore la précision par rapport aux autres motifs étudiés ;
- l'estimateur d'ellipses duales présente l'écart-type le plus faible ;
- l'augmentation de l'angle entraine une amélioration de la précision de la mesure, tant que les cibles sont vues avec suffisamment de pixels ;
- avec des cibles circulaires, dans le contexte de cette recherche, la mesure de déplacements sub-millimétriques est plus précise que la mesure de forme.

6.4 Application à une plaque

La phase de dimensionnement et l'étude de sensibilité étant réalisées, l'étape suivante est l'application expérimentale. Ainsi, une mesure sur une plaque a été effectuée et les résultats ont été comparés à ceux obtenus avec des systèmes de référence : un accéléromètre et un vibromètre laser à balayage.

6.4.1 Matériel utilisé

Montage pseudo stéréoscopique

Une photographie de l'agencement du système est présentée sur la **figure 6.13**. La caméra ultra-rapide est montée sur un trépied Manfrotto 117B, avec une rotule Manfrotto 229. Le miroir central est un miroir *knife edge*, présentant deux surfaces réfléchissantes carrées de 25 mm de côté. Il est fixé sur une platine de rotation, elle-même vissée à une platine de translation afin de faciliter le réglage. L'ensemble est posé sur un support attaché à un tube permettant de régler la hauteur du miroir.

Les miroirs latéraux, plaqués aluminium, sont des plaques carrées de 200 mm de côté environ. Chacun est attaché à un tube à l'aide d'un support pour miroir décalé. Ce tube est fixé sur une platine de rotation, elle-même vissée sur une glissière. Les deux glissières sont posées sur un rail de 1 m de large (provenant de Melles Griot). La liste du matériel utilisé pour le système pseudo stéréoscopique est affichée dans le **tableau 6.5**.

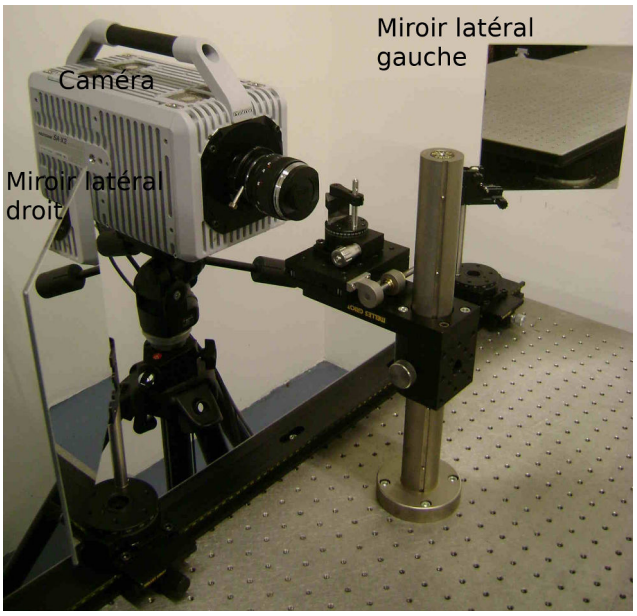


FIGURE 6.13 – Photographie du système pseudo stéréoscopique monté sur une table optique

En ce qui concerne la planéité des miroirs, les valeurs de tolérance données par le constructeur, mesurées par des techniques interférométriques, sont :
- pour les miroirs centraux : un huitième de longueur d'onde à 632,8 nm, *i.e.* 79 nm ;
- pour les miroirs latéraux : cinq longueurs d'onde à 632,8 nm, *i.e.* 3,16 µm.

Tableau 6.5 – Références du matériel utilisé

Elément	description	fournisseur	référence
Miroirs centraux	miroir *knife edge* à angle droit	Thorlabs	MRAK25-P01
	platine de rotation Ø 35 mm	Thorlabs	CR1/M
	pince ajustable pour prisme	Thorlabs	PM4/M
	platine de translation	Newport	M-433
Miroirs latéraux	miroir carré en aluminium 203,2 mm	Thorlabs	ME8S-G01
	platine de rotation Ø 73 mm	Newport	M-481-A
	support de miroir décalé	Newport	M1-RQ
	glissière pour rail de guidage	Laser Components	LCG.CRL-2

Objet mesuré, éclairage et instrumentation

L'objet mesuré est une plaque en aluminium de dimensions 330 mm × 330 mm × 1,5 mm, encastrée aux quatre coins sur un chassis réalisé avec des tubes Norcan. Quatre auto-collants imprimés avec une imprimante A4 sont collés, et présentent un quadrillage de 31 × 31 cibles, espacées de 10 mm, avec un motif point-Gaussien. La plaque est éclairée avec un projecteur Led Ledvance IP65 Noir 200 W 4000 K, avec un angle de diffusion de ≈100°, placé entre le miroir central et la plaque (*cf.* **figure 6.14** *gauche*).

Une tête d'impédance mécanique, composée d'un accéléromètre et d'un capteur de force, est placée derrière le point de mesure situé sur la 20e ligne et la 13e colonne du quadrillage de cibles (*cf.* **figure 6.14** *droite*). Les données obtenues avec l'accéléromètre permettent d'avoir une référence de vibration en un point, tandis que le capteur de force permet de calculer des fonctions de réponse en fréquence, déplacement sur force d'excitation. Un pot vibrant, servant de source d'excitation, est relié à la tête d'impédance à l'aide d'une corde de piano et suspendu à une barre verticale. La plaque et son support, ainsi que le pot vibrant, sont placés sur une petite table optique posée sur quatre plots anti-vibrations sur la grande table optique afin de limiter la transmission de vibrations vers les miroirs. Le logiciel Intac [2], ainsi qu'une carte National Instrument sont utilisés pour générer le signal d'excitation et enregistrer les données obtenues avec l'accéléromètre et le capteur de force. Un amplificateur Audiolab est utilisé pour amplifier le signal électrique d'excitation à appliquer aux bornes du pot vibrant.

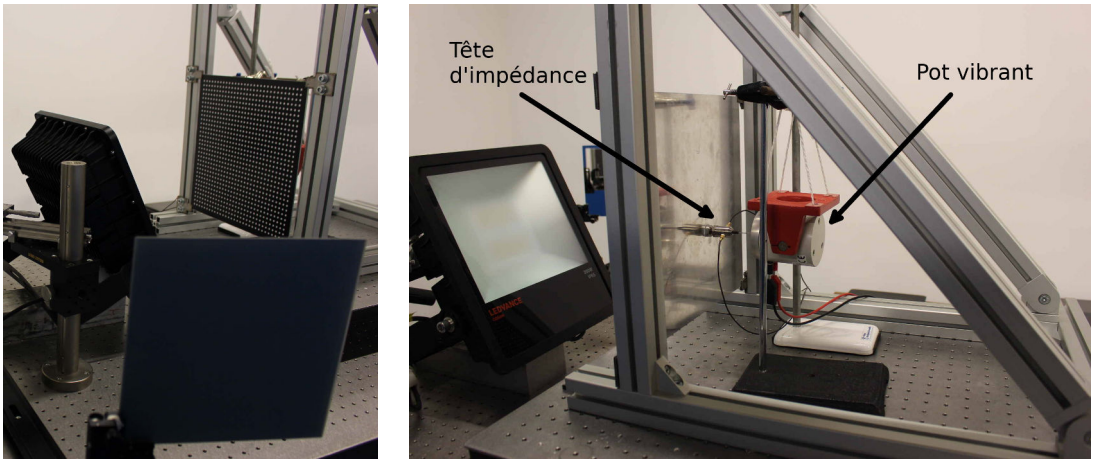


Figure 6.14 – Photographie de la plaque vue de face (*gauche*) et du pot vibrant et de la tête d'impédance (*droite*)

2. http://www.cttm-lemans.com/fr/acoustique-vibrations/produits/logiciels.html

6.4.2 Acquisition des images, calibrage et bruit de fond

Afin de calibrer la caméra et le système de manière indépendante, le protocole présenté en **section 6.2.3** a été appliqué. Les étapes en sont :

1. réglage du montage pseudo stéréoscopique afin d'obtenir les images désirées, avec une taille d'image de 672 × 1024 pixels ;
2. réalisation de clichés de la mire à travers le montage ;
3. enregistrement des vidéos de mesure de vibrations ;
4. déplacement les éléments afin de prendre les clichés de la mire pour calibrer la caméra.

Pour cette mesure, la mire de calibrage asymétrique recommandée lors de l'utilisation des procédures de calibrage avec la librairie OpenCV [20] a été utilisée. La mire est réalisée avec un auto-collant imprimé avec une imprimante A4 et collé sur une plaque d'aluminium de dimensions 200 mm × 140 mm × 2 mm, usinée par le service technique de le Mans Université. Le quadrillage asymétrique est constitué de quarante-quatre points de 13 mm de diamètre. La **figure 6.15** présente un cliché réalisé avec la mire de calibrage devant la plaque.

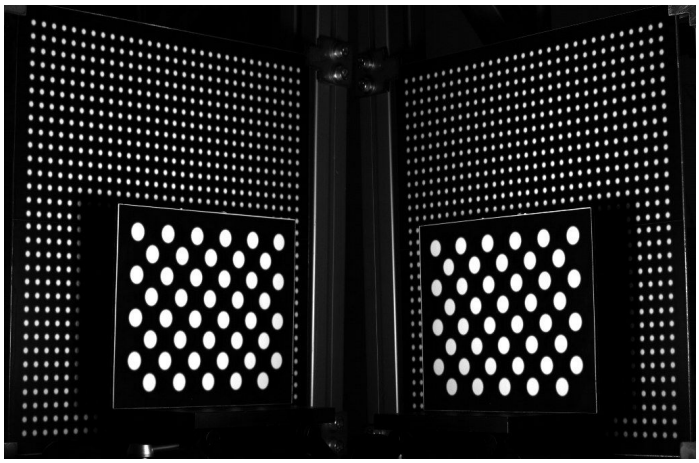


FIGURE 6.15 – Image obtenue avec le montage pseudo stéréoscopique avec la mire de calibrage devant la plaque

Le calibrage de la caméra donne une focale de 56,12 mm. Le calibrage du système donne une base B égale à 2 185 mm avec des angles de rotation autour des axes $\overrightarrow{y}$, $\overrightarrow{x}$ puis $\overrightarrow{z}$ de 104,01°, 0,81° et -1,20° respectivement. Notons qu'en prenant en compte les données du calibrage obtenues, à savoir la focale (56,12 mm) et un angle $2\delta = 104/2°$, pour des images de dimension 672 × 1024 pixels et une hauteur de surface mesurée de 330 mm, la longueur de la base obtenue avec les équations du dimensionnement (*cf.* **section 6.2.2**) est $B = 2192$ mm. L'écart entre la valeur obtenue avec le calibrage et avec la simulation est de 7 mm, soit 0,3%.

Cependant, en calculant la distance épipolaire, *i.e.* la distance entre un point homologue et la droite épipolaire correspondante (*cf.* **section 3.3.3**), la valeur moyenne obtenue pour les points de la mire est de ≈ 0,3 pixel, avec une valeur maximale de ≈ 1,2 pixel. Ces valeurs sont relativement importantes et sont probablement imputables à la taille des cibles de la mire couplée au fort angle panoramique utilisé.

6.4.3 Mesure de vibrations - évaluation de la précision

La caméra rapide utilisée est la Photron SA-X2 de l'IRT Jules Vernes. Notons que cette caméra a été achetée il y a quelques années, et que l'option *quiet fan* qui permet d'arrêter momentanément les ventilateurs n'était pas encore disponible. La cadence d'acquisition de la caméra rapide est réglée à

12 500 images par seconde. Une première mesure est réalisée, sans excitation, avec 3 000 images afin de vérifier le bruit de fond. L'écart-type du déplacement mesuré pour chaque point, idéalement nul, est calculé. La carte résultante est affichée sur la **figure 6.16** et présente un minimum de 1,4 µm au centre de la plaque et un maximum de 2,8 µm aux coins.

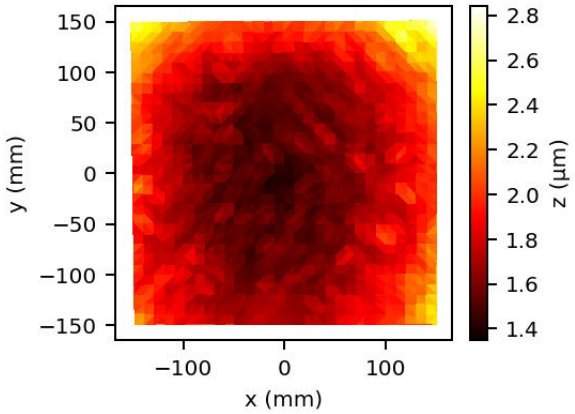


FIGURE 6.16 – Carte de l'écart-type du déplacement mesuré sur la plaque au repos

Puis la plaque est excitée à l'aide du pot vibrant, afin de calculer la fonction de réponse en fréquence de chaque point. Le signal d'excitation utilisé est un sinus glissant linéaire, allant de 20 Hz à 2 500 Hz en 3,5 secondes. Comparé à un signal aléatoire qui répartit de manière uniforme l'énergie sur la bande de fréquences d'intérêt, le sinus glissant a l'avantage de maximiser l'énergie, à un instant donné, sur une petite bande de fréquences. Il en résulte un déplacement plus important à chaque fréquence, ce qui améliore le rapport signal à bruit.

Le spectrogramme du signal mesuré selon l'axe $\overrightarrow{z}$ avec le montage pseudo stéréoscopique correspondant à la cible située au niveau de l'excitation est affiché sur la **figure 6.17**. Il est obtenu en calculant une transformée de Fourier sur 250 ms de signal, toutes les 20 ms, et permet de tracer l'évolution du spectre du signal dans le temps. Ici, la fréquence du signal augmente avec le temps et l'amplitude diminue, comme on pouvait s'y attendre (*cf.* **chapitre 1**). Des lignes verticales sont cependant visibles et correspondent à des vibrations mono-fréquentielles continues. En observant le spectrogramme du signal mesuré par l'accéléromètre [3], sur la **figure 6.18**, ces lignes ne sont pas visibles. En observant les spectrogrammes correspondants pour les axes $\overrightarrow{x}$ et $\overrightarrow{y}$ affichés sur les **figures 6.19** et **6.20**, les lignes verticales sont plus marquées vers 20 Hz, 75 Hz, 130 Hz et 270 Hz. L'hypothèse est que ces vibrations sont celles de la caméra rapide engendrées par les quatre ventilateurs. L'amplitude observée sur le spectrogramme est, ici, de l'ordre de quelques dixièmes de micromètre pour l'axe $\overrightarrow{z}$ et d'environ 3 µm pour les axes $\overrightarrow{x}$ et $\overrightarrow{y}$.

Le capteur de force inclus dans la tête d'impédance fournit un signal de référence, une fonction de réponse en fréquence déplacement sur force, qui peut être calculée avec l'estimateur H1 (*cf.* **section 1.3.1**) à partir des signaux obtenus (*cf.* **figure 6.21**) avec le montage pseudo stéréoscopique (en *noir*) et avec l'accéléromètre (en *rouge*). Les courbes résultantes présentent une bonne corrélation visuelle jusqu'à 1000 Hz environ, le coefficient de corrélation étant de 99.8%. La comparaison fine des deux courbes donne plusieurs informations :

— vers 20 Hz et 75 Hz, soit pour des valeurs déjà identifiées sur la **figure 6.17** comme correspondant à la vibration de la caméra, des écarts de l'ordre de 40% sont visibles ;

3. L'accéléromètre fournit un signal d'accélération, le spectre est alors divisé par $-(2\pi f)^2$ afin de calculer le spectre de déplacement, comme expliqué au **chapitre 1**.

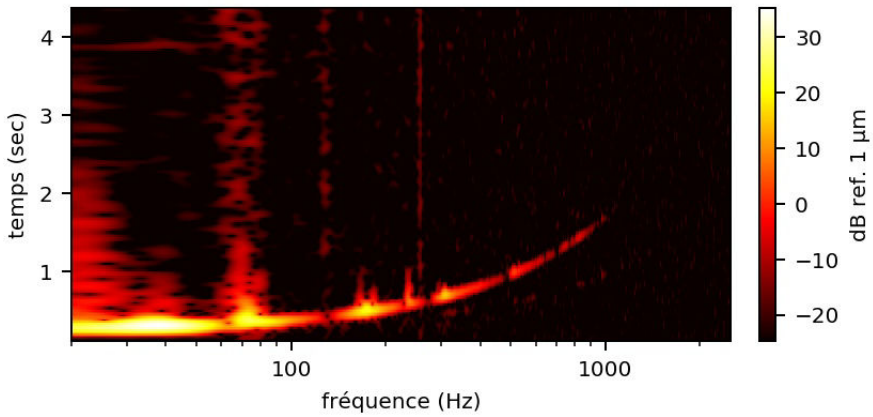


FIGURE 6.17 – Spectrogramme calculé pour le signal mesuré selon l'axe $\vec{z}$ avec le montage pseudo stéréoscopique à l'emplacement de l'excitation

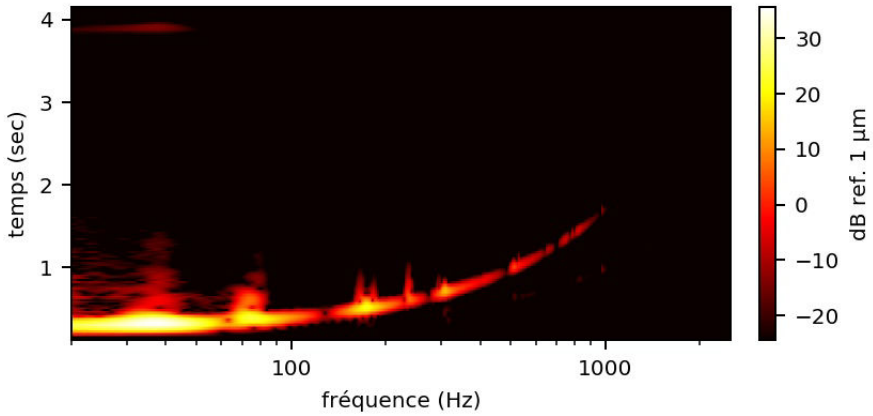


FIGURE 6.18 – Spectrogramme calculé pour le signal mesuré avec l'accéléromètre à l'emplacement de l'excitation

— jusqu'à 300 Hz, la différence est globalement inférieure à 10%, en omettant les fréquences d'anti-résonances[4] et l'impact de la vibration de la caméra ;
— plus la fréquence augmente, plus le rapport signal à bruit augmente pour le signal obtenu avec le montage pseudo stéréoscopique.

Finalement, la comparaison avec un accéléromètre, présentée sur la **figure 6.21**, valide l'amplitude des résultats obtenus avec le montage pseudo stéréoscopique pour un point, mais ne peut donner d'information sur une mesure plein champ. Pour ce faire, des mesures ont été réalisées avec un vibromètre laser à balayage. Le signal d'excitation est identique, mais répété une à plusieurs fois par point de mesure, car le vibromètre effectue une moyenne sur plusieurs mesures si le rapport signal à bruit est insuffisant. Rappelons ici qu'il y a 961 points de mesure, ce qui correspond à environ trois heures de mesures successives. Notons aussi que pour des raisons logistiques, les mesures avec le vibromètre ont été réalisées dans une salle différente, avec un écart de température d'environ dix degrés Celsius par rapport à la salle de l'expérience de vision. Cet écart engendre potentiellement des décalages dans la réponse de la plaque [175, 9].

Des déformées modales opérationnelles mesurées avec le montage pseudo stéréoscopique et le vi-

4. Il s'agit de la fréquence pour laquelle le point de mesure est situé sur une ligne nodale.

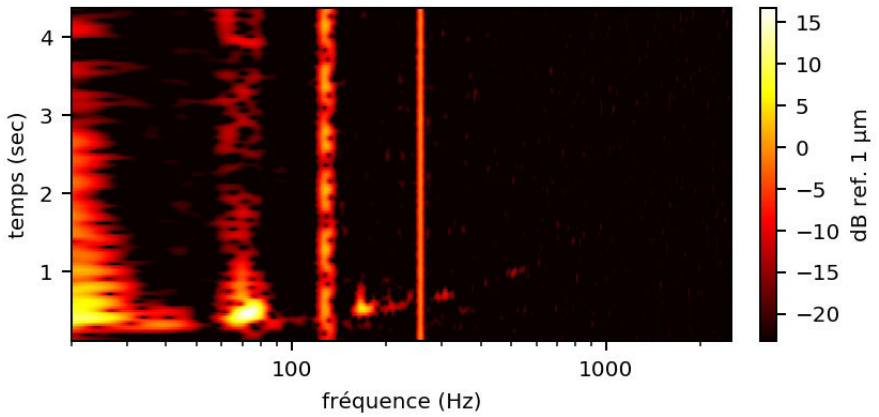


FIGURE 6.19 – Spectrogramme calculé pour le signal mesuré selon l'axe $\overrightarrow{x}$ avec le montage pseudo stéréoscopique à l'emplacement de l'excitation

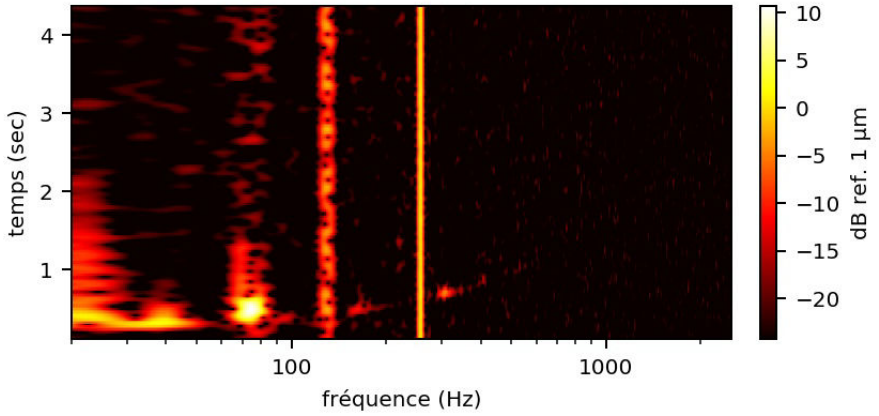


FIGURE 6.20 – Spectrogramme calculé pour le signal mesuré selon l'axe $\overrightarrow{y}$ avec le montage pseudo stéréoscopique à l'emplacement de l'excitation

bromètre laser sont présentées sur la **figure 6.22**. Afin de calculer les cartes présentées, les valeurs des fonctions de réponse en fréquence sont divisées par la valeur complexe du point ayant l'amplitude de vibration la plus forte à la fréquence étudiée. La partie réelle correspond ainsi à un cliché de la déformée modale animée (*cf.* **section 1.3.2**). Les fréquences de comparaison affichées sont représentatives des résultats et se situent à 76 Hz, 317 Hz, 519 Hz et 976 Hz avec des coefficients de corrélation associés de 97,7%, 99,1%, 96,5% et 98,4%. Ces hautes valeurs de corrélation valident la méthode, c'est-à-dire l'utilisation de cibles avec un montage pseudo stéréoscopique avec un jeu de quatre miroirs pour la mesure de vibrations de structures planes. Notons aussi que la non-planéité des miroirs, induisant des distances épipolaires de $\approx$ 1,2 pixel pour le calibrage, n'a pas d'impact clairement visible, dans ce cas, sur la mesure de déformées modales opérationnelles.

6.4.4 En bref...

Cette section a présenté en détail le montage pseudo stéréoscopique à caméra unique et jeu de quatre miroirs réalisé ainsi que les résultats obtenus. Les données de calibrage valident la démarche de dimensionnement, puisque la base obtenue estimée et la base calculée avec les équations de dimensionnement, à partir de la focale et des angles estimés par le calibrage, concordent à 0,3% près. L'écart-type du bruit de mesure dans le domaine temporel, estimé avec une mesure sans excitation,

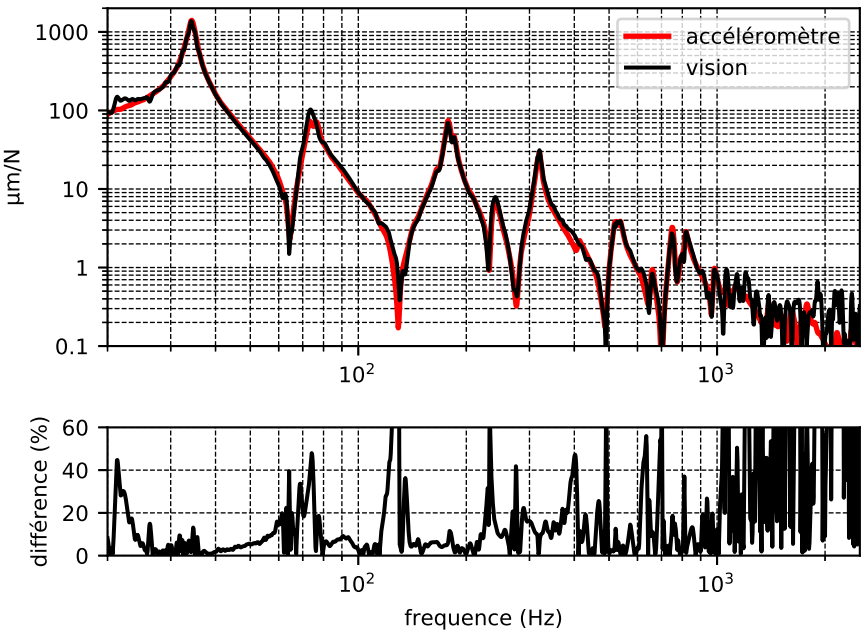


FIGURE 6.21 – Fonctions de réponse en fréquence mesurées avec le montage pseudo stéréoscopique et l'accéléromètre à l'emplacement de l'excitation

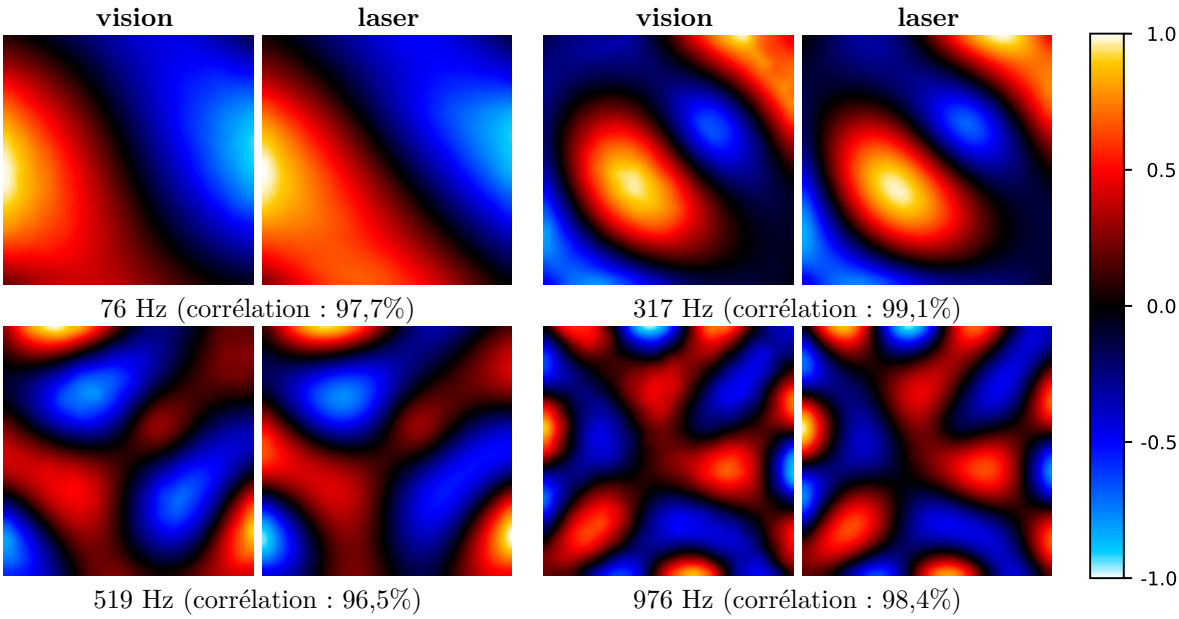


FIGURE 6.22 – Déformées modales opérationnelles de la plaque en aluminium mesurées avec le montage pseudo stéréoscopique et un vibromètre à balayage, à 76 Hz, 317 Hz, 519 Hz et 976 Hz

est inférieur à 3 µm (*cf.* **figure 6.16**).

La comparaison des résultats obtenus avec le montage pseudo stéréoscopique et un accéléromètre montre une bonne concordance dans le domaine fréquentiel (99,8%) jusqu'à 1000 Hz. Les déformées modales opérationnelles mesurées avec le montage et un vibromètre laser à balayage montrent elles

aussi une bonne corrélation, supérieure à 96% pour les fréquences étudiées.

Les résultats présentés attestent que la méthode de pseudo stéréo-vision permet d'obtenir des données vibratoires, à partir de plusieurs centaines de points, pour une surface quasi plane, tel qu'une caisse de guitare, des capots ou des portières de voitures, des ailes d'avions, etc. De plus, la mesure étant synchrone, *i.e.* effectuée simultanément pour tous les points de mesure, les phénomènes étudiés peuvent être non stationnaires et non répétables. Notons aussi que la mesure fournit des données dans les trois axes de l'espace. Ceci peut être intéressant pour certains systèmes complexes, par exemple des amortisseurs dynamiques chaotiques [162, 51, 79, 33, 11].

6.5 Application à un haut-parleur

Suite à la validation du montage appliqué à la mesure de vibrations d'une plaque, un deuxième test est mené sur un objet à géométrie plus complexe dans le cadre de la mesure de vibrations : un haut-parleur. La surface n'étant pas plane, il n'est pas possible de choisir une configuration angulaire aussi grande que pour la plaque. La littérature présente des exemples de mesure de déformées modales avec des techniques de mesure 1D [25, 21, 83]. La méthode associée au montage pseudo stéréoscopique présentée dans ce chapitre donne, quant à elle, des résultats 3D, rapportés dans cette section.

6.5.1 Description du montage

Adaptation du montage aux spécificités du haut-parleur

Le haut-parleur choisi est un woofer AB Sound TW 1041 avec un cône de 190 mm de diamètre externe, 100 mm de diamètre interne et une profondeur de 30 mm environ. La surface du cône est donc inclinée d'environ 30° par rapport à l'axe de déplacement principal de la membrane, ici défini comme l'axe $\vec{z}$. L'angle entre les caméras virtuelles et l'axe principal de déplacement est alors réglé à environ 30°, ce qui amène un angle maximal de 60° environ entre chaque caméra virtuelle et la surface de la membrane.

Le saladier du haut-parleur est fixée au support utilisé pour la mesure de la plaque, toujours posé sur la petite table optique et les plots anti-vibrations. Une photographie de l'ensemble est présentée sur la **figure 6.23**. La distance entre la caméra réelle et le plan situé au niveau de l'extérieur du cône du haut-parleur est de 770 mm environ.

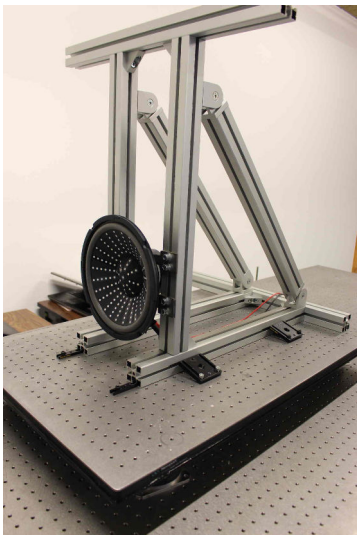


FIGURE 6.23 – Photographie du haut-parleur fixé à la structure positionnée sur la petite table optique.

Agencement des cibles

Une image du haut-parleur avec les cibles, obtenue avec le montage pseudo stéréoscopique réalisé, est affichée sur la **figure 6.24**. Vingt-quatre bandes de six cibles sont collées sur le cone et vingt-cinq cibles sur le dôme. Au total, 169 cibles sont collées sur la membrane. Les cibles ont été imprimées avec une imprimante laser sur un papier adhésif plastique standard, qui induit visiblement des réflexions spéculaires sur l'image finale, perturbant la détection de certaines cibles. On observe que deux points du dôme ne sont pas exploitables. Notons aussi que, avec environ 1 cm^2 par cible, et 80 g/m^2 pour le

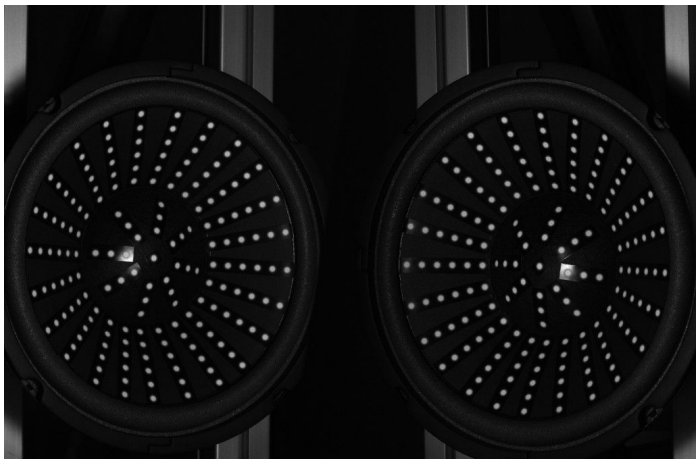


FIGURE 6.24 – Image du haut-parleur avec les cibles obtenues avec le montage pseudo stéréoscopique

papier, en négligeant la colle, la masse ajoutée est de 1,4 g environ. La masse mobile du haut-parleur, mesurée avec la méthode de Thiele & Small [157], étant de $\approx$ 28 g, l'alourdissement amené par le papier est donc de 5% environ et pourrait ainsi s'avérer à terme non-négligeable.

Appariement initial

Afin de ne pas trier les cibles une à une manuellement, un algorithme est créé. Cet algorithme est utilisé pour chaque image virtuelle. Les étapes en sont :

1. 8 points aux bords du cône sont sélectionnés manuellement ;
2. une ellipse est calculée à partir des points sélectionnés ;
3. les 24 points les plus proches de l'ellipse sont sélectionnés ;
4. le tri angulaire des points sélectionnés est réalisé en exprimant leurs coordonnées dans un repère centré autour de l'ellipse calculée ;
5. ces points sont ensuite supprimés de la liste des points à trier ;
6. boucler à l'étape 2 cinq fois ;
7. boucler à l'étape 2 trois fois en cherchant les 8 points les plus proches ;
8. le point le plus proche du centre de l'ellipse finale est finalement sélectionné.

La **figure 6.25** présente le tri initial, résultant de l'algorithme créé, pour les deux images virtuelles. Les lignes *jaunes* rejoignent les points de mesures en *rouge*.

6.5.2 Données de calibrage

Comme pour l'expérience menée avec la plaque, la caméra est calibrée après les mesures de vibrations. La distance focale obtenue pour la caméra est de 56,4 mm. Les angles de rotation estimés autour des axes $\overrightarrow{y}$, $\overrightarrow{x}$ puis $\overrightarrow{z}$ sont de 65,24°, -0,06° et 0,02° respectivement. La distance calculée entre les caméras virtuelles est de 1124 mm.

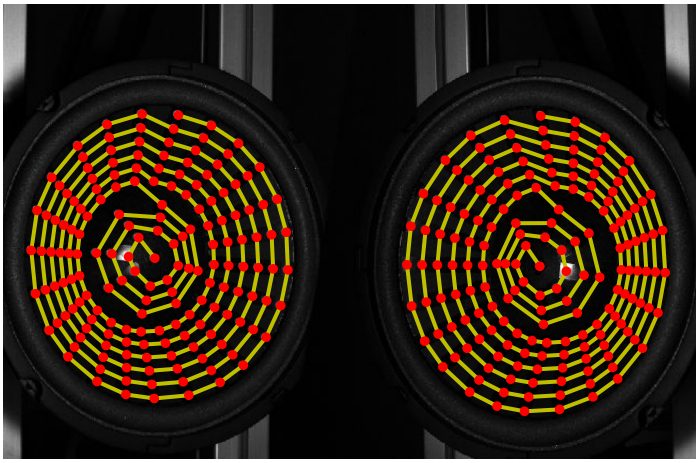


FIGURE 6.25 – Illustration du tri initial des cibles dans les deux images virtuelles

En regardant la **figure 6.24**, la hauteur totale du haut-parleur (256 mm) est imagée sur environ 590 pixels, et le bord extérieur de la suspension (à environ 110 mm du centre) arrive au bord de l'image. En prenant en compte la focale et un angle $2\delta = 32{,}5°$, la base B obtenue est de 1157 mm et correspond, à 3% près, à la distance estimée par le calibrage (1124 mm). Ceci valide encore une fois les équations de calibrage.

Au niveau de la mire, les distances épipolaires observées sont en moyenne de 0,2 pixel, avec un maximum de 0,93 pixel.

6.5.3 Mesure de vibrations - évaluation de la précision

Spectrogrammes

Le signal d'excitation est un sinus glissant de 1 V d'amplitude allant de 20 Hz à 430 Hz en quatre secondes. À 400 Hz, le déplacement global de la membrane est donc de l'ordre de 5 µm. La cadence d'acquisition est alors réglée à 5 000 images par seconde pour avoir au moins 10 points par période pour la fréquence la plus élevée. Les spectrogrammes pour les axes $\overrightarrow{x}$, $\overrightarrow{y}$ et $\overrightarrow{z}$, calculés à partir du signal correspondant à la cible en *rouge* sur la **figure 6.26**, sont affichés sur les **figures 6.27**, **6.28** et **6.29**. Cet emplacement a été choisi car les amplitudes des déformées modales sont plus importantes en périphérie du cône et donc davantage visibles sur les spectrogrammes.

FIGURE 6.26 – Image affichant l'emplacement du point correspondant au signal étudié (en *rouge*)

Le bruit observé pour la plaque vers 20 Hz est encore une fois présent sur les signaux des trois axes de l'espace. Cependant, la ligne verticale vers 76 Hz n'est pas visible et les lignes verticales vers 130 Hz et

270 Hz ne sont visibles que pour les axes $\overrightarrow{x}$ et $\overrightarrow{y}$. De plus, vers 70 Hz, 85 Hz et 95 Hz, on observe une rémanence de l'énergie alors que la fréquence instantanée du signal d'excitation a dépassé ce stade. Ces traînées correspondent fort probablement à des *rocking modes*, c'est-à-dire à des modes de rotation de la membrane (cette hypothèse est validée ci-après, en traçant les déformées modales opérationnelles).

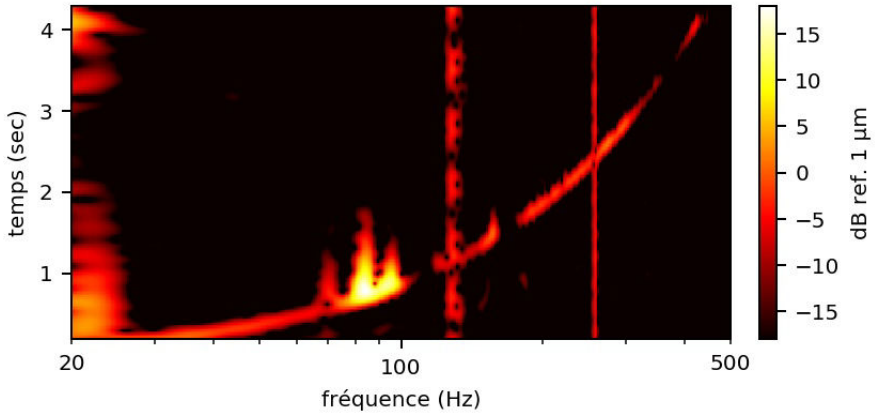


FIGURE 6.27 – Spectrogramme calculé pour le point choisi selon l'axe $\overrightarrow{x}$

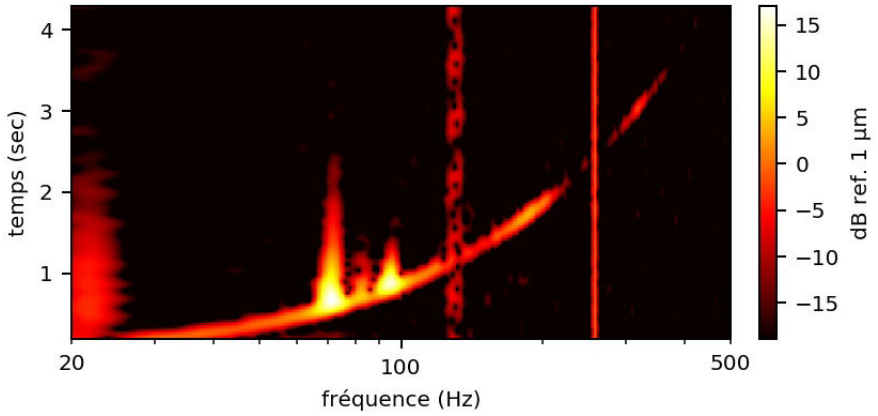


FIGURE 6.28 – Spectrogramme calculé pour le point choisi selon l'axe $\overrightarrow{y}$

Le repère de triangulation utilisé est le repère défini par la mire, pour le cliché où la mire est alignée avec le plan moyen du haut-parleur. Sur les spectrogrammes observés, vers 40 Hz, les amplitudes pour les axes $\overrightarrow{x}$, $\overrightarrow{y}$ et $\overrightarrow{z}$ sont égales à $\approx$ 0,8 µm, $\approx$ 1,1 µm et $\approx$ 98 µm respectivement. En supposant qu'à cette fréquence le mouvement du déplacement est selon l'axe principal de vibration du haut-parleur, l'angle entre l'axe $\overrightarrow{z}$ et l'axe principal est de 0,8° environ et le décalage correspondant $(1-cos(0,81°))$ est d'environ 0,1‰. Il n'est alors pas nécessaire de passer par le repère du plan moyen de l'objet pour obtenir des valeurs valides.

Notons que ces amplitudes ne correspondent pas à l'amplitude maximale réelle mesurée ($\approx$ 400 µm, *cf.* **figure 6.30**) puisque la fréquence instantanée du signal n'est pas constante.

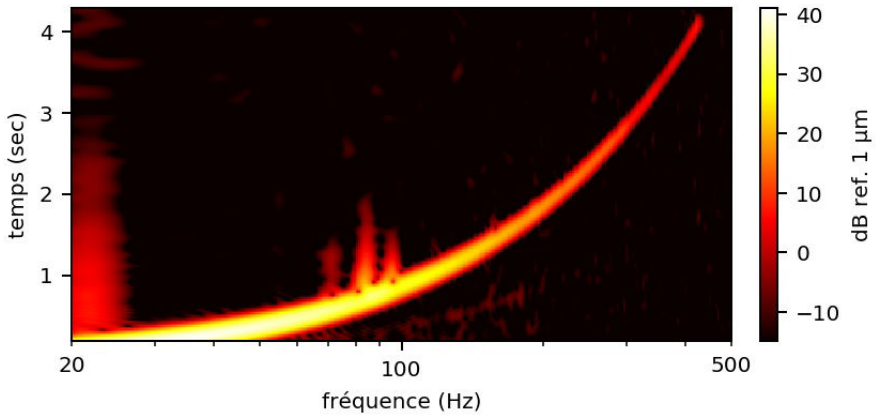


FIGURE 6.29 – Spectrogramme calculé pour le point choisi selon l'axe $\vec{z}$

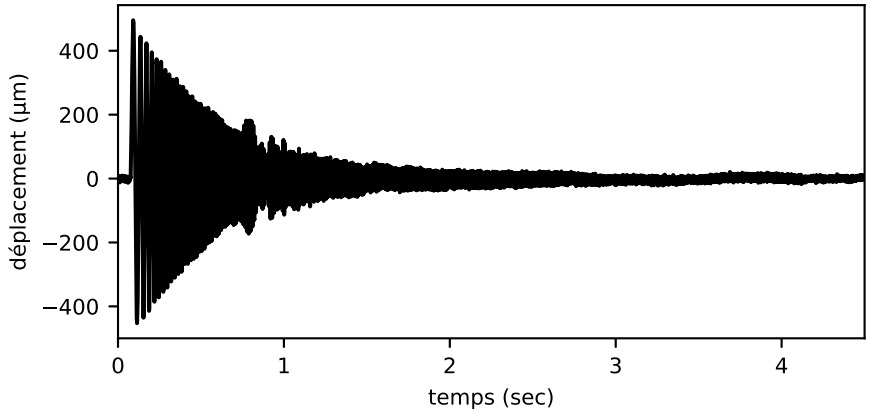


FIGURE 6.30 – Signal de déplacement mesuré pour un point selon l'axe $\vec{z}$

Fonction de réponse en fréquence

La fonction de réponse en fréquence a été calculée à partir des données obtenues avec le montage pseudo stéréoscopique et un vibromètre laser, afin d'effectuer une comparaison. Le point de mesure choisi est le point au centre du dôme puisque le déplacement associé est supposé colinéaire avec l'axe de déplacement de la bobine du haut-parleur. Les courbes obtenues avec les données du vibromètre et du montage pseudo stéréoscopique sont respectivement tracées en *rouge* et *noir* sur la **figure 6.31**. En omettant l'impact de la composante basse fréquence visible sur les spectrogrammes ci-dessus, et les écarts au niveau des creux entre 70 Hz et 100 Hz dus aux *rocking modes*, l'écart est inférieur à 2% jusqu'à 250 Hz environ. Le niveau de bruit augmente visiblement après cette fréquence pour la méthode de pseudo stéréo-vision.

Déformées modales opérationnelles

L'écart-type des auto-spectres des signaux mesurés permettent de déterminer les fréquences correspondant à des déformées modales opérationnelles. Une courbe est calculée pour chaque axe de l'espace selon la formule :

$$\sigma_{S_{xx}}(f) = \sqrt{\frac{\sum_i^N \overline{S_{xx}^i}(f)}{N-1}} \tag{6.15}$$

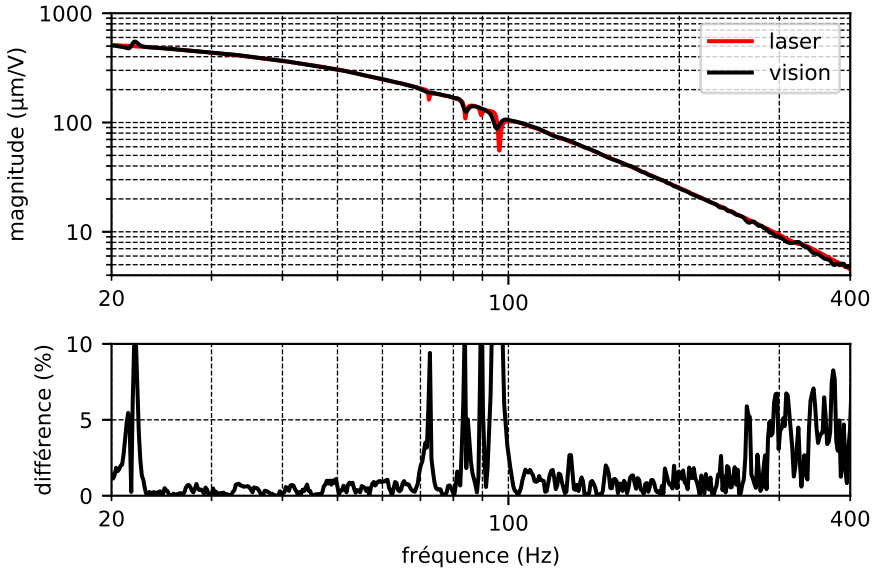


FIGURE 6.31 – Fonction de réponse en fréquence mesurée au centre du haut-parleur avec le montage pseudo stéréoscopique et le vibromètre laser

avec $\overline{S^i_{xx}}(f)$ l'auto-spectre du signal du point i, auquel le signal moyen a été soustrait.

Les écarts-types des auto-spectres des signaux mesurés, en fonction de la fréquence, sont tracés sur la **figure 6.32**, en *rouge* pour l'axe $\overrightarrow{x}$, en *vert* pour l'axe $\overrightarrow{y}$ et en *noir* pour l'axe $\overrightarrow{z}$. Les pics correspondent à des déformées modales opérationnelles (*cf.* **section 1.3.2**) et sont clairement visibles sur les trois courbes pour la plupart des fréquences de résonance observables ici. Le bruit de chaque courbe donne aussi une information sur le bruit de fond dans le domaine fréquentiel pour chaque axe de l'espace, à savoir un écart-type de ≈ 0,2 µm pour l'axe $\overrightarrow{z}$ et ≈ 0,1 µm pour les axes $\overrightarrow{x}$ et $\overrightarrow{y}$.

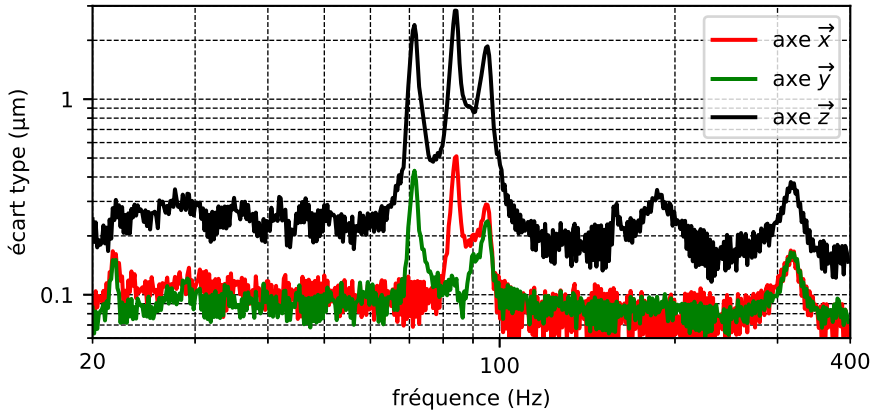


FIGURE 6.32 – Courbes des écarts-types des auto-spectres des signaux mesurés en fonction de la fréquence

Notons que la valeur absolue de la transformée de Fourier du signal d'excitation (affichée sur la **figure 6.33**) présente une valeur moyenne de ≈ 47 mV entre 30 Hz et 400 Hz. L'écart-type de l'erreur attendue pour les fonctions de réponse en fréquence, utilisées pour calculer les déformées modales opérationnelles, est alors de l'ordre de 2 µm/V pour les axes $\overrightarrow{x}$ et $\overrightarrow{y}$ et de 4 µm/V pour l'axe $\overrightarrow{z}$.

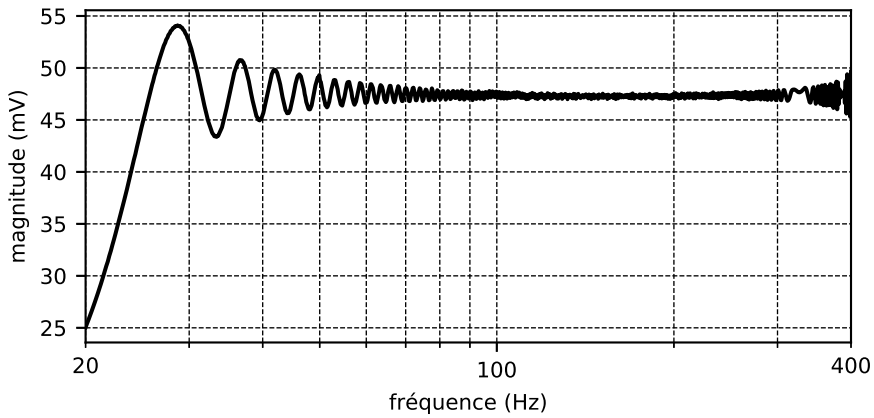


FIGURE 6.33 – Valeur absolue de la transformée de Fourier du signal d'excitation

Les fonctions de réponse en fréquence sont ensuite calculées pour les trois axes de l'espace afin d'afficher les déformées modales opérationnelles, aux fréquences des pics des courbes des écarts-types des auto-spectres (*cf.* **figure 6.32**), à savoir 72 Hz, 84 Hz, 95 Hz, 158 Hz, 187 Hz et 318 Hz. De plus, pour une lecture plus simple des cartes obtenues, le signal de déplacement du point central est soustrait au signal de chaque point de mesure avant le calcul de la fonction de réponse en fréquence correspondante. Les cartes résultantes sont affichées sur les **figures 6.34** à **6.39** au format cliché de l'animation du mode, en prenant comme référence angulaire le point présentant la valeur absolue la plus haute selon l'axe $\overrightarrow{z}$.

Les déformées calculées à 72 Hz, 84 Hz et 95 Hz correspondent à des *rocking modes*, avec une ligne nodale respectivement horizontale, verticale et oblique. Comme on peut l'observer sur les courbes sur la **figure 6.32**, les déformées mesurées selon l'axe $\overrightarrow{x}$ à 72 Hz et selon l'axe $\overrightarrow{y}$ à 84 Hz ne correspondent pas à des pics mais à du bruit, avec des écarts-types égaux à $\approx 0{,}5$ µm/V et $\approx 0{,}7$ µm/V respectivement, soit des valeurs légèrement inférieures à la valeur moyenne estimée ci-dessus (≈ 2 µm/V).

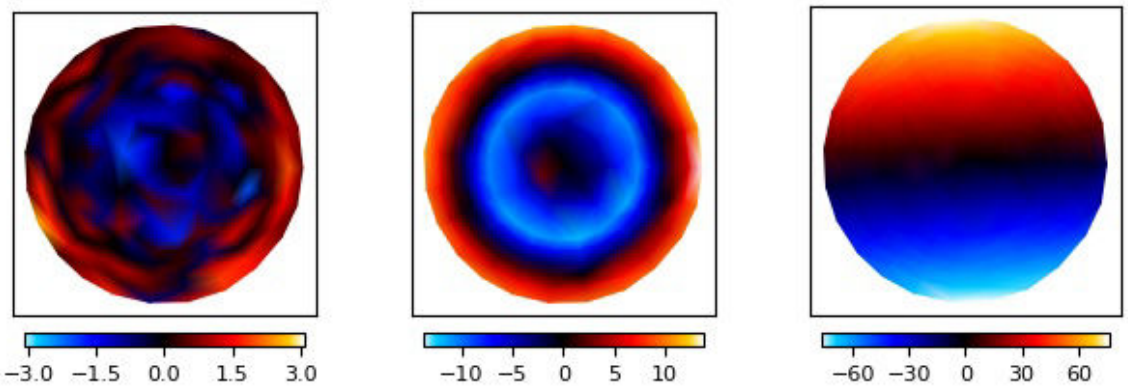


FIGURE 6.34 – Cartes de la déformée opérationnelle mesurée du haut-parleur à 72 Hz ; (*gauche*) axe $\overrightarrow{x}$, (*milieu*) axe $\overrightarrow{y}$, (*droite*) axe $\overrightarrow{z}$; les valeurs sont exprimées en µm/V

Pour les déformées à 158 Hz et 187 Hz, une allure de *rocking mode* est visible, respectivement sur les axes $\overrightarrow{y}$ et $\overrightarrow{z}$ dans le premier cas, et $\overrightarrow{x}$ et $\overrightarrow{z}$ dans le second cas, possiblement mélangé avec un mode de membrane. Les valeurs des cartes pour les axes $\overrightarrow{x}$ (à 158 Hz) et $\overrightarrow{y}$ (à 187 Hz) sont visuellement très proches du bruit de fond, et les potentiels modes de membranes ne sont pas clairement identifiables.

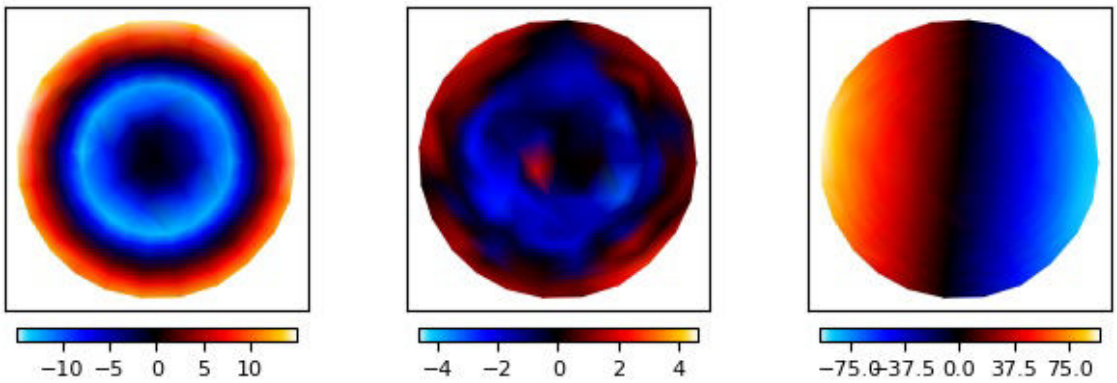


FIGURE 6.35 – Cartes de la déformée opérationnelle mesurée du haut-parleur à 84 Hz ; (*gauche*) axe $\overrightarrow{x}$, (*milieu*) axe $\overrightarrow{y}$, (*droite*) axe $\overrightarrow{z}$; les valeurs sont exprimées en µm/V

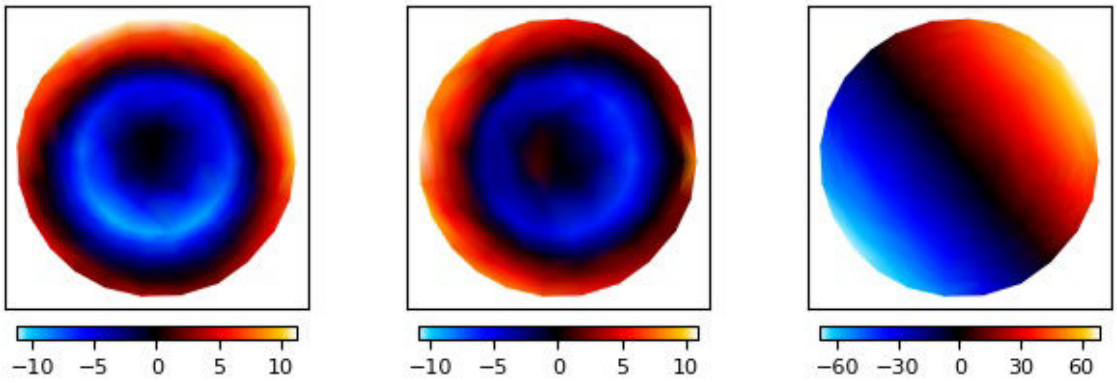


FIGURE 6.36 – Cartes de la déformée opérationnelle mesurée du haut-parleur à 95 Hz ; (*gauche*) axe $\overrightarrow{x}$, (*milieu*) axe $\overrightarrow{y}$, (*droite*) axe $\overrightarrow{z}$; les valeurs sont exprimées en µm/V

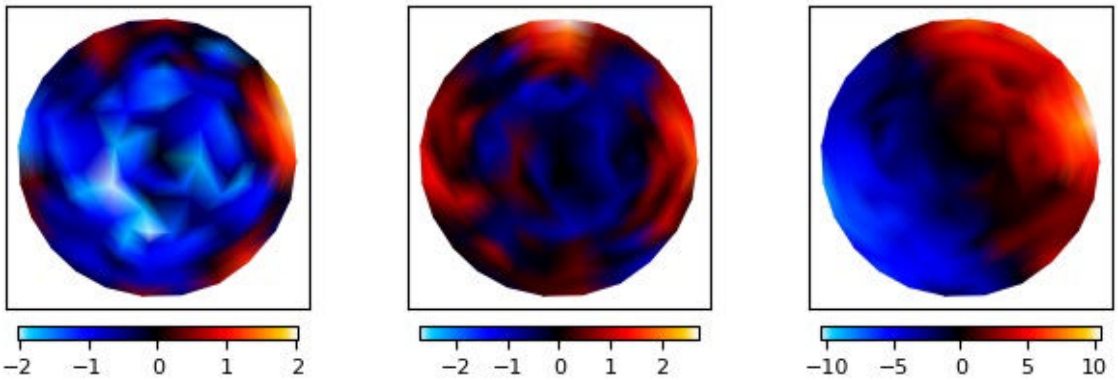


FIGURE 6.37 – Cartes de la déformée opérationnelle mesurée du haut-parleur à 158 Hz ; (*gauche*) axe $\overrightarrow{x}$, (*milieu*) axe $\overrightarrow{y}$, (*droite*) axe $\overrightarrow{z}$; les valeurs sont exprimées en µm/V

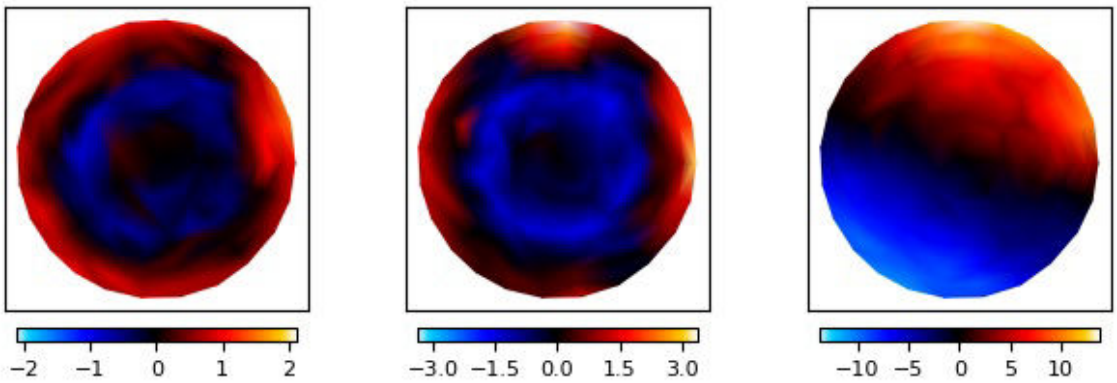


FIGURE 6.38 – Cartes de la déformée opérationnelle mesurée du haut-parleur à 187 Hz ; (*gauche*) axe $\overrightarrow{x}$, (*milieu*) axe $\overrightarrow{y}$, (*droite*) axe $\overrightarrow{z}$; les valeurs sont exprimées en µm/V

Finalement, le mode observé à 318 Hz correspond à un mode de membrane à géométrie relativement complexe. Le dôme ne se déforme pas, mais le cône présente un mode radial avec cinq longueurs d'ondes. Contrairement aux *rocking modes*, les amplitudes selon les trois axes de l'espace sont relativement proches, entre ≈ 9 µm et ≈ 17 µm, l'axe des vibrations semble lié à la normale locale de la surface. Néanmoins, avec la disposition des cibles, les cinq longueurs d'ondes radiales sont mesurées avec vingt quatre points, ce qui correspond à 4,8 points par longueur d'onde. Cette valeur est relativement faible pour interpoler correctement les cartes de déformées : la valeur recommandée pour les simulations numériques en acoustique est de six à dix nœuds par longueur d'onde [155]. On retrouve ainsi les déformées, mais avec des erreurs de reconstruction.

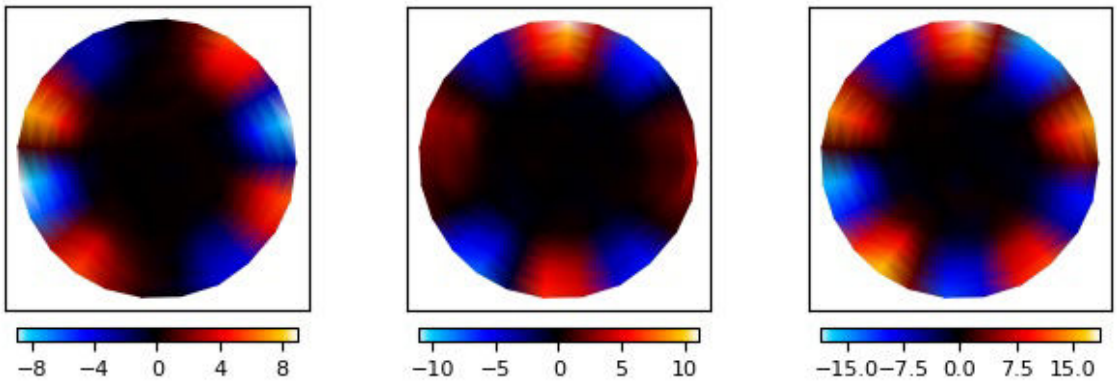


FIGURE 6.39 – Cartes de la déformée opérationnelle mesurée du haut-parleur à 318 Hz ; (*gauche*) axe $\overrightarrow{x}$, (*milieu*) axe $\overrightarrow{y}$, (*droite*) axe $\overrightarrow{z}$; les valeurs sont exprimées en µm/V

6.5.4 Bilan

Cette section a présenté les résultats de mesure obtenus pour un haut-parleur. La géométrie de cet objet d'étude induit deux contraintes :

— l'angle entre chaque caméra virtuelle et l'axe principal de vibration est limité à ≈ 30-35°, afin que l'angle maximal entre le repère caméra et la surface mesurée n'excède pas 60° environ ;

— les cibles doivent être collées selon un schéma permettant un tri semi-automatique.

Ainsi, pour cette expérimentation, l'angle entre les caméras virtuelles est réglé à $\approx 60°$ et est estimé à 65,24° par le calibrage. Un algorithme a été spécifiquement créé pour trier des cibles réparties en bandes collées à différents angles par rapport au centre de la membrane.

Les spectrogrammes calculés à partir des signaux mesurés, dans les trois axes de l'espaces (*cf.* **figures 6.27** à **6.29**), pour une cible sur le bord de la membrane, permettent de mettre en évidence le bruit induit par les ventilateurs de la caméra. De plus, les *rocking modes* entre 70 Hz et 100 Hz sont clairement visibles car ils induisent une rémanence de l'énergie aux fréquences correspondantes.

La comparaison de la réponse en fréquence mesurée au centre du haut-parleur avec la méthode de pseudo stéréo-vision et un vibromètre laser (*cf.* **figure 6.31**) donne une différence entre les deux courbes inférieure à 2% jusqu'à 250 Hz et inférieure à 5% jusqu'à 400 Hz.

Les courbes des écarts-types des auto-spectres en fonction de la fréquence (*cf.* **figure 6.32**), calculés à partir des signaux mesurés dans les trois axes de l'espace, permettent de localiser les déformées modales mesurables avec la méthode et indiquent que le bruit de la mesure dans le domaine fréquentiel est de l'ordre de 0,2 µm.

Les déformées modales localisées ont donc été tracées (*cf.* **figures 6.34** à **6.39**) et correspondent à des *rocking modes* et à un mode de membrane. Le mode de membrane mesuré est radial pour le cône et montre une limitation de l'agencement choisi pour les cibles, à savoir un échantillonnage spatial trop faible. Notons qu'un agencement en spirale permettrait d'augmenter le nombre de points de mesure, mais sa réalisation semble *a priori* délicate. Rappelons aussi qu'avec 169 cibles, la masse mobile du haut-parleur est alourdie d'environ 5%. L'idéal serait donc, avec la méthode de suivi de cibles, de dessiner les cibles sur la membrane ou de les projeter.

6.6 Conclusion

En conclusion, ce chapitre a présenté le principe et le dimensionnement du montage pseudo stéréoscopique à caméra unique et jeu de quatre miroirs. Différents protocoles de calibrage, ainsi que les avantages et inconvénients associés, ont été présentés. Le choix s'est finalement porté sur un calibrage de caméra en déplaçant les miroirs, une fois le montage (pré-)réglé, soit juste avant le calibrage du système pseudo stéréoscopique, soit juste après les mesures de vibrations.

Une étude de sensibilité de la précision de la mesure à certains paramètres a été menée à l'aide de simulations réalisées avec le logiciel Blender, aboutissant :

— premièrement, au choix d'un motif pour les cibles utilisées pour la mesure de vibrations, mais aussi de deux détecteurs associés ;
— deuxièmement, à la sélection d'un agencement optimisant la mesure de vibrations, grâce à la simulation d'images d'une plaque en vibration vue à travers le système pseudo stéréoscopique pour différentes configurations angulaires.

Finalement, cette étude de sensibilité a permis de démontrer que, pour la mesure de vibrations avec une méthode de suivi de cibles, il est préférable d'utiliser un motif point-Gaussien, avec l'estimateur d'ellipses duales ainsi qu'un angle entre les caméras virtuelles et la surface à mesurer de l'ordre de 50°. Notons aussi que les résultats de l'étude attestent que, dans la configuration sélectionnée, la mesure de petits déplacements est plus précise que la mesure de forme, et que la mesure de vibrations l'est encore davantage en passant dans le domaine fréquentiel.

Un montage pseudo stéréoscopique a finalement été réalisé afin de comparer les résultats de mesure, dans une configuration tenant compte de l'étude de sensibilité, à ceux obtenus avec un accéléromètre et un vibromètre laser à balayage. Pour une plaque de 330 mm × 330 mm encastrée aux quatre coins, l'écart-type de l'erreur de la mesure statique, dans le domaine temporel, est inférieur à 3 µm. La mesure avec une excitation de type sinus glissant a mis à jour l'impact de la vibrations des ventilateurs de la caméra sur la mesure de vibrations à certaines fréquences. Néanmoins, la corrélation entre les réponses en fréquence mesurées avec le montage pseudo stéréoscopique et l'accéléromètre est de 98,8%

entre 20 Hz et 1000 Hz. Les résultats ont aussi été comparés avec ceux obtenus avec un vibromètre laser à balayage et affichent des coefficients de corrélation supérieurs à 96% pour les modes étudiés. La limite fréquentielle de la méthode, 1000 Hz pour cette expérimentation, est plus faible que pour un vibromètre, mais offre une mesure 3D de phénomènes vibratoires sans limitation quant à la non stationnarité et la non répétabilité des phénomènes étudiés, de par la mesure synchrone de plusieurs centaines de points de mesure.

Le montage a ensuite été appliqué à la mesure de vibrations d'un objet à géométrie plus complexe, un haut-parleur, dont les caractéristiques requièrent une adaptation des paramètres angulaires et de l'agencement des cibles. La fonction de réponse en fréquence du point central mesurée avec le montage pseudo stéréoscopique a été comparée avec celle obtenue avec un vibromètre laser et montre un écart inférieur à 2% jusqu'à 250 Hz et 5% jusqu'à 400 Hz. L'analyse temps-fréquence des signaux mesurés pour un point dans les trois axes de l'espace a mis en évidence l'impact des vibrations de la caméra sur la mesure et la rémanence de l'énergie des *rocking modes* du haut-parleur entre 70 Hz et 100 Hz. Les courbes des écarts-types des auto-spectres en fonction de la fréquence, permettent, pour cet objet d'étude, d'obtenir une erreur de mesure pour les trois axes de l'espace dans le domaine fréquentiel, approximativement égale à 0,1 µm ou 0,2 µm selon les axes. Ces courbes permettent aussi de détecter les déformées modales opérationnelles observables qui ont ensuite été affichées et décrites. Finalement, une amélioration de l'agencement des cibles est requise pour correctement décrire les déformées modales complexes du haut-parleur, avec l'utilisation préférentielle de cibles dessinées afin de limiter l'alourdissement de la structure légère qu'est sa membrane. Ces réflexions ont abouti au choix d'un motif aléatoire pour les phases de mesure ultérieures, et par conséquent à l'utilisation de la corrélation d'images. S'est alors posé le problème de l'appariement des images initiales, avec un objet d'étude montrant une forte profondeur et le choix d'un angle panoramique de plus de 60°. Le chapitre suivant présente la solution trouvée : la conception d'une méthode originale de rectification d'images.

Chapitre 7

Mesure par corrélation d'images : méthode 3D Itérative de Rectification d'images (IRIs)

« *Take a method and try it. If it fails, admit it frankly, and try another. But by all means, try something.* »
Franklin D. Roosevelt

Sommaire

7.1 Introduction **119**
7.2 Méthode IRIs **120**
7.2.1 Préambule 120
7.2.2 Principe général 122
7.3 Validation expérimentale **125**
7.4 Validation numérique de la méthode **129**
7.5 Non-planéité des miroirs **130**
7.6 Automatisation de l'appariement initial de la méthode IRIs **131**
7.7 Conclusion **136**

7.1 Introduction

Comme exposé au **chapitre 6**, les mesures réalisées avec le montage pseudo stéréoscopique ont donné des informations sur le comportement vibratoire d'un haut-parleur. Cependant, les cibles collées alourdissent la membrane très légère (la masse mobile avoisine les 28 g), et l'agencement testé n'est pas optimal pour décrire correctement les déformations. Une autre configuration permettrait possiblement de palier le manque d'information spatiale constaté lors de ce premier test. De même, l'utilisation d'un crayon blanc afin de dessiner les cibles à la place de les coller, pourrait résoudre le problème d'alourdissement. Néanmoins, dessiner des cibles avec un motif point-Gaussien a paru peu évident, sachant que la répartition de ces disques à bords adoucis doit être très précise. En conséquence, le choix s'est porté sur un passage des cibles à un motif aléatoire et donc à la corrélation d'images.

Rappelons que l'objet d'étude est un objet non-plan et que l'angle choisi entre les points de vue est supérieur à 60°. Dans ce contexte, les méthodes habituelles d'appariement initial, *i.e.* de rectification

d'image avec la géométrie épipolaire et de corrélation aux abords de la droite épipolaire (*cf.* **section 4.6.2**), ne permettent pas de réaliser un appariement automatique. En effet, pour la mesure de forme, un angle panoramique compris entre 10° et 30° est recommandé pour limiter les différences de perspective [151]. La conception d'une méthode originale, adaptée à ces besoins spécifiques, a alors semblé nécessaire. Dans sa thèse, Garcia [49] propose d'interpoler, dans les images brutes, la déformation des subsets initialement estimée dans des images rectifiées, avant d'effectuer une optimisation avec la méthode Newton-Raphson [23, 151]. Cette approche a servi de base de travail : une mesure de forme grossière a été interpolée afin de calculer des images corrigeant les distortions de perspective, ceci permettant de réaliser un appariement automatique d'un grand nombre de points de mesure, avec une répartition souhaitée dans le plan moyen de l'objet à mesurer. Une méthode 3D Itérative de Rectification d'Images (notée méthode IRIs) a ainsi été conçue, avec les codes associés [1].

Ce chapitre est organisé de la façon suivante : dans un premier temps, la méthode IRIs est décrite en détail (**section 7.2**). Est ensuite présentée la mesure de forme du haut-parleur étudié au **chapitre 6** avec le montage pseudo stéréoscopique à quatre miroirs et un motif aléatoire (**section 7.3**), afin d'obtenir une validation expérimentale de la méthode. La **section 7.4** correspond à l'étape de simulation réalisée pour vérifier les codes et valider numériquement la méthode (**section 7.5**). Cette phase de travail a permis de mettre à jour la non-planéité des miroirs utilisés. Finalement, la méthode a été améliorée afin d'aboutir à un appariement automatique. Ce protocole complémentaire est présenté en **section 7.6**.

7.2 Méthode IRIs

7.2.1 Préambule

Avant d'effectuer la mesure de vibrations, une mesure de forme doit être réalisée afin de définir et d'appairer des subsets dans les images initiales, qui, ici, correspondent aux sous-images obtenues avec le montage pseudo stéréoscopique avec un jeu de quatre miroirs, réglé pour observer le haut-parleur AB Sound étudié en **section 6.5**, avec un angle entre les deux caméras virtuelles de $\approx 75°$. L'image initiale est affichée sur la **figure 7.1**.

FIGURE 7.1 – Image initiale du haut-parleur avec mouchetis obtenue avec un montage pseudo stéréoscopique avec un jeu de quatre miroirs

Comme présenté en **section 4.6.2**, la rectification épipolaire permet d'effectuer un appariement de subsets avec les outils de corrélation d'images. Cependant, dans le cas d'un haut-parleur, et avec un angle panoramique important, les différences de perspective ne sont pas suffisamment corrigées. En illustration, les images rectifiées sont affichées sur la **figure 7.2** : les images des bords du dôme et du

1. Notons que des méthodes utilisent un modèle CAO de l'objet mesuré afin d'initialiser l'appariement des points de mesure.

cône sont bien circulaires, mais les subsets présentés sont corrélés à seulement 15,7%. Leur appariement manuel est possible, à deux ou trois pixels près. *A contrario*, l'appariement automatique génère des associations aberrantes, dues à la forte différence de perspective. Par exemple, sur la **figure 7.3**, la plus haute valeur de coefficient de corrélation, pour un point proche de la droite épipolaire, est de 28,6% pour un subset appairé situé sur le dôme, alors que le subset de référence était situé sur le cône (*cf.* **figure 7.2**, à *gauche*).

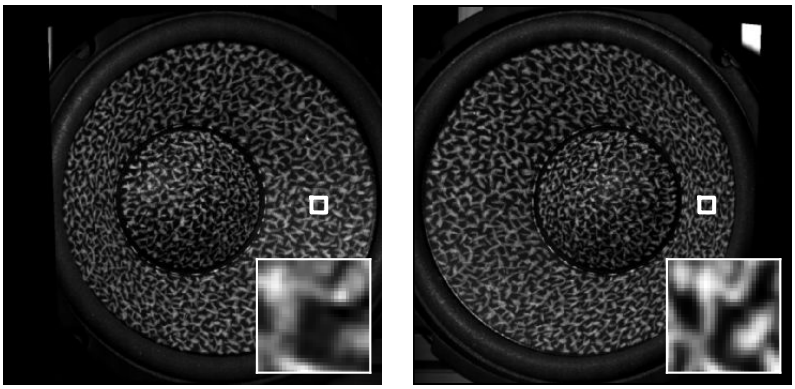


FIGURE 7.2 – Images rectifiées à l'aide de la géométrie épipolaire ; les subsets affichés (encadrés en bas à droite de chaque image) sont corrélés à 15,7%

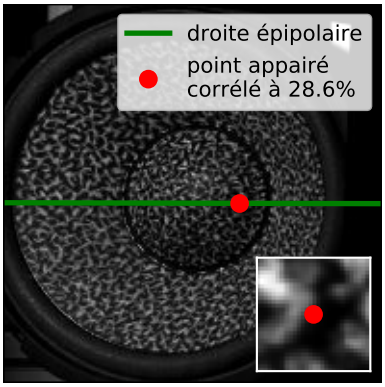


FIGURE 7.3 – Centre du subset appairé présentant la plus haute valeur de corrélation avec le subset de référence (affiché sur la **figure 7.2** à *gauche*)

Le problème étant identifié, plusieurs pistes ont été explorées afin de trouver la solution la plus adaptée à ce contexte.

Mentionnons premièrement l'existence de logiciels industriels de corrélation d'images, bien que cette option ait été rapidement rejetée. En effet, le haut-parleur a une géométrie très spécifique qui, couplée aux forts angles stéréoscopiques requis pour optimiser la précision de la mesure de vibrations, nécessite un processus d'appariement adapté qui n'est pas forcément intégré dans les algorithmes disponibles sur le marché. De plus, les licences de ces logiciels ont un coût certain, incompatible avec l'utilisation libre, à terme, de la méthode au sein des différentes structures finançant ce travail de recherche.

Deuxièmement, des méthodes du type Newton-Raphson sont souvent utilisées. Basées sur l'estimation des déformations de subsets, leur implémentation est complexe et chronophage, ce qui les a fait écarter des solutions envisageables à court terme.

Finalement, restait l'option consistant à concevoir une méthode adaptée au contexte spécifique de cette recherche, et n'engendrant *in fine* aucun surcoût. La méthode IRIs développée repose ainsi

sur l'estimation, non pas de la déformation de chaque subset, mais de la forme globale de la surface mesurée afin de corriger les distortions de perspective et ainsi d'effectuer un appariement automatique d'un grand nombre de points de mesure, répartis selon le schéma souhaité dans le repère lié au plan moyen de la surface mesurée. La méthode est relativement simple d'implémentation lorsque les outils basiques de corrélation d'images, de triangulation et d'interpolation d'images sont déjà à disposition.

7.2.2 Principe général

À l'instar de la stéréo-corrélation fine améliorée proposée par Garcia [49], la méthode utilise une mesure initiale interpolée. Cependant, ici, la forme de la surface mesurée est estimée à l'aide d'un calcul itératif d'ortho-images. Le diagramme synoptique détaillé de la méthode est affiché sur la **figure 7.4**. Les phases principales sont rappelées sur la gauche de l'image, et associées à différentes couleurs. Une *mesure de forme initale*, grossière (en *noir*) est réalisée à partir d'images rectifiées à l'aide de la géométrie épipolaire (*cf.* **section 4.6.2**). Cette forme sert de référence pour effectuer un *calcul d'ortho-images* (en *vert*), une par point de vue, corrigeant partiellement les distortions de perspective. La paire d'ortho-images permet ensuite d'appairer automatiquement des subsets par *corrélation d'images* (en *bleu*). Les positions des points de mesure appairés sont ensuite *projetées vers les images initiales* (en *rouge*), afin d'estimer, par triangulation, un nouvelle forme de référence (en *gris*). Le processus peut alors être réitéré avec la nouvelle forme de référence, estimée à chaque itération. Lorsque les calculs convergent, la projection de points vers les images initiales permet de calculer des subsets de suivi dans les images initiales (en *noir*), pour les séquences temporelles d'images. Les étapes détaillées figurant dans l'organigramme respectent ce code couleur afin de faciliter la lecture de la figure, et sont expliquées dans les paragraphes suivants.

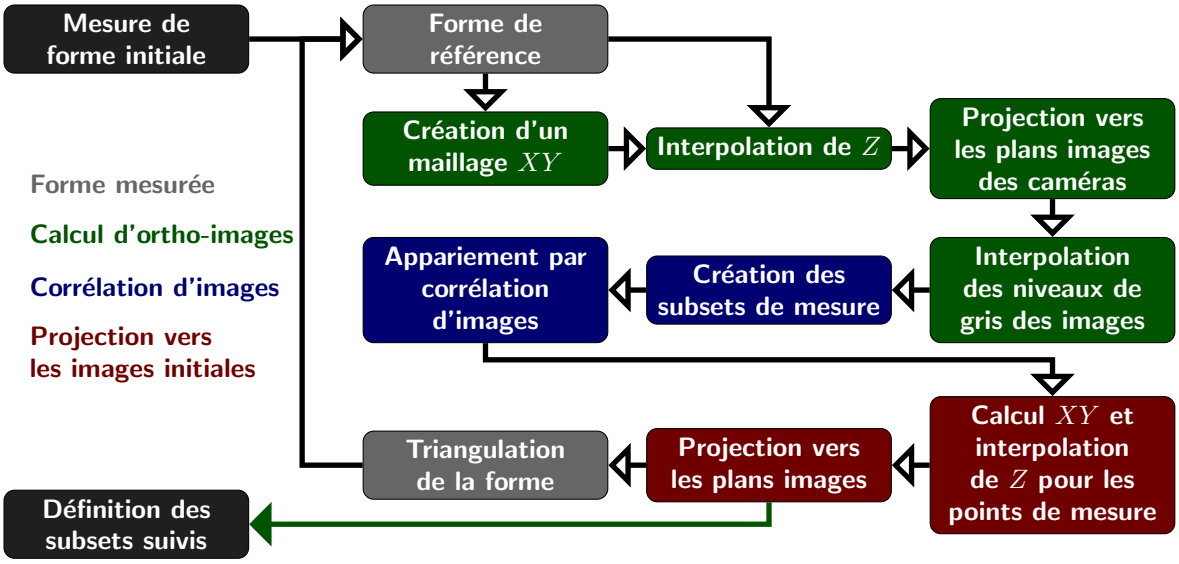


FIGURE 7.4 – Diagramme synoptique détaillé de la méthode IRIs

Mesure de forme initiale

La mesure de forme initiale est réalisée à partir d'images rectifiées calculées à l'aide de la géométrie épipolaire (*cf.* **section 4.6.2**). Si l'appariement automatique ne donne pas des résultats satisfaisants, quelques points de mesure sont définis, à des positions permettant de décrire approximativement la forme observée, et appairés manuellement avec des entrées graphiques. Les droites épipolaires sont alors tracées pour faciliter l'opération.

Calcul d'ortho-images

Une ortho-image est un image vue par une caméra orthoptique située en face de la surface à mesurer. Son modèle est différent du modèle de caméra sténopé : les rayons perspectifs ne convergent pas vers le centre optique, mais sont tous parallèles à l'axe $\overrightarrow{z}$ du repère image associé [173]. La **figure 7.5** illustre la différence entre une image issue d'un modèle sténopé, une image rectifiée et une ortho-image, à partir de la profondeur d'un objet projetée sur les différents plans images. Notons que la forme prise en exemple présente une symétrie de révolution et que le code couleur choisi pour afficher la profondeur va du *bleu* au *jaune*. À *gauche*, la forme vue par un modèle sténopé, placé sur le côté de l'objet avec un angle proche de 30°, est ellipsoïdale et de fortes distortions de perspectives, non linéaires, sont visibles, (*i.e.* les centres des ellipses images pour différentes profondeurs de l'objet ne sont pas confondus). Au *milieu*, sur l'image rectifiée avec la méthode basée sur la géométrie épipolaire (*cf.* **section 4.6.2**), chaque niveau de couleur décrit bien un cercle, mais les différents cercles, correspondant à différentes profondeurs, ne sont pas concentriques. À *droite*, l'ortho-image affiche une forme circulaire, centrée autour du centre de l'image, quelle que soit la profondeur de l'objet.

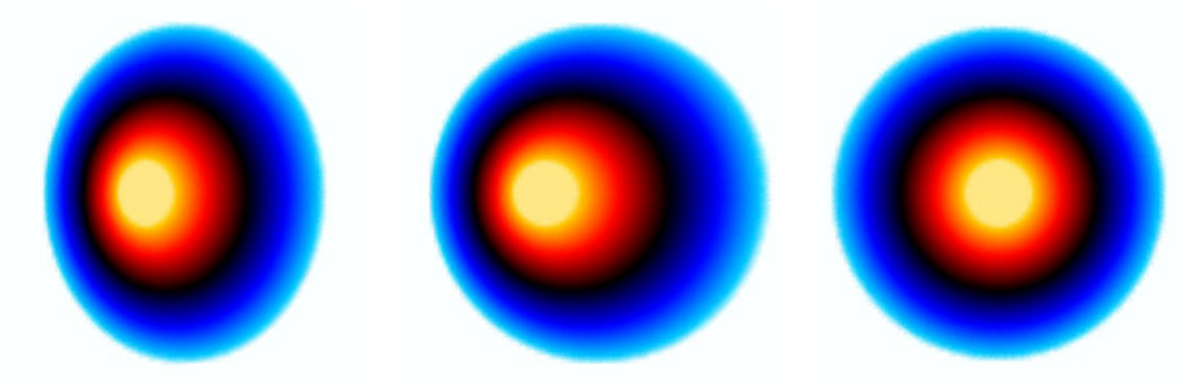


FIGURE 7.5 – Images de la profondeur d'un objet ; (*gauche)* image vue par une caméra sténopé sur le côté, (*milieu*) image rectifiée, (*droite*) ortho-image

Une ortho-image est calculée à partir d'une image obtenue par une caméra de position et d'orientation connues par rapport à une forme estimée, dite de référence. La **figure 7.6** en illustre le principe pour le point de vue de la caméra gauche. Les pixels de l'ortho-image (pour illustration les deux points *noir* en bas à droite de l'image) sont virtuellement positionnés devant la forme estimée. La profondeur de la forme mesurée est ensuite interpolée (flèches *noires*), afin d'obtenir une position 3D (points *rouges*), pour chaque pixel. Ces positions 3D sont ensuite projetées (flèches *rouges*) dans le repère image initial de la caméra avant d'effectuer une interpolation des niveaux de gris.

Le protocole de calcul est le suivant :

1. calculer le plan moyen de la surface mesurée ;
2. définir la matrice de pixels dans le plan moyen de la surface ;
3. interpoler la profondeur de la surface à partir de la forme 3D de référence ;
4. effectuer une rotation et translation des positions 3D vers le repère de la caméra ;
5. projeter des positions 3D des pixels dans le repère image ;
6. appliquer les distortions optique aux positions 2D et interpoler les niveaux de gris ;
7. filtrer l'image obtenue à l'aide d'un filtre gaussien.

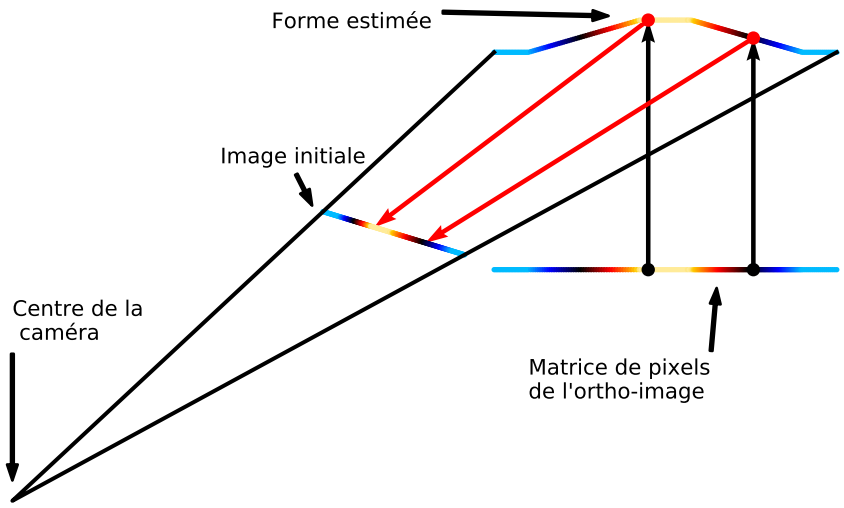


FIGURE 7.6 – Schéma de principe du calcul d'une ortho-image

Le plan moyen de la surface est ainsi estimé afin d'obtenir la normale et de calculer les matrices de rotation R_{pm} et de translation T_{pm} correspondantes (*cf.* **section 3.3.4**). Les positions 2D (X_i^{pm}, Y_i^{pm}) dans le plan moyen, associées aux positions 2D (u_i, v_i) de la matrice de pixels définie, de dimension $(n_{rows} \times n_{cols})$, sont calculées telles que :

$$(X_i^{pm},\, Y_i^{pm}) = \left(\alpha * \frac{wdth_{obj}}{n_{cols}} u_i - \overline{X},\, \alpha * \frac{hght_{obj}}{n_{rows}} v_i - \overline{Y}\right) \tag{7.1}$$

avec α un coefficient (ici égal à 1,1), $wdth_{obj}$, $hght_{obj}$, $\overline{X}$ et $\overline{Y}$ respectivement la largeur, la hauteur, et les valeurs moyennes des positions (X, Y) de la forme estimée. Les valeurs de profondeur Z_i^{pm} sont ensuite calculées par interpolation de la profondeur de référence aux positions (X_i^{pm}, Y_i^{pm}). Pour des positions (X_i^{pm}, Y_i^{pm}) en dehors de la surface de référence, les valeurs de profondeur Z_i^{pm} correspondantes sont définies comme égales à la valeur de profondeur du point le plus proche dans la surface de référence. Les positions 3D $(X_i^{pm}, Y_i^{pm}, Z_i^{pm})$ obtenues sont ensuite exprimées dans le repère du capteur stéréoscopique avec les matrices R_{pm} et T_{pm}, et projetées dans les repères images des caméras, en tenant compte des distortions optiques des objectifs.

Corrélation d'images

De nouveaux points de mesure sont ainsi définis dans l'image de référence. Les positions des points homologues dans la seconde ortho-image sont obtenus au pixel près avec une corrélation 2D (*cf.* **section 4.6.1**) avec des subsets de forme circulaire (*cf.* **section 4.6.1**). Une corrélation sub-pixélique est ensuite calculée avec la méthode basée sur le gradient. Les distortions de perspective sont partiellement corrigées dans les ortho-images calculées pour les deux caméras, et permettent une corrélation automatique de plusieurs centaines de points de mesure si la forme initiale est estimée avec une précision et un nombre de points suffisants. Dans le cas contraire, quelques points peuvent encore être repositionnés avec une entrée graphique.

Projection vers les images initiales

Une fois les points de mesure appairés dans les ortho-images, les positions 2D sont projetées sur le plan initial de la même manière que pour le calcul des ortho-images : les positions 3D correspondantes sont estimées avec la forme de référence puis les positions 2D dans le repère image de chaque caméra sont calculées. Ces points permettent d'estimer une nouvelle forme par triangulation.

Processus itératif et sortie de la boucle

La forme estimée à l'aide des ortho-images est ensuite utilisée pour calculer un nouvelle paire d'ortho-images. Les distortions de perspective étant davantage corrigées, un appariement plus précis est effectué, permettant d'estimer une nouvelle forme, et ainsi de suite. Lorsque les ortho-images sont suffisamment corrélées, les points de mesure finaux sont choisis et la forme finale estimée.

Pour la mesure de vibrations, des subsets de suivi sont alors définis dans les images initiales, à partir de huit points situés sur un cercle autour de chaque point de mesure dans la dernière paire d'ortho-images. Ces points sont projetés vers les images initiales afin de calculer une ellipse, pour chaque point de mesure, dans chaque image, pour que les subsets suivis dans les images des deux points de vue correspondent à la même zone de mesure sur l'objet. La **figure 7.7** illustre le processus : le point de mesure est en *rouge* et les huit points en *bleu* sont utilisés pour calculer les ellipses en *vert*. Les pixels à l'intérieur des ellipses servent de subsets et sont suivis dans la ou les séquences temporelles d'images. De par la perspective, le subset de gauche est visuellement plus étroit que le subset de droite, mais l'image à l'intérieur est bien similaire, à des distortions près.

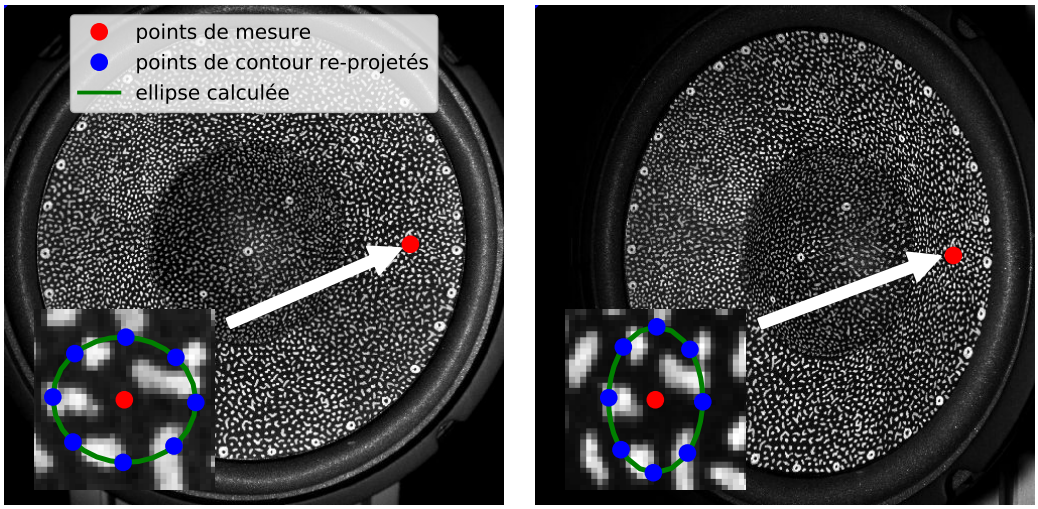


FIGURE 7.7 – Subsets définis pour un point de mesure dans les deux images

7.3 Validation expérimentale : mesure de forme initiale du haut-parleur

Une validation expérimentale est effectuée en mesurant la forme du haut-parleur étudié au **chapitre 6** avec le montage pseudo stéréoscopique. Un motif aléatoire, constitué de traits de ≈ 2 mm de large, a été dessiné sur la membrane du haut-parleur à l'aide d'un crayon pastel blanc. Afin d'optimiser la mesure par stéréo-corrélation d'images, il est habituellement recommandé d'avoir des taches de trois à cinq pixels dans les images [151]. Ici, le tracé est délibérément plus gros afin de faciliter l'appariement manuel de certains points.

La caméra utilisée est la Photron SA-X2 du LAUM. Le calibrage de la caméra est réalisé avec une mire carrée constitué de 19×19 cibles, espacées de ≈ 10 mm. La distance focale obtenue est de ≈ 2322 pixels ($\approx 46{,}4$ mm). Le calibrage du système stéréoscopique est ensuite réalisé à partir de cinq images d'une mire en spirale (*cf.* **section 5.3**) de ≈ 210 mm de diamètre pour le cercle externe, avec 176 cibles. La distance estimée entre les caméras est de 997 mm, et les angles de rotation autour des axes $\overrightarrow{y}$, $\overrightarrow{x}$ et $\overrightarrow{z}$ sont de 75,7°, 2,7° et -3,4° respectivement. Au niveau de la mire,

l'erreur de re-projection est en moyenne de 0,13 pixel, avec un maximum de 0,44 pixel. Ces distances épipolaires sont environ deux fois plus petites que pour le calibrage réalisé avec 26 images de la mire asymétrique de la bibliothèque OpenCV (*cf.* **section 6.5**), pour lequel l'erreur de re-projection était de 0,2 pixel en moyenne avec un maximum de 0,93 pixel. L'utilisation de la mire en spirale conduit à une amélioration du calibrage, par rapport à l'utilisation de la mire asymétrique d'OpenCV. Cet écart s'explique vraisemblablement par la différence du nombre de cibles, quatre fois supérieur pour la mire en spirale, et par le choix de cibles plus petites, induisant un moins grand décalage entre les centres des images des cibles et les images de leurs centres.

Mesure de forme initiale

Les images initiales sont tout d'abord rectifiées. La forme à mesurer présentant une symétrie de révolution, les points de mesure sont placés sur plusieurs cercles : un au bord du cône, un au bord du dôme, un entre ces deux cercles et un sur le dôme. Sur soixante-trois points de mesure (en *rouge* sur la **figure 7.8** à *gauche*) définis dans l'image du point de vue de la caméra de gauche, vingt-et-un sont mal appairés (*cf.* **figure 7.8** au *milieu*) et sont positionnés manuellement à l'aide d'une entrée graphique (*cf.* **figure 7.8** à *droite*).

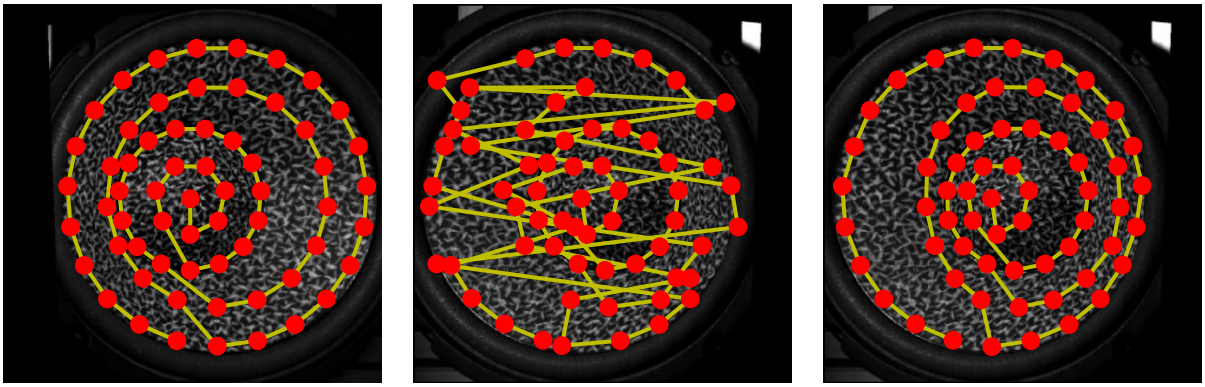


FIGURE 7.8 – Images rectifiées et points de mesure ; (*gauche*) point de vue gauche et points de mesure définis ; (*milieu*) point de vue droite et points de mesure appairés par corrélation autour des droites épipolaires ; (*droite*) point de vue droite et points repositionnés manuellement

Première ortho-image

Les ortho-images calculées avec les soixante-trois points de mesure (*cf.* **figure 7.8**) sont affichées sur deux couleurs (*rouge* et *vert*) sur la **figure 7.9**. La dimension des ortho-images est réglée à 750 × 750 pixels, soit une matrice carrée de côté approximativement 10% plus grande que la hauteur des images initiales (voir la **figure 7.1**). Les distortions de perspective sont partiellement corrigées, et le coefficient de corrélation des ortho-images, calculé dans le cercle *bleu*, est de 53,8%. Les ortho-images se ressemblent alors suffisamment pour effectuer une corrélation automatique de davantage de points de mesure.

Ces nouveaux points de mesure sont positionnés dans l'image de référence (points *rouge* sur la **figure 7.10** à *gauche*), ici, l'ortho-image correspondant à la caméra de gauche, sur des cercles concentriques. Le cercle extérieur est défini à partir de trois points sélectionnés graphiquement. Les positions des points homologues dans la seconde ortho-image sont affichés sur la **figure 7.10** à *droite*. Sur 322 points, moins de dix points sont à re-positionner manuellement avec une entrée graphique. Les positions projetées dans les images initiales sont affichées sur la **figure 7.11**.

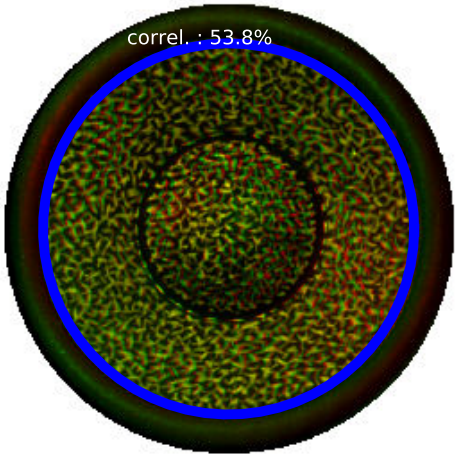


FIGURE 7.9 – Ortho-images combinées sur deux couleurs (*rouge* et *vert*) ; le coefficient de corrélation des images dans le cercle *bleu* est de 53,8%

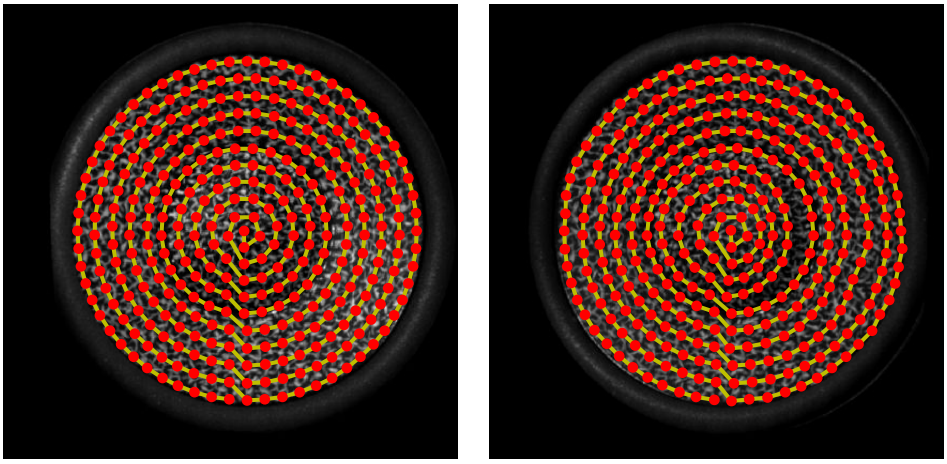


FIGURE 7.10 – Ortho-images et points de mesure appairés ; (*gauche*) image calculée pour le point de vue de gauche ; (*droite*) image calculée pour le point vue de droite

Processus itératif

Après la première itération, la paire d'ortho-images est corrélée à 88,7%, et les 1201 points de mesure définis sont appairés automatiquement. Après la deuxième itération, les ortho-images sont corrélées à 92,0% : sur la **figure 7.12**, les mouchetis se superposent presque parfaitement. Les différences sont principalement dues aux réflexions spéculaires, correspondant aux zones plus rouges ou plus vertes, visibles à différents endroits et notamment sur le dôme. L'évolution du nombre de points de mesure et du coefficient de corrélation des ortho-images, pour le cas présenté, est affichée dans le **tableau 7.1**.

TABLEAU 7.1 – Évolution du nombre de points de mesure et du coefficient de corrélation des ortho-images

	mesure initiale	itération 1	itération 2
Nombre de points de mesure	63	322	1201
Corrélation des ortho-images résultantes	53,8%	88,7%	92,0%

FIGURE 7.11 – Points de mesure projetés vers les images initiales

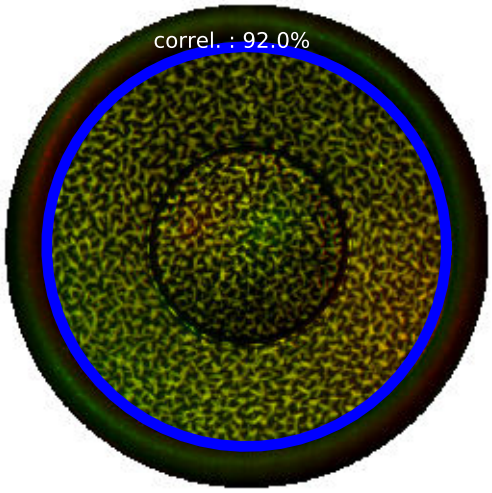


FIGURE 7.12 – Ortho-images à l'itération 3, combinées sur deux couleurs (*rouge* et *vert*) ; le coefficient de corrélation des images dans le cercle *bleu* est de 92,0%

Les images convergeant à plus de 90%, le nombre de points de mesure peut alors être choisi en fonction de la résolution des images initiales brutes et des besoins de la mesure. Pour la mesure de vibrations, il n'est en effet pas toujours nécessaire d'avoir un très grand nombre de points de mesure. Pour rappel, l'idéal en simulation numérique est d'avoir au moins dix points par longueur d'onde, et six points sont considérés comme un minimum pour minimiser les erreurs de calcul.

Distances épipolaires

Les ortho-images convergent correctement, ce qui est en soit un indicateur de convergence de l'algorithme et apporte une validation expérimentale de la méthode. Cependant, lorsque l'on regarde la carte des distances épipolaires des points de mesure pour la mesure de forme finale, affichée sur la **figure 7.13**, des zones pour lesquelles la distance épipolaire est comprise entre $\approx$ 0,6 à 0,9 pixel sont visibles (en *jaune* et *blanc* sur la figure). Leur regroupement en deux ou trois taches principalement fait suspecter un problème potentiel au niveau de la planéité des miroirs. Une simulation a alors été réalisée avec le logiciel Blender [18] afin d'obtenir conjointement une seconde validation pour la méthode proposée et des informations sur la cause de ces distances épipolaires localement assez élevées.

Les résultats sont présentés dans les sections suivantes.

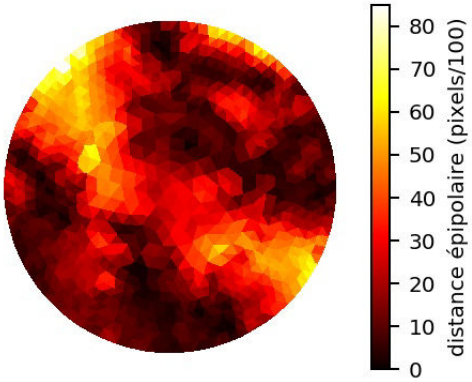


FIGURE 7.13 – Carte des distances épipolaires calculées au niveau du haut-parleur pour la mesure de forme finale

7.4 Validation numérique de la méthode

Bien que les résultats obtenus lors des mesures réalisées semblent cohérents, une seconde validation, numérique cette fois, a été recherchée, par le biais d'un modèle créé sous Blender. Rappelons que l'objectif de cette simulation est aussi de déterminer l'origine des zones d'augmentation des distances épipolaires observées lors de la validation expérimentale (*cf. supra*).

Un modèle est ainsi réalisé à partir des images et des données obtenues expérimentalement, afin d'obtenir des images réalistes du haut-parleur avec un montage pseudo stéréoscopique. Premièrement, une image du haut-parleur avec le mouchetis est acquise avec la Photron SA-X2, positionnée en face de l'objet d'étude, avec un objectif de 50 mm, donnant un angle de champ de vision de $\approx 23°$. Un matériau *image texture* est ensuite généré, avec une image codant le profil, et utilisé pour déformer la surface plane, avec un *object modifier*, afin d'obtenir une forme réaliste du dôme et du cône. Le système pseudo stéréoscopique est simulé (*cf.* **section 6.3.3**) avec des miroirs plans et un angle entre les caméras virtuelles et la surface réglé à 40°. Une capture d'écran du système résultant est affichée sur la **figure 7.14**.

Des images sont ensuite simulées de manière similaire à l'expérimentation décrite en **section 7.3**, à savoir cinq images d'une mire carrée pour calibrer la caméra, cinq images d'une mire en spirale pour calibrer le système pseudo stéréoscopique et une image pour effectuer la mesure de forme avec la méthode IRIs. Ces données sont utilisées pour appliquer le protocole global de mesure de forme de la méthode. Après le calibrage de la caméra et du système, la mesure initiale est effectuée à partir d'images rectifiées une première fois, puis les ortho-images sont calculées itérativement. Une forme 3D mesurée est finalement obtenue, dont on trace un profil et que l'on compare au profil simulé. Les résultats sont affichés sur la **figure 7.15**, avec en ligne *rouge* pleine le profil simulé et en ligne *noire* discontinue le profil mesuré, à partir des images simulées, après quatre itérations. Hormis une légère discontinuité au bord du dôme, on note une très bonne correspondance visuelle.

Afin de quantifier cette concordance, l'erreur de mesure, correspondant à la différence entre la valeur simulée et celle mesurée à chaque point de mesure, est alors calculée. Rappelons que les valeurs

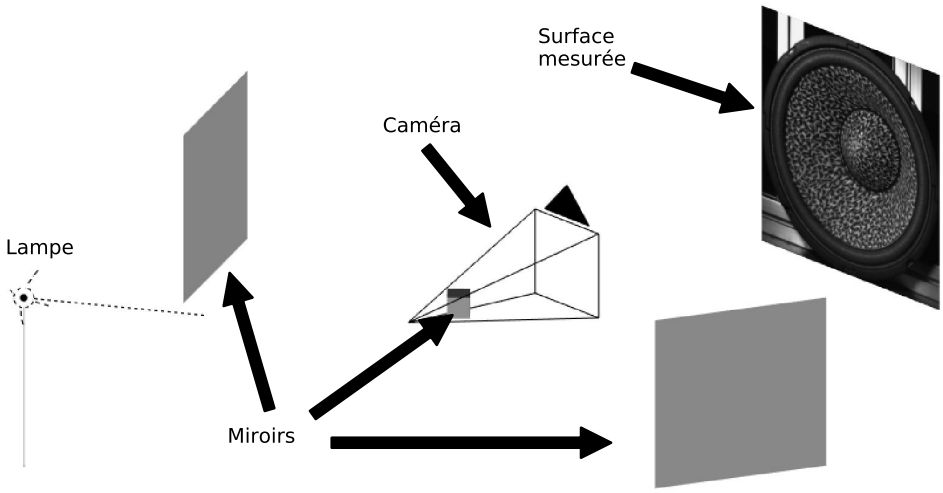


FIGURE 7.14 – Capture d'écran de la simulation réalisée sous Blender

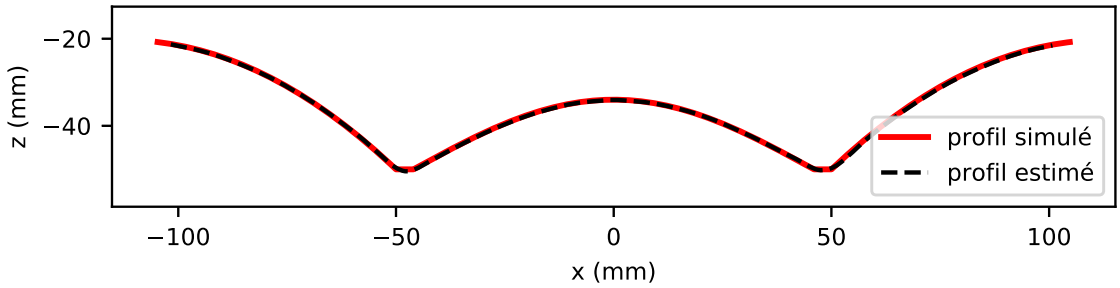


FIGURE 7.15 – Profil simulé et profil estimé à partir des images simulées

mesurées sont obtenues à partir d'un système simulé et sont ainsi très probablement en deçà des valeurs que l'on obtiendrait pour une mesure effectuée avec un système réel. Leur intérêt ici est à titre indicatif. La **figure 7.16** affiche les résultats. Les zones correspondant au dôme et au cône sont facilement identifiables, ainsi que la zone de discontinuité de la forme aux bords du dôme, même si l'on note une légère torsion de la forme mesurée. On constate que la moyenne de la valeur absolue de l'erreur est de $\approx$ 52 µm et que l'écart-type est de $\approx$ 79 µm avec des différences maximales de $\approx$ 300 µm au niveau de la discontinuité, qui ne peut être correctement décrite avec le mouchetis utilisé et le peu de pixels à disposition (soit environ 500 $\times$ 400 pixels). Ces valeurs confirment la très bonne concordance des deux profils et apportent une validation numérique de la méthode IRIs utilisée.

7.5 Non-planéité des miroirs

Le modèle Blender créé pour obtenir une validation numérique de la méthode IRIs a été utilisé afin de chercher l'origine des taches visibles sur la carte des distances épipolaires (*cf.* **figure 7.13**). Les distances épipolaires observées au niveau de la surface sont calculées pour la mesure effectuée à partir des images simulées avec des miroirs plans. La **figure 7.17** affiche les résultats obtenus, à *gauche*. Aucune tache n'est visible et les valeurs sont relativement faibles : $\approx$ 0,03 pixel en moyenne avec un écart-type de $\approx$ 0,04 pixel. Pour comparaison, la carte des distances épipolaires associées aux mesures effectuées avec le montage pseudo stéréoscopique réel est affichée sur la **figure 7.17** au *milieu*. On y voit nettement plusieurs taches, dont les valeurs sont comprises, pour rappel, entre $\approx$ 0,6 et 0,9 pixel.

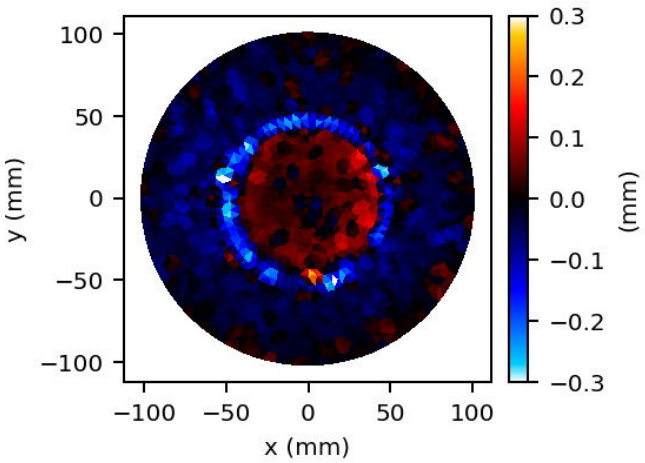


FIGURE 7.16 – Carte de l'erreur de la mesure de forme

Une potentielle non-planéité des miroirs est alors envisagée comme origine de cette différence.

Afin de confirmer cette hypothèse, les miroirs ont été mesurés avec un interféromètre ZYGO [46] à SupOptique (Palaiseau). Les résultats montrent des écarts de profondeur légèrement supérieurs à 3 µm, valeur de tolérance indiquée par le fournisseur des miroirs. Un second modèle a finalement été conçu, avec des paramètres identiques au premier, sauf en ce qui concerne les miroirs. Ceux-ci ont été modifiés afin d'intégrer des déformations similaires à celles mesurées avec l'interféromètre. De nouvelles images ont ensuite été générées. Rappelons à nouveau que les valeurs obtenues par simulation sont souvent en deçà des valeurs réelles. L'intérêt principal de l'approche réside dans la possibilité d'une comparaison de différentes configurations, avec un ordre de grandeur associé pour les valeurs calculées.

Le protocole déjà utilisé est appliqué à nouveau à la seconde série d'images simulées. La carte de distances épipolaires résultante, affichée sur la **figure 7.17** à *droite* présente de petites taches, correspondant à des valeurs de l'ordre de 0,7 pixel. On constate que la granularité de ces taches est globalement du même niveau que sur la carte du *milieu* et que le bruit de fond est également similaire. Ceci laisse à penser que les taches observées sont induites par la non-planéité des miroirs et permet de valider une fois de plus, indirectement, la méthode IRIs.

L'impact de la non-planéité des miroirs sur la mesure de vibrations aurait aussi pu être vérifié par simulation. Néanmoins, le travail de simplification du montage, mené en parallèle, ayant abouti à la suppression des miroirs pour certains objets (*cf.* **chapitre 9**), ce point n'a pas été étudié.

7.6 Automatisation de l'appariement initial de la méthode IRIs

Comme vu précédemment, la méthode IRIs permet d'effectuer un appariement d'un grand nombre de points de mesure pour des objets comportant d'importantes variations de profondeur et ce avec un grand angle entre les caméras d'un système (pseudo) stéréoscopique. Toutefois, la méthode requiert une mesure de forme initiale qui, dans les sections présentées ci-dessus, a dû être effectuée en partie manuellement avec des entrées graphiques. La présente section propose une amélioration de la méthode qui conduit à l'automatisation quasi-totale de cette mesure de forme initiale : à l'instar du système stéréoscopique trinoculaire ATSIS [164], un troisième point de vue est utilisé pour initialiser la mesure. Les trois points de vue requis peuvent être obtenus à partir d'un montage pseudo stéréoscopique et

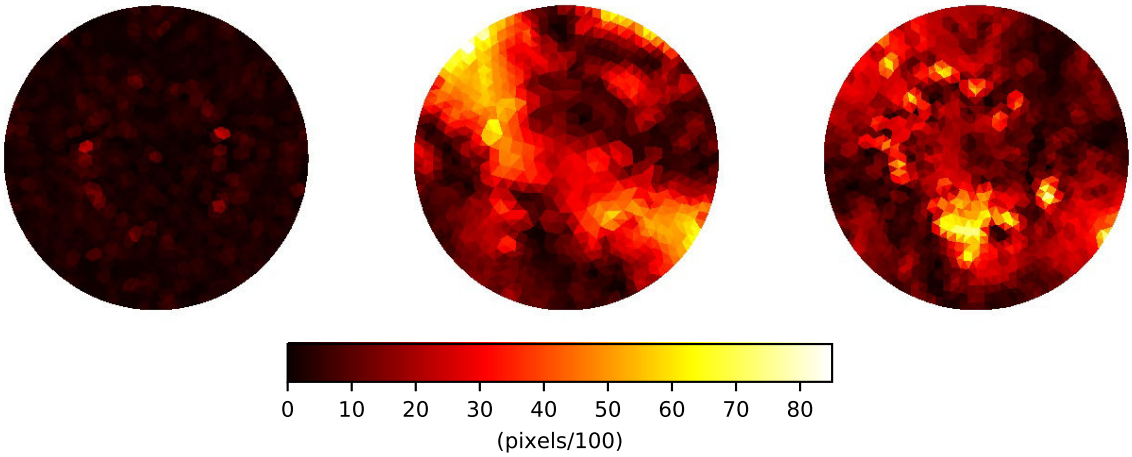


FIGURE 7.17 – Cartes de distances épipolaires ; (*gauche*) pour la simulation avec miroirs plans, (*milieu*) pour la mesure avec le système pseudo stéréoscopique, (*droite*) pour la simulation avec les miroirs déformés

d'une deuxième caméra ou bien à partir de trois caméras. C'est cette seconde option qui est décrite ici.

Montage

Le système stéréoscopique principal est réalisé avec les caméras Phantom et Photron du LAUM, avec des objectifs de 25 mm et 50 mm respectivement. La troisième caméra est une caméra Mako U-503B avec un objectif de 17 mm. La taille des images est de 1024 × 1024 pixels pour les caméras Photron et Phantom, et de 1944 × 2592 pixels pour la Mako. La surface de l'image de la membrane du haut-parleur au niveau de la Mako est, pour des raisons pratiques d'agencement avec les objectifs à disposition, de 1200 × 1200 pixels environ. Une photographie du montage est présentée sur la **figure 7.18** : les caméras Phantom et Photron sont positionnées juste devant le haut-parleur, et un spot LED est inséré entre les deux. Sur la droite, la caméra Mako est fixée sur un trépied, et légèrement surélevée pour ajuster la vue de l'objet d'étude.

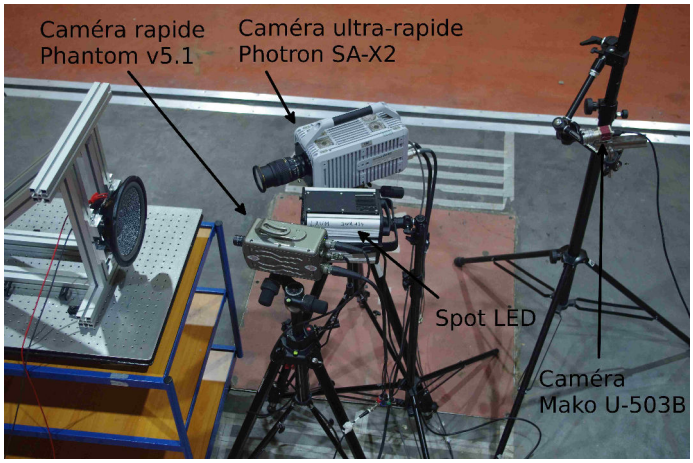


FIGURE 7.18 – Photographie du montage stéréoscopique trinoculaire réalisé

Calibrage

Quatre clichés d'une mire en spirale, comportant 784 points espacés de $\approx$ 5,3 mm, sont acquis, à différentes positions, par chaque caméra, avec un intervalle de temps inférieur à une seconde. Ces images servent à calibrer deux sous-systèmes stéréoscopiques : le premier est constitué des caméras Photron et Phantom et sert pour une mesure de vibrations, le second est formé par la paire de caméras Phantom - Mako, et est utilisé pour effectuer une première mesure de forme, qui sert d'initialisation à la méthode IRIs.

Les angles de rotation des différentes caméras par rapport à la mire sont donnés dans le tableau **tableau 7.2**.

Tableau 7.2 – Angles de rotation des différentes caméras par rapport à la mire

Caméra	rotation $\overrightarrow{y}$	rotation $\overrightarrow{x}$	rotation $\overrightarrow{z}$
Photron	34,8°	0,4°	1,5°
Phantom	-10,1°	5,6°	-0,6°
Mako	9,1°	12,5°	4,6°

Les angles de rotation autour des axes $\overrightarrow{y}$ et $\overrightarrow{x}$ induisent des distorsions de perspective et sont respectivement égaux, pour la paire Phantom-Mako, à $(9,1°+10,1°)$ et $(12,5°-5,6°)$, soit 19,2° et 6,9°. Ces angles sont inférieurs à 30°, valeur maximale habituellement préconisée pour la mesure de forme. Ces valeurs assez faibles induisent alors suffisamment peu de distorsions de perspective et permettent un appariement automatique des points de mesure dans la paire d'images initiales rectifiées en utilisant la méthode basée sur la géométrie épipolaire présentée en **section 4.6.2**.

Mesure initiale

Une première mesure de forme est ainsi réalisée avec le sous-système stéréoscopique Phantom-Mako afin d'initialiser la méthode IRIs, appliquée par la suite aux images acquises par le système Photron-Phantom. La caméra de référence choisie pour le premier système est la Phantom car elle est commune aux deux capteurs, ce qui permet d'optimiser le lien entre les positions 3D estimées avec la paire de caméras Phantom-Mako et le repère associé au système Photron-Phantom. De plus, l'objectif est d'obtenir un appariement initial avec un minimum d'entrées de la part de l'utilisateur. Ainsi, un algorithme adapté aux spécificités du haut-parleur a été créé, tenant compte de la symétrie de révolution de la forme de la membrane :

1. trois points (au minimum) (en *rouge* sur la **figure 7.19** à *gauche*) sont sélectionnés dans l'image de référence ;
2. les points homologues, dans la seconde image (en *rouge* sur la **figure 7.19** à *droite*), sont appairés manuellement à l'aide des droites épipolaires ;
3. la position 3D de ces points est triangulée ;
4. un cercle (2D) est ensuite estimé dans le plan moyen de ces positions 3D ;
5. des points sur ce cercle, ici 88, sont alors définis puis projetés dans les images initiales (*i.e.* les ellipses *jaunes* sur la **figure 7.19**) puis dans les images rectifiées (*i.e.* les cercles *jaunes* sur la **figure 7.20**).

Notons que les positions 2D estimées dans les images initiales (les ellipses *jaunes* sur la **figure 7.19**) permettent de vérifier visuellement le positionnement des points de mesure. Dans les images rectifiées, de dimension 1024 $\times$ 1024, affichées sur la **figure 7.20**, les positions 2D correspondantes sont utilisées pour délimiter la zone de mesure (les cercles *jaunes*). Pour cette mesure, 297 points sont définis, en plus des 88 points du cercle, selon un quadrillage asymétrique. Les points mal appairés étant recalés manuellement, le nombre de points de mesure est volontairement limité afin de limiter le temps nécessaire au recalage de ces points. Sur les 385 points de mesure, seulement trois doivent être repositionnés

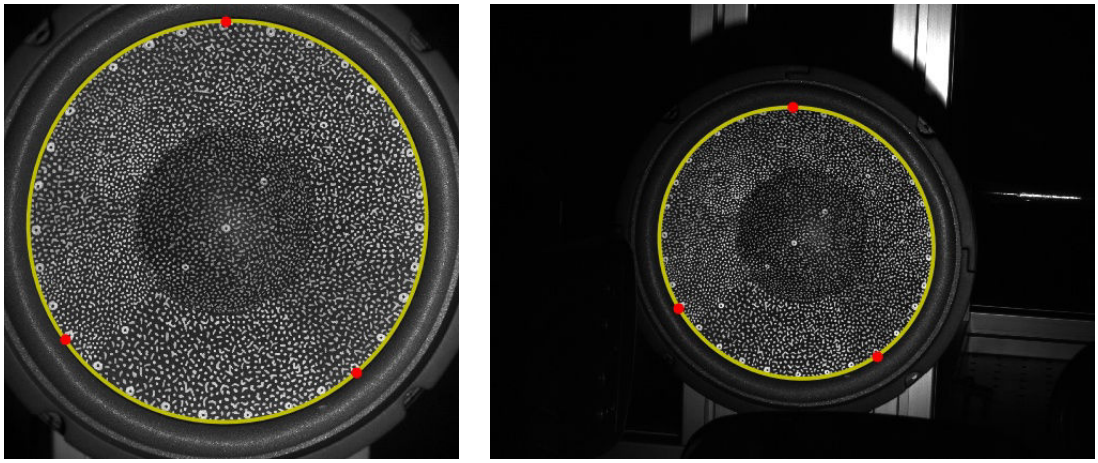


FIGURE 7.19 – Images initiales acquises avec la caméra Phantom (*gauche*) et Mako (*droite*) ; les points *rouges* sont sélectionnés manuellement, appairés à l'aide des droites épipolaires et servent à estimer un cercle dans le plan moyen des positions 3D triangulées ; ce cercle est ensuite projeté dans les images (ellipses *jaunes*)

avec une entrée graphique après la phase d'appariement automatique ; ils pourraient néanmoins être négligés avec un algorithme adapté. Finalement, les distances épipolaires correspondantes sont en moyenne de $\approx 0{,}37$ pixel avec un écart-type de $\approx 0{,}24$ pixel.

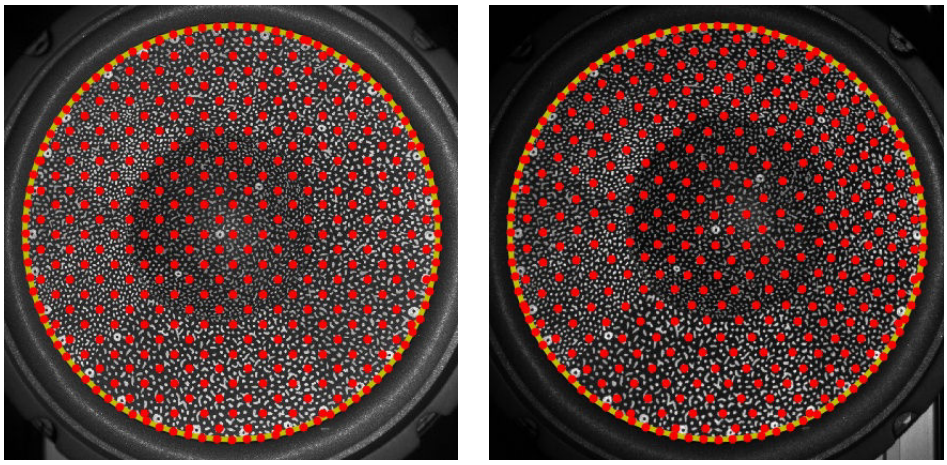


FIGURE 7.20 – Images rectifiées ; les 385 points de mesure en *rouge* sont à l'intérieur de la zone délimitée par le cercle initial *jaune* projeté dans chaque image rectifiée

Processus itératif

La méthode IRIs est ensuite utilisée avec le système Photron-Phantom, avec des ortho-images calculées sur 1024×1024 pixels. La différence des paramètres externes calculés pour les deux calibrages est, ici, négligée : l'écart est inférieur à 0,02° pour les angles de rotation et inférieur à 100 µm pour la translation, comme l'affiche le **tableau 7.3**.

Tableau 7.3 – Différence entre les paramètres externes calculés pour la caméra Phantom entre les deux calibrages

Axe	x	y	z
Rotation	0,016°	-0,011°	0,001°
Translation (mm)	0,06	0,09	0,01

La première paire d'ortho-images (*cf.* **figure 7.21** à *gauche*) présente un degré de corrélation de 85,4% pour la zone de mesure indiquée par le cercle *bleu*. Les 1741 points de mesure définis selon des cercles concentriques sont alors appairés automatiquement. À l'itération 1, les ortho-images sont corrélées à 96,8% (*cf.* **figure 7.21** à *droite*), et 2698 points sont de nouveau définis, amenant à une deuxième paire d'ortho-images aussi corrélées à 96,8%. Les principales différences sont dues à des différences de réflexions spéculaires au niveau des bords du dôme du côté gauche de l'image. L'évolution du nombre de points de mesure et du degré de corrélation des ortho-images est présentée dans le **tableau 7.4**.

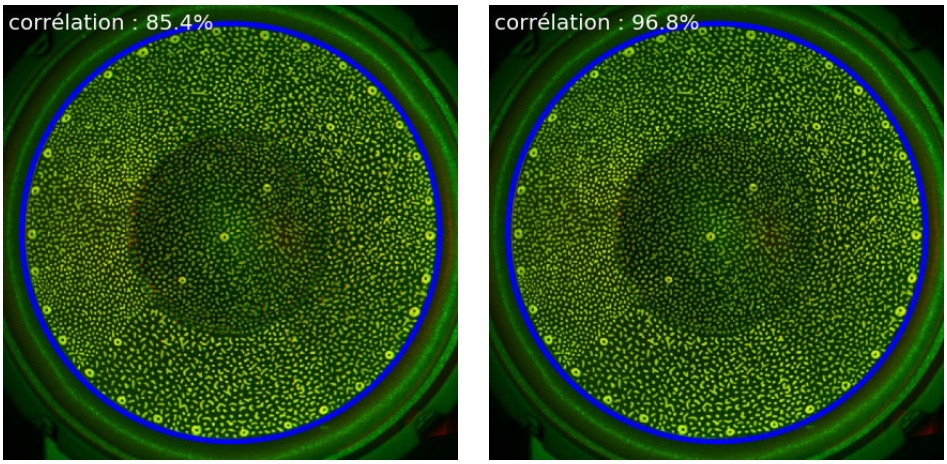


Figure 7.21 – Images rectifiées à la première (*gauche*) et seconde (*droite*) itérations, combinées sur deux couleurs (*rouge* et *vert*)

Tableau 7.4 – Évolution du nombre de points de mesure et de la corrélation des ortho-images

	initiale	itération 1	itération 2	itération 3
Nombre de points de mesure	385	1741	2698	3777
Degré de corrélation	85,4%	96,8%	96,8%	97,9%

À la troisième et dernière itération, 3777 points sont définis dans une zone légèrement plus petite afin que tous les points soient à l'intérieur de la surface estimée à l'itération précédente. À titre indicatif, la paire d'ortho-images résultante affiche un degré de corrélation de 97,9%. La forme et la carte des distances épipolaires associées sont affichées sur la **figure 7.22**. La profondeur mesurée est cohérente et les distances épipolaires sont en moyenne de $\approx$ 0,03 pixel, avec un écart-type de $\approx$ 0,03 pixel. Ces valeurs sont approximativement 10 fois inférieures aux valeurs obtenues pour l'appariement initial, qui étaient en moyenne de $\approx$ 0,37 pixel avec un écart-type de $\approx$ 0,24 pixel. Aucune tache correspondant à des valeurs élevées n'est identifiable. On remarque seulement quelques points visibles au niveau du bord du dôme et sur le bord supérieur de la membrane, avec des valeurs de $\approx$ 0,2 pixel, ce qui semble confirmer encore une fois que les zones de distances épipolaires supérieures à 0,6 pixel sont induites par la non-planéité des miroirs (*cf.* **section 7.5**). L'agencement de ces points laisse ici plutôt à penser que les écarts sont principalement dus à un facteur géométrique induit par la disposition des deux

caméras. Finalement, ces résultats valident le protocole proposé afin d'automatiser la mesure initiale requise pour l'application de la méthode IRIs présentée.

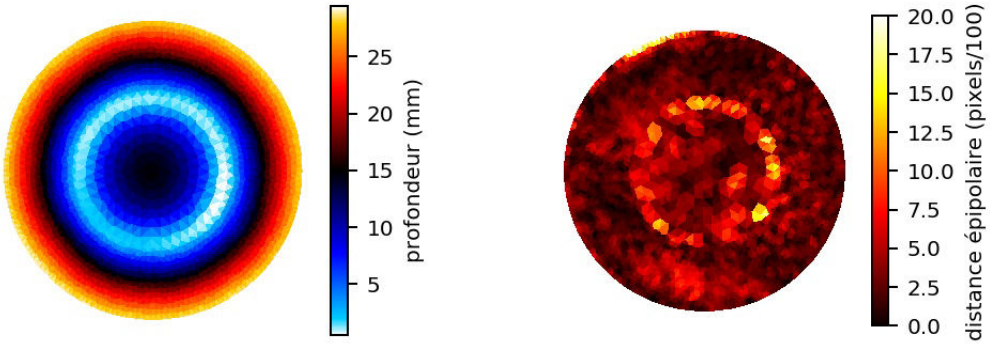


FIGURE 7.22 – (*gauche*) Profondeur mesurée, (*droite*) carte des distances épipolaires

7.7 Conclusion

Ce chapitre a exposé une méthode permettant d'appairer, à partir d'une mesure de forme initiale, des points de mesure avec un angle supérieur à 60° entre les caméras d'un système stéréoscopique, et ce même pour des objets présentant d'importantes variations de profondeur. Comparativement aux méthodes usuelles estimant la déformation des subsets de mesure, cette méthode est relativement simple d'implémentation. À l'instar de la méthode de stéréo-corrélation fine améliorée [49], les calculs sont effectués sur des valeurs de niveaux de gris directement interpolées dans les images initiales, réduisant ainsi les imprécisions dues aux interpolations successives. Notons aussi que la méthode proposée estime de manière itérative la forme globale de la surface mesurée, et non la déformation locale de chaque subset, ce qui permet d'utiliser une mesure de forme initiale très grossière avec des points appairés avec un précision relativement faible, éventuellement en partie avec des entrées graphiques. La méthode a été validée expérimentalement et numériquement avec un haut-parleur comme objet d'étude.

Afin de limiter au maximum l'intervention de l'usager, une solution permettant d'automatiser l'appariement pour la mesure de forme initiale a été conçue en complément. Elle utilise, à l'instar du système stéréoscopique trinoculaire ATSIS [164], un troisième point de vue présentant un angle panoramique relativement faible avec l'une des deux caméras du système stéréoscopique positionné avec un angle panoramique important. La mesure initiale n'ayant pas besoin d'être extrêmement précise, il n'est pas nécessaire d'avoir une caméra très onéreuse. La caméra choisie pour cette expérimentation (la Mako U-503B) coûte, à titre indicatif, 450 €. Notons que moins de la moitié des pixels a été utilisée : une caméra moins performante aurait pu aussi tout à fait convenir. La méthode finale, intégrant cette automatisation de la phase de mesure de forme initiale, permet d'initialiser la mesure de vibrations en calculant des subsets dans les images initiales à partir d'un cercle défini dans le plan moyen de la surface mesurée. Les résultats obtenus valident l'algorithme développé.

Une technique d'appariement des points de mesure adaptée aux spécificités de l'objet d'étude choisi, à savoir un haut-parleur, étant désormais à disposition, une simplification du montage pseudo stéréoscopique avec un jeu de quatre miroirs a été étudiée, notamment afin d'éliminer les miroirs centraux et les inconvénients associés. Ce travail a abouti à la conception d'un montage simplifié à deux miroirs, présenté dans le chapitre suivant.

Chapitre 8

Montage pseudo stéréoscopique à deux miroirs

« *Ce qui est sain est simple.* »
René Mey

Sommaire

8.1 Introduction **137**
8.2 Principe **138**
8.3 Application à un haut-parleur **138**
8.3.1 Fonctions de réponse en fréquence 138
8.3.2 Déformées modales opérationnelles 139
8.4 Application à la mesure plein champ d'une plaque **140**
8.4.1 Montages 141
8.4.2 Calibrage 141
8.4.3 Fonctions de réponse en fréquence 143
8.4.4 Déformées modales opérationnelles 143
8.5 Conclusion **146**

8.1 Introduction

Les différentes mesures de vibration menées avec le capteur pseudo stéréoscopique à quatre miroirs ont mis en évidence la difficulté de maniabilité du montage. En effet, le prisme central est particulièrement complexe à régler, d'autant qu'il est nécessaire de le positionner très près de la caméra. De plus, comme évoqué en **section 6.2.3**, une forte contrainte est imposée par ce miroir central sur le calibrage du système. En conséquence, un des objectifs du travail présenté dans ce manuscrit a consisté à tenter une simplification du montage afin d'en faciliter l'utilisation. Le réagencement des éléments optiques a abouti à la suppression du miroir central.

Ce chapitre présente le montage pseudo stéréoscopique à deux miroirs réalisé. La **section 8.2** en décrit le principe et les **sections** 8.3 et 8.4 montrent les résultats obtenus respectivement sur un haut-parleur et sur une plaque, afin d'obtenir une validation pour un seul point et une validation plein champ de la méthode proposée. Les fonctions de réponse en fréquence et déformées modales opérationnelles sont affichées dans les deux cas.
Une phase expérimentale ayant été menée sur les montages présentés dans ce manuscrit afin d'effectuer une comparaison des résultats obtenus et d'évaluer les méthodes proposées les unes par rapport aux autres, notons que les données sur les montages, les protocoles et le calibrage des systèmes sont données dans ce chapitre, de façon transversale, en **section 8.4.1**.

8.2 Principe

La **figure 8.1**, à *gauche*, présente le schéma de base du montage, comparé au montage classique affiché sur la **figure 8.1** à *droite*.

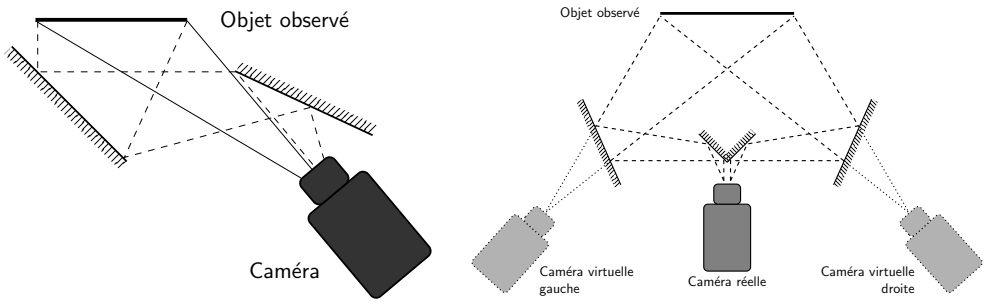


FIGURE 8.1 – Schémas des montages pseudo stéréoscopiques ; (*gauche*) avec jeu de deux miroirs et (*droite*) avec jeu de quatre miroirs

On constate que les deux miroirs sont orientés de manière que la caméra ultra-rapide soit placée latéralement par rapport à l'objet à mesurer. Ceci permet d'obtenir facilement un point de vue réel sensible au déplacement hors plan. L'autre point de vue, virtuel, est généré par les deux miroirs et est situé approximativement en face de l'objet : il est ainsi sensible aux déplacements dans le plan moyen de l'objet. Finalement, l'angle panoramique est réduit de moitié, ce qui simplifie la phase d'appariement initiale tout en permettant de conserver une vue adaptée à la mesure de déplacements. Ce montage est ainsi spécifiquement conçu pour la mesure de vibrations. Il offre une plus grande maniabilité, puisque seulement deux miroirs latéraux sont à orienter, avec un calibrage du système facilité. La zone d'ombre centrale présente dans les images obtenues avec le montage classique à quatre miroirs est réduite.

8.3 Application à un haut-parleur

Le montage à deux miroirs a tout d'abord été testé sur le haut-parleur AB Sound TW1041 choisi précédemment comme objet d'étude, la méthode IRIs (*cf.* **chapitre 7**) étant utilisée afin d'initialiser le processus d'appariement. Le montage expérimental est présenté sur la **figure 8.2** ; la **figure 8.3** montre un exemple d'image vue à travers le système. Les résultats ont été comparés à ceux obtenus avec un vibromètre laser pour un point de mesure, et à ceux obtenus avec le montage classique à quatre miroirs pour la mesure ponctuelle mais aussi plein champ, afin d'obtenir une première validation du protocole.

8.3.1 Fonctions de réponse en fréquence

Les fonctions de réponse en fréquence du point de mesure situé au centre du haut-parleur sont calculées à partir du signal d'excitation appliqué aux bornes et à partir des signaux de déplacement mesurés avec les montages à deux et quatre miroirs, ainsi qu'avec le vibromètre laser. Les résultats des méthodes de vison sont affichés sur la **figure 8.4**, en *rouge* pour le montage à quatre miroirs et en *vert* pour le montage à deux miroirs, et concordent bien visuellement avec ceux obtenus avec la technique de référence. Les résultats obtenus avec le montage à deux miroirs semblent, pour les courbes présentées, aussi précis que ceux du montage classique, avec une différence entre les valeurs mesurées avec la méthode de référence et avec la méthode associée au montage à deux miroirs globalement inférieure à 5% jusqu'à 300 Hz. Pour les fréquences supérieures, la différence est en moyenne de l'ordre de 10%.

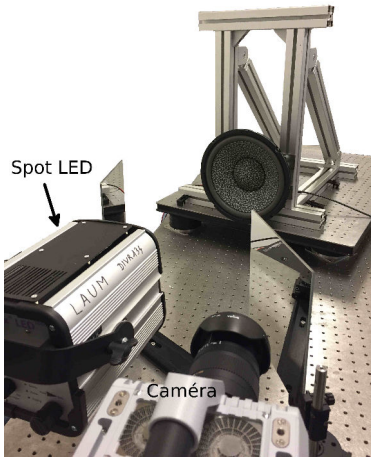


FIGURE 8.2 – Image du montage expérimental

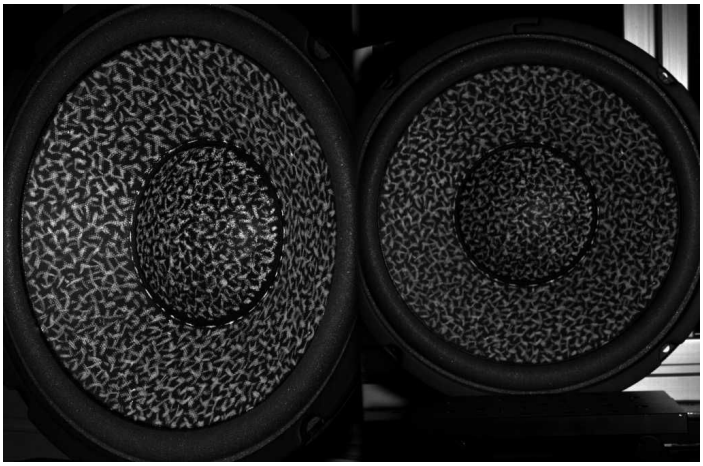


FIGURE 8.3 – Exemple d'image vue à travers le montage pseudo stéréoscopique à deux miroirs

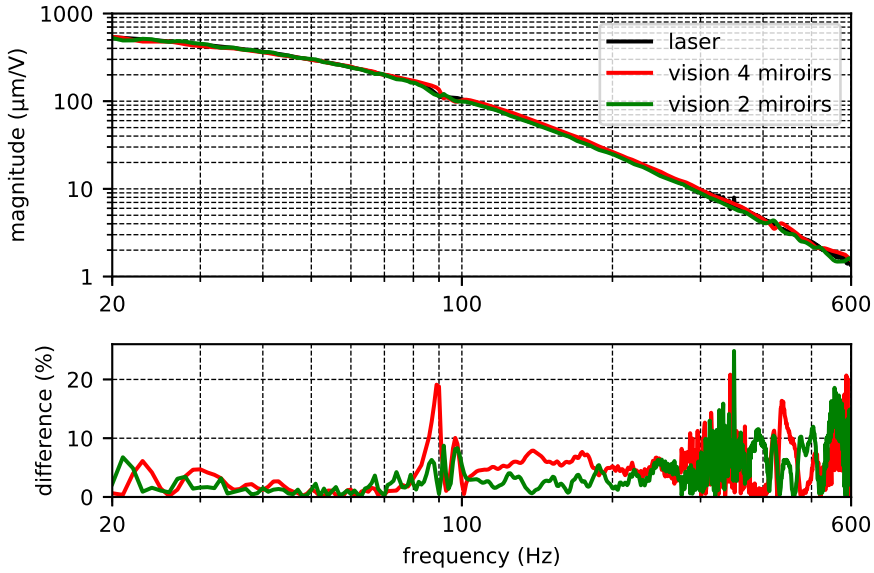


FIGURE 8.4 – FRFs mesurées avec le vibromètre laser et les montages pseudo stéréoscopiques à deux et quatre miroirs

8.3.2 Déformées modales opérationnelles

À partir des FRFs de tous les points, les déformées modales opérationnelles peuvent être tracées. Elles sont affichées sur la **figure 8.5** pour les systèmes pseudo stéréoscopiques à deux (*haut*) et quatre (*bas*) miroirs. Bien qu'un écart de niveau soit à noter à chaque fréquence, de l'ordre de $\approx$ 4,5 µm/V pour le *rocking mode* présenté à *gauche* et de $\approx$ 0,8 µm/V pour le mode de membrane à *droite*, les images correspondent bien visuellement, ce qui valide le montage présenté. Afin de vérifier la pertinence de la méthode et d'obtenir une seconde validation, le montage à deux miroirs a ensuite été appliqué à la mesure de vibrations d'une plaque.

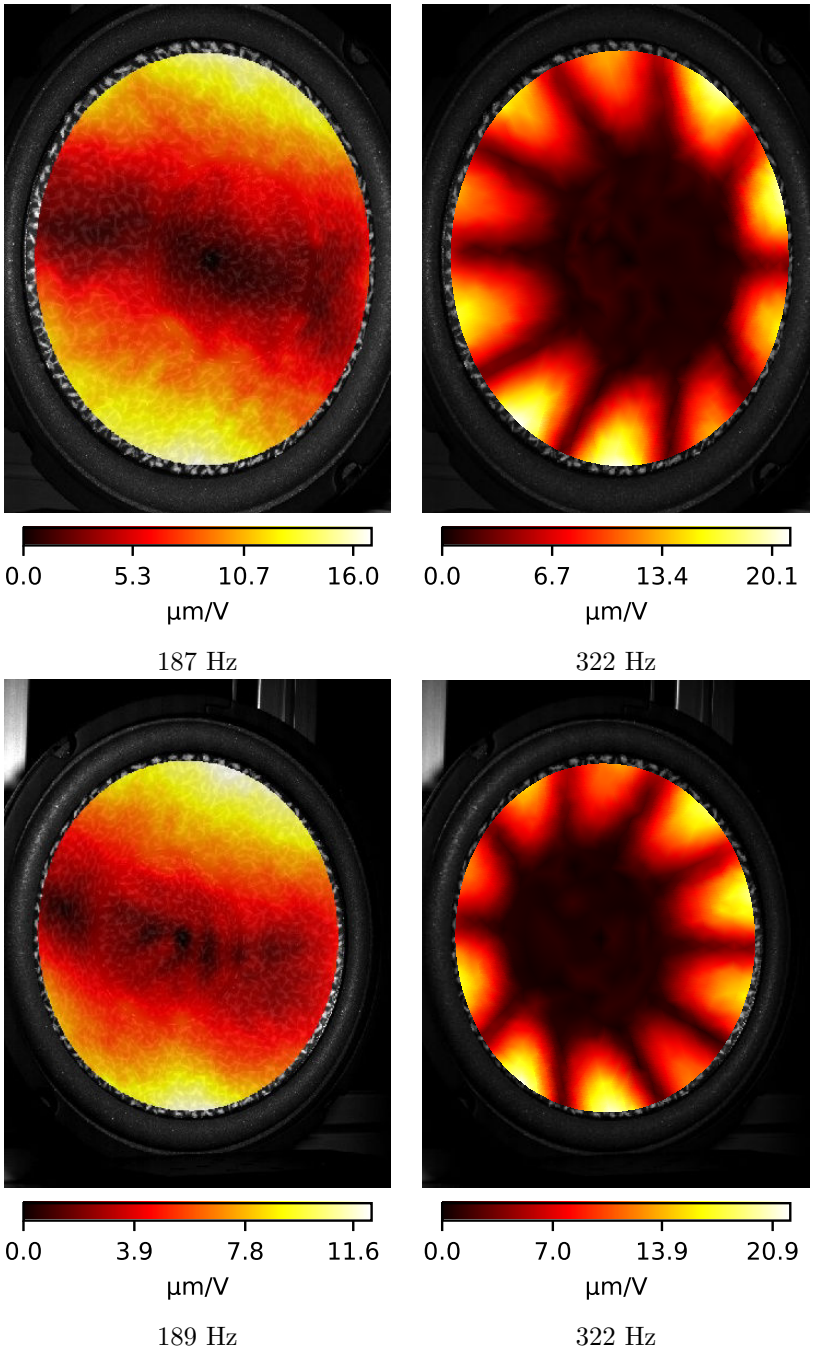


FIGURE 8.5 – Déformées modales opérationnelles obtenues avec le montage à deux miroirs (*haut*) et avec le montage à quatre miroirs (*bas*)

8.4 Application à la mesure plein champ d'une plaque

L'approche a ainsi été testée sur une plaque en vibration afin d'obtenir une validation pour une mesure plein champ par comparaison avec les résultats obtenus avec le montage pseudo stéréoscopique classique avec un jeu de quatre miroirs, considéré ici comme référence. Les autres montages ayant aussi été testés, l'ensemble des données est rapporté dans la section suivante, de façon transversale.

8.4.1 Montages

L'objet d'étude est une plaque en aluminium de 420 mm × 420 mm × 1,5 mm, suspendue à un cadre monté avec des tubes NORCAN (*cf.* **figure 8.6** à *gauche*). Une force est appliquée à la plaque avec un pot vibrant situé à l'arrière de la plaque (*cf.* **figure 8.6** à *droite*), le signal d'excitation étant un sinus glissant, allant de 5 Hz à 1300 Hz en 2,25 secondes. De plus, une tête d'impédance mécanique (capteur de force et accéléromètre) est positionnée entre la plaque et le pot vibrant, afin de calculer des fonctions de réponse en fréquence et d'avoir une référence ponctuelle à chaque mesure. La zone de mesure correspond à la zone centrale carrée de 180 mm × 180 mm. Les photographies du montage

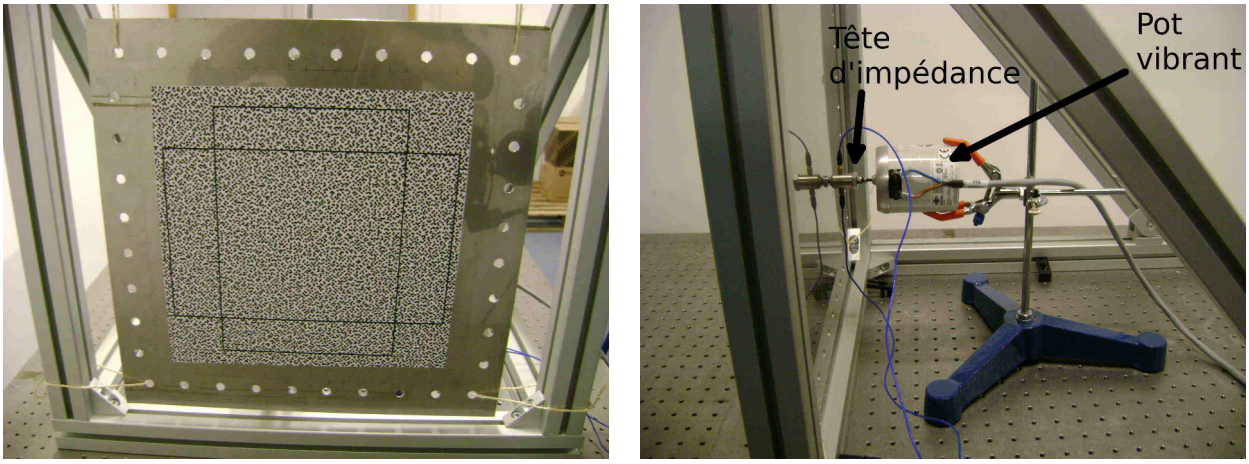


FIGURE 8.6 – (*gauche*) Photographie de la plaque suspendue ; (*droite*) photographie du pot vibrant et de la tête d'impédance

pseudo stéréoscopique classique et du montage à deux miroirs réalisés sont affichées sur la **figure 8.7** respectivement à *gauche* et à *droite*. Notons que la caméra et les miroirs doivent être positionnés à hauteur du centre de la plaque. La caméra ultra-rapide est le modèle Photron SA-X2 du LAUM, placée sur un trépied Manfrotto. La cadence d'acquisition est réglée à 12 500 images par seconde. Le même montage est utilisé avec la méthode à caméra unique sans miroir et la méthode à caméras asynchrones. Une photo est présentée sur la **figure 8.8** à *droite*. Les deux caméras utilisées sont la Phantom v5.1 et la Photron SA-X2 du LAUM, toutes deux montées sur des trépieds Manfrotto. Pour la méthode à caméra unique, le modèle Phantom est seulement utilisé pour la mesure de forme. La séquence temporelle d'images acquises avec la caméra Photron est utilisée pour les deux méthodes. Les cadences d'acquisition sont de 1200 images par seconde pour la caméra Phantom, et de 12 500 images par seconde pour la caméra Photron.

8.4.2 Calibrage

Les paramètres externes des différentes vues sont donnés dans le **tableau 8.1** ci-dessous. Au total, 5408 subsets de mesure sont définis sur la surface de mesure, à des positions proches, *i.e.* à environ un dixième de millimètre, entre le montage pseudo stéréoscopique et le montage à deux caméras. Les distances épipolaires au niveau de la zone de mesure sont de ≈ 0,22 pixel en moyenne avec un écart-type de ≈ 0,20 pixel pour le montage pseudo stéréoscopique avec jeu de quatre miroirs, de ≈ 0,11 pixel en moyenne avec un écart-type de ≈ 0,12 pixel pour le montage à deux miroirs, et de ≈ 0,07 pixel en moyenne avec un écart type de ≈ 0,05 pixel pour le montage à deux caméras. Les formes estimées de la plaque sont affichées sur la **figure 8.9**, à *gauche* pour le montage à quatre miroirs, au *milieu* pour le montage à deux miroirs et à *droite* pour le montage à deux caméras. La forme mesurée avec le montage à deux caméras sert de référence, puisque qu'elle n'est pas obtenue à travers des miroirs et est ainsi

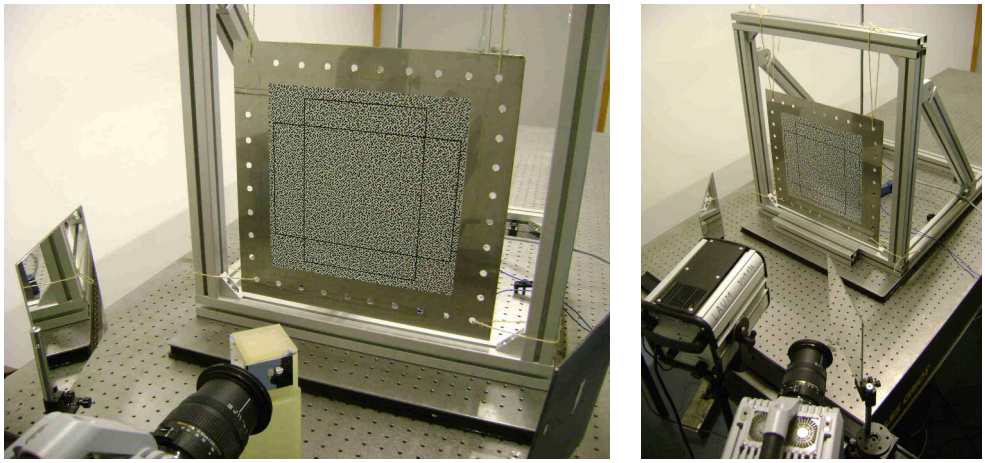


FIGURE 8.7 – Photographie du montage pseudo stéréoscopique avec quatre miroirs (*gauche*) et avec deux miroirs (*droite*)

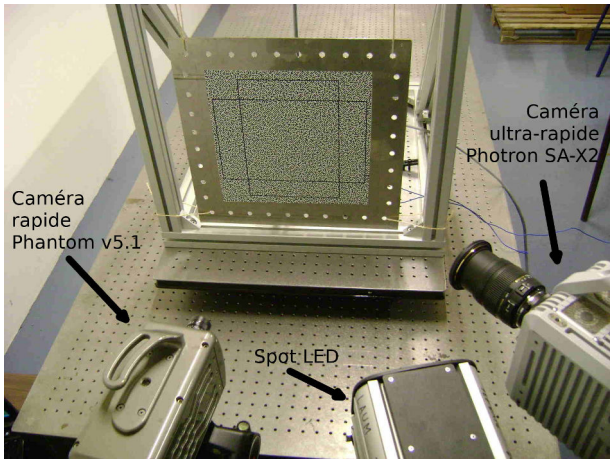


FIGURE 8.8 – Photographie du montage stéréoscopique pour les mesures effectuées avec la méthode sans miroir et avec la méthode à caméras asynchrones

TABLEAU 8.1 – Paramètres externes des différentes vues par rapport à la surface de mesure

		axe de rotation			translation (mm)		
Montage	vue	$\overrightarrow{y}$	$\overrightarrow{x}$	$\overrightarrow{z}$	$\overrightarrow{x}$	$\overrightarrow{y}$	$\overrightarrow{z}$
À quatre miroirs	gauche	-49,9°	-5,3°	8,2°	586	-157	607
	droite	48,9°	-2,4°	-12,3°	-550	-157	629
À deux miroirs	gauche	1,3°	1,9°	0,6°	45	35	1016
	droite	29,9°	1,1°	1,2°	-491	22	584
À deux caméras	gauche	-8,6°	0,9°	0,6°	86	8	468
	droite	40,9°	1,3°	1,4°	-344	10	375

soumise à moins de sources de distorsions et d'erreurs. Les formes calculées avec les montages pseudo stéréoscopiques semblent assez perturbées par la non-planéité des miroirs. Celle-ci est aussi visible sur les cartes correspondantes des distances épipolaires : sur la **figure 8.10**, à *gauche* et au *milieu*, des élévations locales sont notables. Ainsi, pour la mesure de forme, si un montage pseudo stéréoscopique est requis, des miroirs spécifiques faits sur mesure semblent nécessaires, avec des valeurs de tolérance plus faibles en termes de planéité.

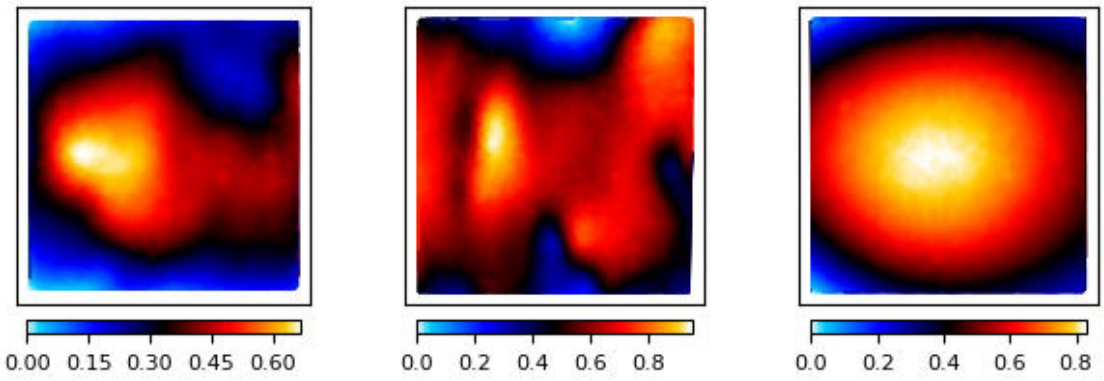


FIGURE 8.9 – Formes calculées (en mm) ; (*gauche*) avec le montage à quatre miroirs ; (*milieu*) avec le montage à deux miroirs ; (*droite*) avec le montage à deux caméras

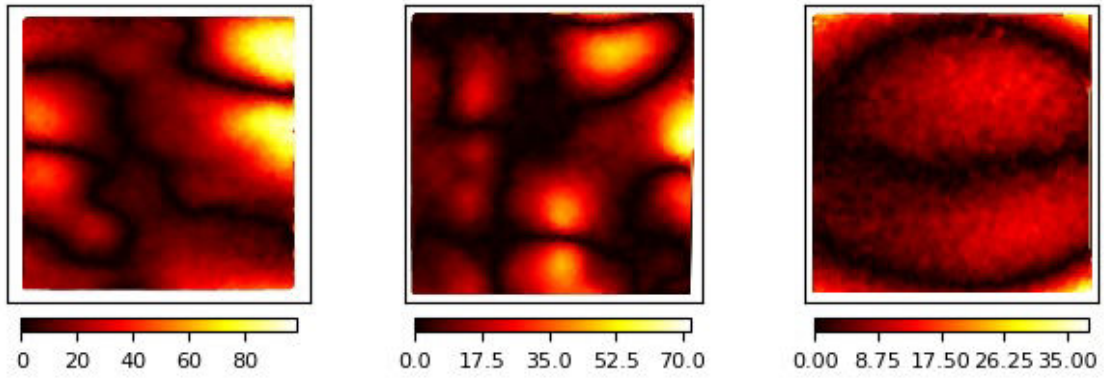


FIGURE 8.10 – Cartes des distances épipolaires (en pixel/100) ; (*gauche*) avec le montage à quatre miroirs ; (*milieu*) avec le montage à deux miroirs ; (*droite*) avec le montage à deux caméras

8.4.3 Fonctions de réponse en fréquence

La tête d'impédance, positionnée entre le pot vibrant et la plaque, est utilisée pour mesurer un signal de force et d'accélération en un point. Le signal de force permet de calculer des fonctions de réponse en fréquence et de comparer les différents montages, et le signal d'accélération permet d'avoir une référence ponctuelle. La position du point de mesure, situé approximativement au niveau de l'accéléromètre, est affichée sur la **figure 8.11**.

Les fonctions de réponse en fréquence obtenues, avec l'accéléromètre, en *rouge*, et avec le montage à deux miroirs, en *noir*, sont affichées sur la **figure 8.12**. Malgré des différences pour les fréquences inférieures à 100 Hz environ, dues principalement à la plus faible sensibilité des accéléromètres en basses fréquences, on note une bonne concordance visuelle. Une comparaison chiffrée ne semble pas pertinente puisqu'il est impossible de parfaitement aligner la position du subset de mesure et de l'accéléromètre situé à l'arrière de la plaque. De plus, le subset de mesure correspond à un disque de 6,6 mm de diamètre sur la plaque, alors que l'accéléromètre est vissé sur une pastille de 20 mm de diamètre : l'intégration spatiale est donc différente.

8.4.4 Déformées modales opérationnelles

Afin de valider les résultats obtenus sur toute la surface de mesure, des déformées modales opérationnelles mesurées avec le montage présenté sont comparées avec celles obtenues avec le montage

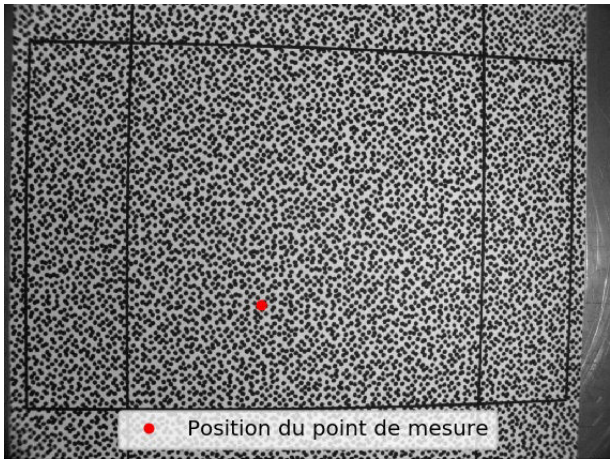


FIGURE 8.11 – Image de la zone de mesure (carré central de 180 mm × 180 mm) et position du point de mesure correspondant à l'accéléromètre

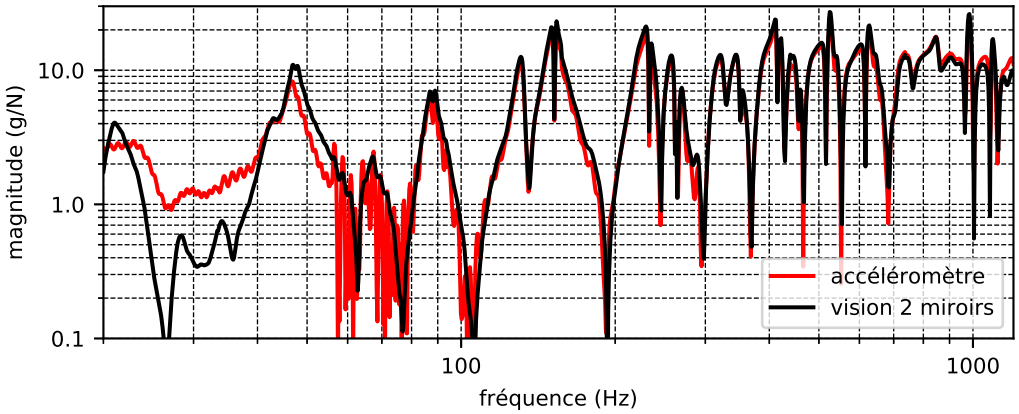


FIGURE 8.12 – Fonctions de réponse en fréquence mesurées pour un point avec un accéléromètre et avec le montage à deux miroirs

pseudo stéréoscopique classique, la rotation de la plaque étant calculée et corrigée dans les deux cas. La **figure 8.13** présente les déformées à 222 Hz, 594 Hz, 888 Hz et 1147 Hz, à *droite* pour le montage à deux miroirs et à *gauche* pour le montage pseudo stéréoscopique classique. Les coefficients de corrélation correspondants sont compris entre 99,4% et 99,9%, ce qui valide les résultats plein champ de la méthode à deux miroirs.

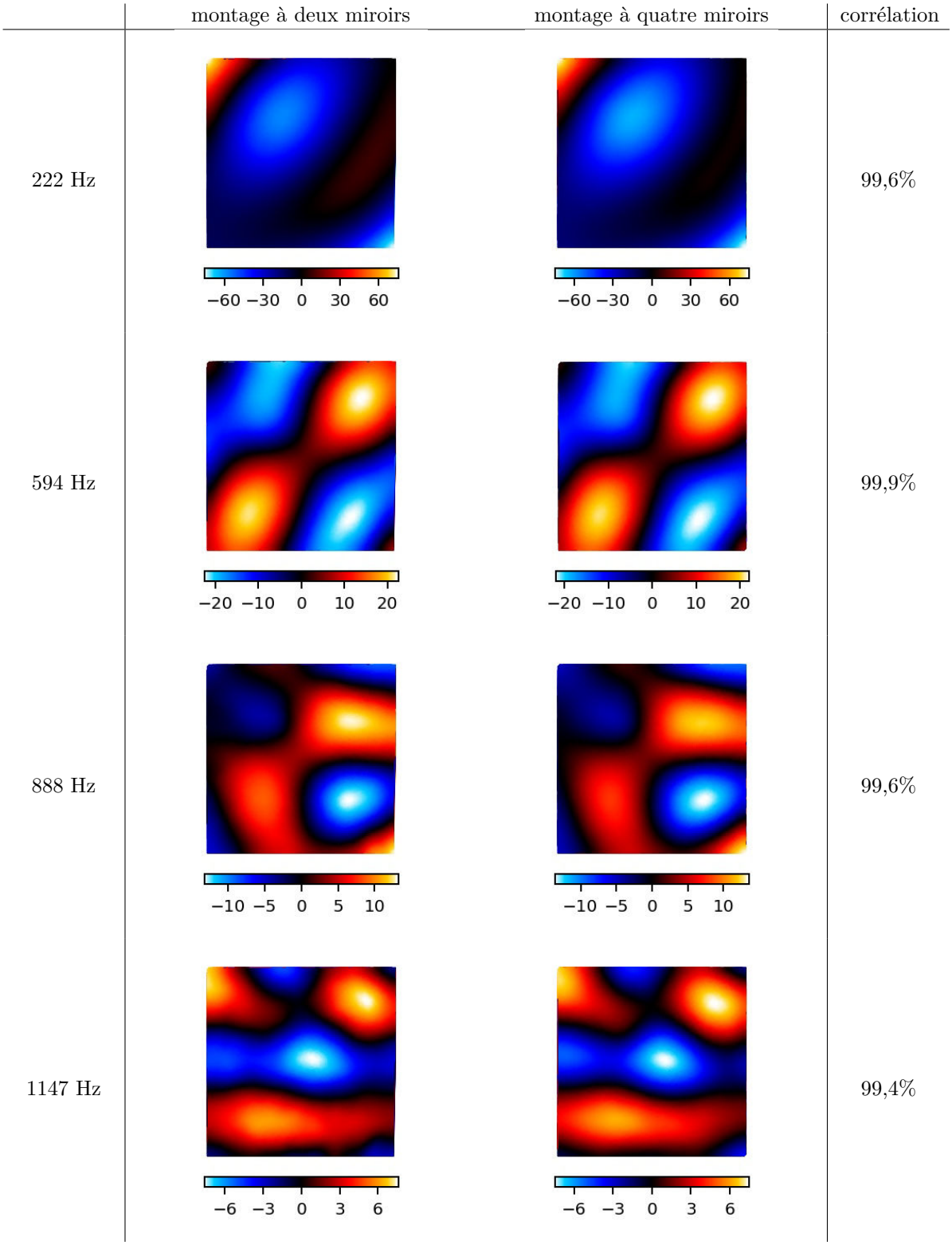


FIGURE 8.13 – Exemples de déformées modales opérationnelles, en µm/N, calculées avec le montage à deux miroirs (à *gauche*) et le montage pseudo stéréoscopique classique (à *droite*) ; de *haut* en *bas* à 222 Hz, 594 Hz, 888 Hz et 1147 Hz

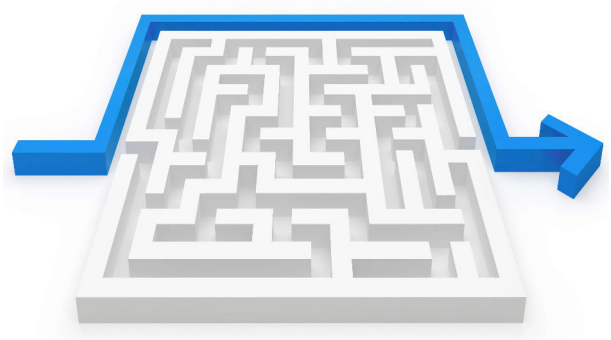
8.5 Conclusion

En conclusion, l'axe de recherche ciblant la simplification du montage pseudo stéréoscopique classique avec un jeu de quatre miroirs a abouti à la suppression du prisme central et ainsi au montage pseudo stéréoscopique avec un jeu de deux miroirs. Les résultats obtenus pour la mesure de déplacement du point central d'un haut-parleur attestent de la validité de la méthode, puisque la différence entre les valeurs obtenues avec le montage à deux miroirs et avec la méthode de référence, ici un vibromètre laser, est globalement inférieure à 5% jusqu'à 300 Hz et environ égale à 10% de 300 Hz à 600 Hz. Les deux montages pseudo stéréoscopiques permettent de retrouver des déformées modales opérationnelles similaires, avec des coefficients de corrélation supérieurs à 99%. Le montage à deux miroirs est beaucoup plus maniable, et la suppression du prisme central ne semble pas impacter la précision de la mesure de vibration dans les tests effectués. Il permet, tout comme le montage pseudo stéréoscopique à quatre miroirs, de mesurer des vibrations multi-axiales ; la forme mesurée est, dans les deux cas, perturbée par la non-planéité des miroirs.

Finalement, on peut arguer que, généralement, les objets vibrent majoritairement selon un seul axe, celui de la normale locale de la surface : une méthode de reconstruction 3D avec un montage pseudo stéréoscopique à quatre ou deux miroirs n'est donc pas forcément nécessaire. Cette hypothèse a été utilisée afin de simplifier le protocole et de concevoir une méthode sans miroir, présentée dans le chapitre suivant, visant à mesurer les vibrations mono-axiales.

Chapitre 9

Méthode de mesure de vibrations mono-axiales



« Ils ne savaient pas que c'était impossible, alors ils l'ont fait. »
Mark Twain

Sommaire

9.1 Introduction **147**
9.2 Principe du montage sans miroir **148**
9.3 Application à un tambourin **149**
9.3.1 Montage et matériel 149
9.3.2 Protocole 149
9.3.3 Calibrage 150
9.3.4 Mesure de forme 150
9.3.5 Validation sur un point avec un laser 151
9.3.6 Mesure plein champ 152
9.3.7 En bref... 153
9.4 Application à la mesure plein champ d'une plaque **154**
9.5 Application à la mesure plein champ d'un haut-parleur **154**
9.6 Conclusion **157**

9.1 Introduction

L'utilisation d'un jeu de miroirs dans les systèmes pseudo stéréoscopiques est peu pratique, même en passant de quatre à deux miroirs, et limite le nombre de pixels disponibles par vue pour toutes les implémentations réalisées dans ce travail. En analysant les spécificités de la mesure de vibration, on note que, généralement, les objets vibrent majoritairement selon un seul axe, celui de la normale locale de la surface [1]. Ainsi, les vibromètres laser à balayage utilisent cette hypothèse pour donner une valeur de vibration normale à la surface de l'objet mesuré à partir d'une mesure effectuée avec un certain angle. La reconstruction 3D que permettent les montages pseudo stéréoscopiques présentés dans les chapitres précédents n'est donc plus nécessaire. En conséquence, cette hypothèse a été utilisée comme base de travail afin de concevoir une méthode de vision 3D plein champ, utilisant une seule

1. Cette hypothèse est valide si le mouvement de corps rigide est négligeable, lorsque l'objet d'étude est encastré par exemple.

caméra ultra-rapide et aucun miroir, adaptée à la mesure de vibration d'objets d'étude affichant des déplacements exclusivement mono-axiaux localement.

Ce chapitre présente tout d'abord le principe du montage (**section 9.2**). La méthode est ensuite appliquée à un objet respectant la contrainte posée, *i.e.* vibrant globalement selon un seul axe : un tambourin (**section 9.3**). La méthode est validée pour une mesure sur un point, par comparaison avec les résultats obtenus avec un vibromètre laser. Dans un deuxième temps, la méthode a été appliquée à la mesure plein champ d'une plaque (**section 9.4**) et validée par comparaison des résultats avec ceux obtenus avec le montage pseudo stéréoscopique à quatre miroirs, considéré ici comme référence. Finalement la méthode a été testée sur un haut-parleur, afin d'étudier son éventuelle pertinence et ses limites pour l'obtention de données initiales sur un comportement dynamique multi-axial (**section 9.5**).

9.2 Principe du montage sans miroir

Diverses méthodes ont déjà été développées pour mesurer des déplacements hors-plan avec une seule caméra positionnée en face d'un objet plan [128, 102]. Toutefois, le choix de positionnement face à l'objet d'étude limite fortement la précision de la mesure hors-plan. Le choix s'est porté, pour la méthode proposée dans ce chapitre, sur un positionnement latéral de la caméra.

Le but est ainsi de supprimer les miroirs restants. Ceci implique, concrètement, qu'un seul point de vue est désormais à disposition, et non plus deux, comme cela était le cas pour les montages (pseudo) stéréoscopiques présentés dans les chapitres précédents. Ce choix impacte particulièrement la phase de triangulation. En effet, comme rappelé en **section 3.3.3**, la triangulation de la position 3D d'un point est fondamentalement réalisée avec deux droites dans l'espace, habituellement deux rayons perspectifs d'un système stéréoscopique calibré. Toutefois, pour un déplacement selon un axe connu, l'une des deux droites peut être calculée mathématiquement. On peut alors mesurer des vibrations mono-axiales avec d'une part, cette droite, qui remplace le premier point de vue, et d'autre part un second point de vue obtenu directement avec une seule caméra rapide positionnée stratégiquement. Il est à noter que la mesure de forme initiale, quant à elle, requiert toujours deux points de vue. À cette fin, une caméra standard est ponctuellement utilisée en plus de la caméra ultra-rapide, mais elle n'est pas conservée lors de la phase de mesure de vibrations.

La **figure 9.1** montre le schéma du montage correspondant à ce principe. La mesure de forme initiale est réalisée à partir de deux points de vue : le premier provient d'une caméra simple (ici un appareil photo, noté première caméra sur le schéma) et le second est obtenu par la caméra ultra-rapide (notée deuxième caméra sur le schéma). La forme initiale mesurée sert à calculer, au niveau de chaque

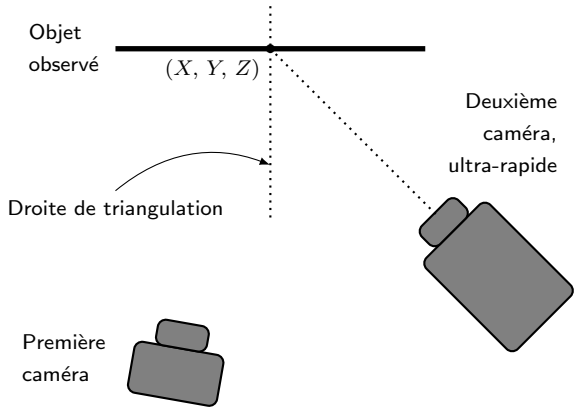


FIGURE 9.1 – Schéma de principe du montage sans miroirs

point de mesure, la normale locale à la surface mesurée, utilisée comme droite de triangulation en

remplacement du premier point de vue. Le rayon perspectif est ensuite calculé pour chaque point de mesure à chaque image (*i.e.* le deuxième point de vue), et la mesure de vibration est effectuée avec la caméra ultra-rapide uniquement.

L'équation (3.38) utilisée pour la triangulation de la position d'un point est alors reprise afin de l'adapter au nouveau contexte :

$$(ku_{1\to 2}, kv_{1\to 2}, k)^T = K_2 R_2 (R_1^T K_1^{-1} \mathrm{x} * z_{cam} + T2 - T1)$$

avec $\mathrm{x} = (u_1, v_1, 1)^T$. En l'analysant, on note que $(R_1^T K_1^{-1}\mathrm{x} - T1)$ correspond à une droite de vecteur $(R_1^T K_1^{-1}\mathrm{x})$ passant par $(-T1)$ le centre du repère caméra 1. Cette droite est remplacée par le vecteur normal à la surface v_n au point $(X, Y, Z)^T$. L'équation (3.38) devient ainsi :

$$(ku_{1\to 2}, kv_{1\to 2}, k)^T = K_2 R_2 (v_n \delta_n + (X, Y, Z)^T + T_2) \tag{9.1}$$

avec v_n le vecteur de la ligne de triangulation dans le repère monde et δ_n la valeur du déplacement. Les bases de la méthode étant posées, un test de validation a été effectué sur un point de mesure.

9.3 Application à un tambourin

La méthode est fondamentalement conçue pour mesurer les vibrations mono-axiales d'objets plans, affichant des déplacements selon la normale de la surface. L'objet d'étude est choisi selon ce critère : le test est mené sur un tambourin. La validation est effectuée en comparant les mesures obtenues simultanément, sur un point, à l'aide de la méthode proposée et d'un vibromètre laser.

9.3.1 Montage et matériel

L'objet d'étude est un tambourin de 300 mm de large, sur la membrane duquel a été dessiné avec un feutre noir un motif aléatoire constitué de traits et de lignes, avec une largeur de tracé de ≈ 1 mm à 2 mm, correspondant à ≈ 2 à 4 pixels sur l'image (*cf.* **figure 9.3**). Une photographie du montage est présentée sur la **figure 9.2**. La caméra ultra-rapide utilisée est le modèle Photron SA-X2 du LAUM, positionnée à environ 40° de la normale de la membrane du tambourin. La caméra du LAUM étant plus récente que celle de l'IRT Jules Vernes, l'option *quiet fan* est intégrée et permet d'arrêter les ventilateurs. Un appareil photo Pentax k-5, avec des pixels de ≈ 4,77 µm de large, est utilisé pour la mesure de forme initiale et placé en face du tambourin. L'excitation est réalisée à l'aide d'une tête d'impact, montée sur un ressort et propulsée par un électro-aimant (*cf.* **figure 9.4**). Le système d'excitation est positionné à l'arrière du tambourin et légèrement excentré (*cf.* **figure 9.3**) afin de fournir de l'énergie aux modes radiaux. En parallèle de la mesure avec le système de vision, la vibration d'un point de contrôle est enregistrée avec un vibromètre laser Polytec.

9.3.2 Protocole

Comme préconisé en **section 3.3.2**, l'appareil photo et la caméra ultra-rapide sont calibrés indépendamment avant de calibrer le système stéréoscopique. Une image par caméra est enregistrée lorsque l'objet est au repos afin de réaliser une mesure de forme initiale à l'aide des outils de corrélation d'images (*cf.* **section 4.6**) et de la méthode IRIs (*cf.* **chapitre 7**). La définition des images est de 896 × 1024 pour la caméra Photron et 720 × 1280 pour l'appareil photo Pentax. Finalement, la vidéo est enregistrée par la caméra Photron, dont le déclencheur contrôle aussi l'électro-aimant de la tête d'impact afin de minimiser le nombre d'images de l'objet au repos. Les subsets sont définis comme des disques de 6 mm de diamètre dans le plan moyen de la surface observée (*cf.* **figure 9.3**).

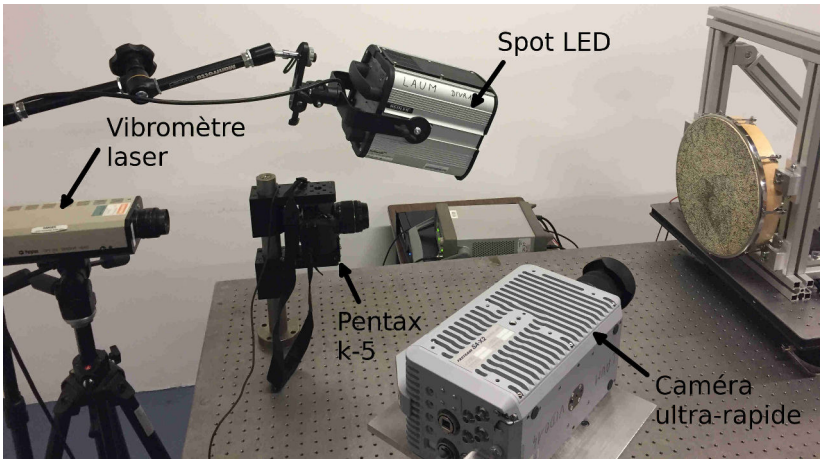


FIGURE 9.2 – Photographie du montage réalisé pour mesurer le tambourin

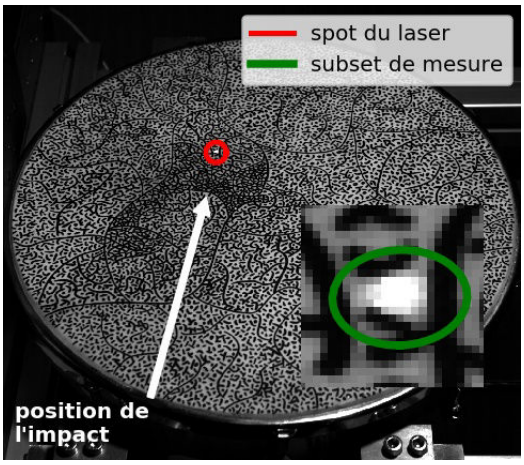


FIGURE 9.3 – Image du tambourin obtenue avec la caméra ultra-rapide, position du spot laser et de l'impact

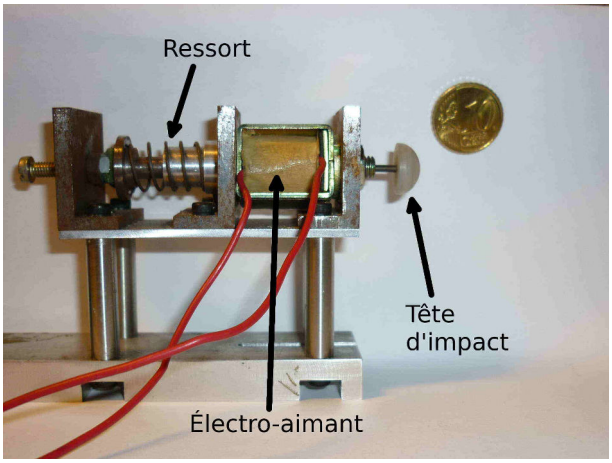


FIGURE 9.4 – Photographie de la tête d'impact, montée sur un ressort et propulsée par un électro-aimant

9.3.3 Calibrage

Le calibrage des caméras est réalisé avec une mire carrée constituée de 12 × 12 cibles, espacées de ≈ 22 mm. Les distances focales obtenues sont de 2274 pixels (≈ 45,5 mm) pour la caméra Photron, et de 1790 pixels (≈ 8,5 mm) pour l'appareil photo Pentax.

Le calibrage du système stéréoscopique est ensuite réalisé à l'aide d'une mire en spirale (*cf.* **section 5.3**) de ≈ 270 mm de diamètre pour le cercle externe, avec 340 cibles. L'angle estimé entre le repère de la mire, approximativement aligné avec la membrane du tambourin, et le plan image de la caméra ultra-rapide est de ≈ 41,8°. Les distances épipolaires obtenues au niveau de la mire, avec cinq images de la mire par caméra, sont de ≈ 0,13 pixel en moyenne, avec un maximum de ≈ 0,55 pixel.

9.3.4 Mesure de forme

La mesure de forme initiale est réalisée avec la méthode IRIs (*cf.* **chapitre 7**) avec une définition de 1024 × 1024 pixels pour les ortho-images. La forme mesurée et les distances épipolaires associées sont affichées sur la **figure 9.5**. On peut constater que le tambourin est légèrement en torsion sous

la contrainte du support, avec des distances épipolaires de ≈ 0,04 pixel en moyenne, et un écart-type de ≈ 0,05 pixel. Les valeurs maximales, allant jusqu'à ≈ 0,46 pixel, sont situées aux bords de la membrane.

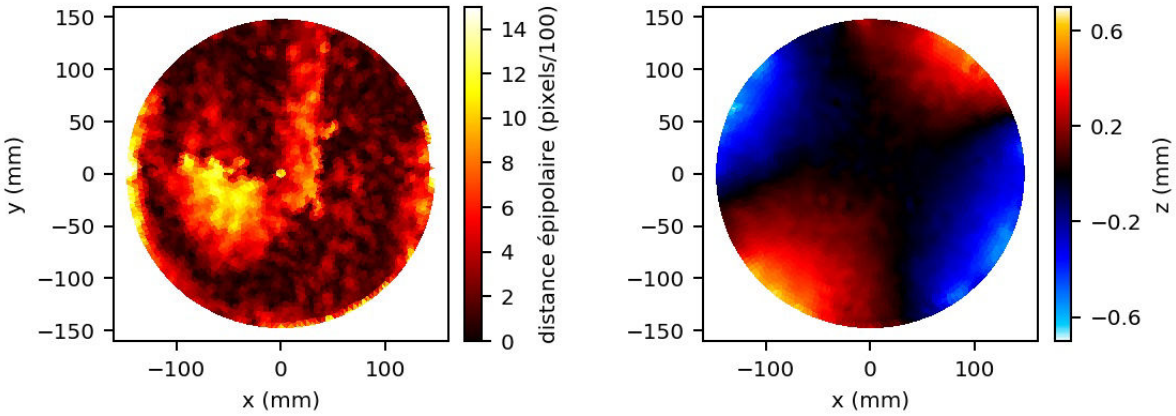


FIGURE 9.5 – Résultats de la mesure de forme ; (*gauche*) distances épipolaires, (*droite*) forme mesurée

9.3.5 Validation sur un point avec un laser

Une mesure de la déformation de la membrane subissant un impact est réalisée simultanément avec la méthode de vision proposée et un vibromètre laser. Pour rappel, le subset de mesure, le spot du laser et la position de l'impact sont affichés sur la **figure 9.3**. La position du point de mesure est choisie arbitrairement, dans une zone de l'image présentant une résolution légèrement inférieure à la valeur moyenne. Les spectres mesurés dans les deux cas sont affichés sur la **figure 9.6**, en *rouge* pour le vibromètre laser et en *noir* pour la méthode de vision. Les courbes sont corrélées à 99,0%, malgré un bruit nettement visible sur la courbe obtenue avec la méthode de vision. Ce bruit semble principalement dû à la saturation induite par le spot du laser. En effet, le spectre mesuré pour un subset adjacent (en *vert* sur la **figure 9.7**) présente un bruit moindre.

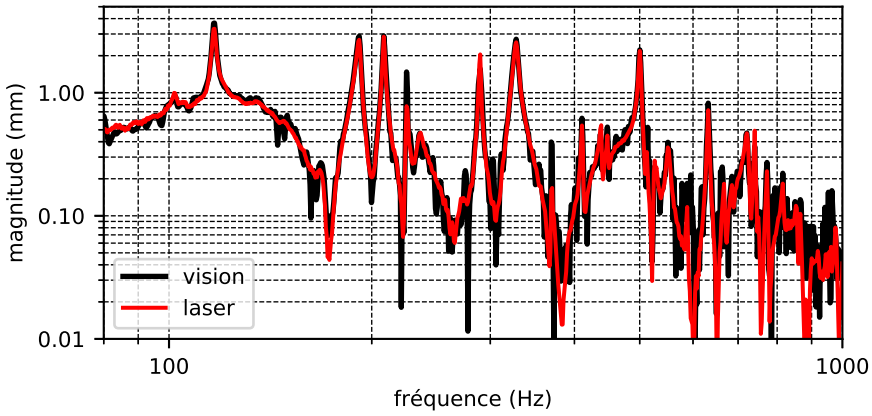


FIGURE 9.6 – Spectres mesurés avec le vibromètre laser et la méthode de vision sur la même zone

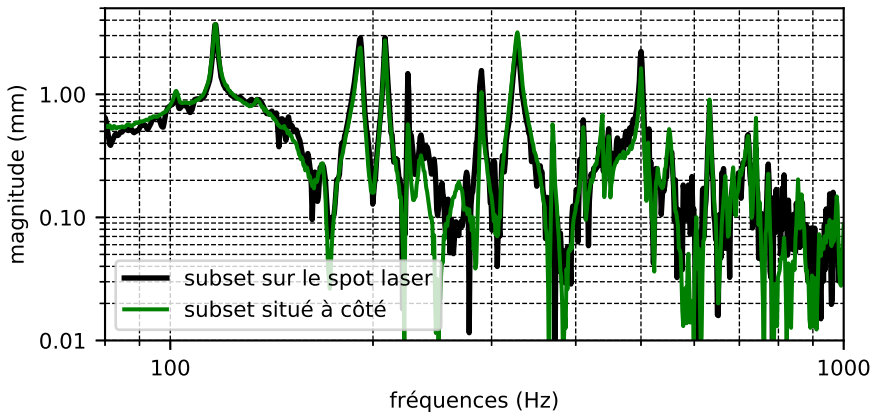


FIGURE 9.7 – Spectres mesurés avec la méthode de vision au niveau du spot laser et d'un subset adjacent

9.3.6 Mesure plein champ

La mesure étant validée pour un point, une vidéo est enregistrée en éteignant le laser. L'évolution de la déformation induite par l'impact est suivie sur ≈ 3000 points de mesure. La **figure 9.8** présente neuf clichés, tous les 0,7 ms, correspondant au début de l'impact. Le déplacement au niveau de l'impact induit une légère rotation de la membrane qui se propage et est réfléchie sur les bords du tambourin. L'écart-type dans le domaine temporel calculé avant l'impact, pour chaque point de mesure, est affiché

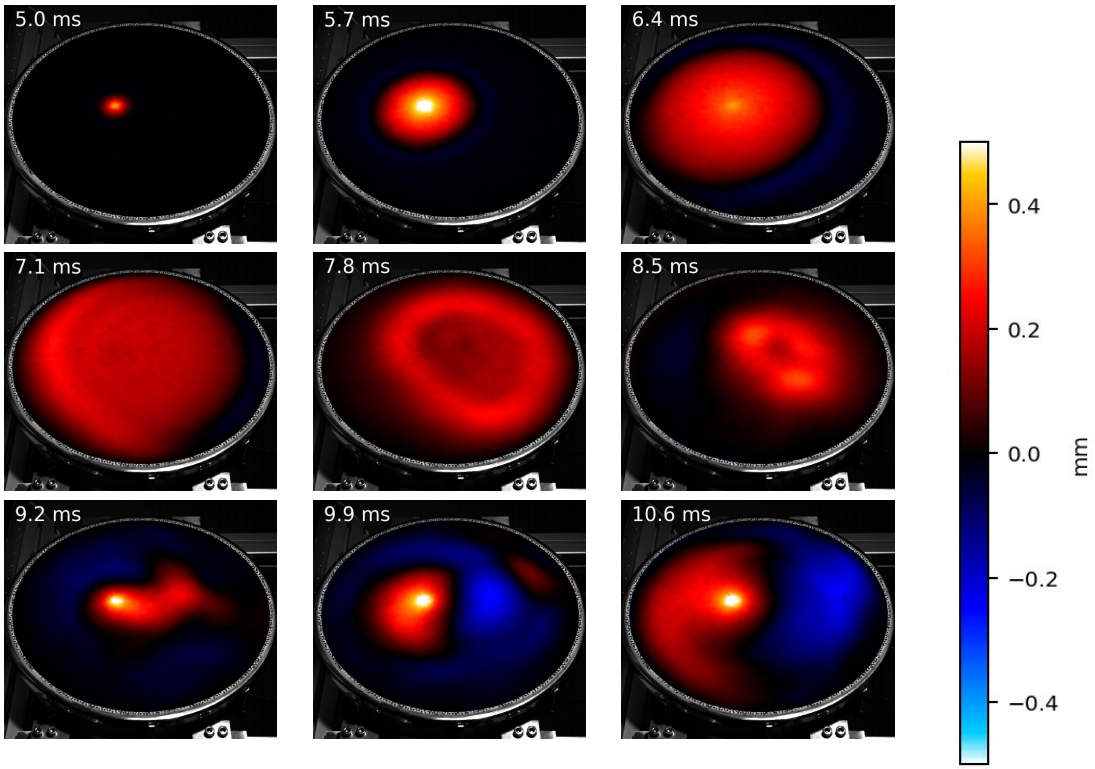


FIGURE 9.8 – Évolution temporelle de l'impact (*i.e.* déplacement de la membrane)

sur la **figure 9.9**. On note des valeurs plus élevées sur les bords du tambourin, en haut et en bas de l'image. L'écart-type moyen est de ≈ 1,1 µm, avec un maximum de ≈ 2,4 µm.

FIGURE 9.9 – Carte de l'écart-type du déplacement mesuré dans le domaine temporel

Comme précédemment, les fréquences des déformées modales opérationnelles sont obtenues à partir de l'écart-type de l'auto-spectre des signaux de déplacement (*cf.* **section 1.3.2**). La transformée de Fourier du signal de chaque point de mesure est alors calculée aux fréquences d'étude, afin d'afficher les déformées. Sont tracés, sur la **figure 9.10**, des exemples à 361 Hz, 412 Hz, 768 Hz et 858 Hz.

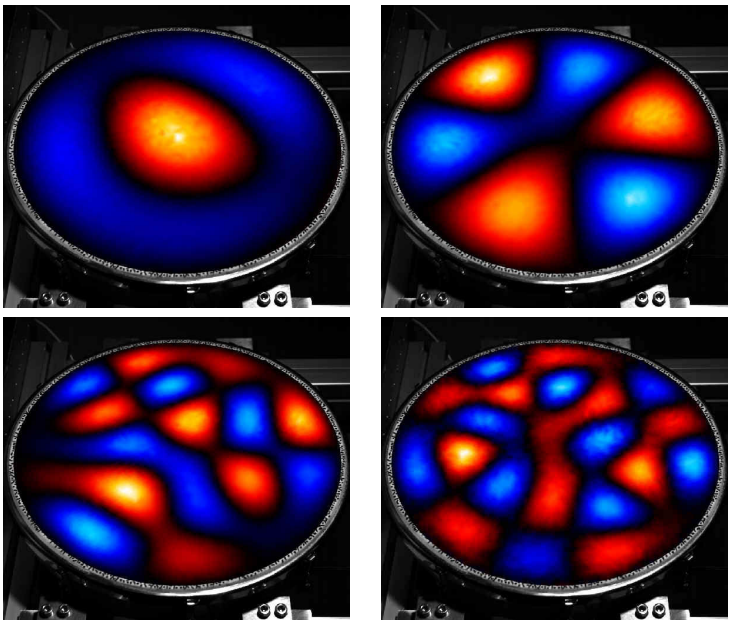


FIGURE 9.10 – Exemples de déformées modales ; (*haut gauche*) à 361 Hz, (*haut droite*) à 412 Hz, (*bas gauche*) à 768 Hz et (*bas droite*) à 858 Hz

9.3.7 En bref...

La comparaison des spectres calculés pour les signaux mesurés avec la méthode de vision sans miroir et un vibromètre laser valide l'approche présentée, permettant la mesure plein champ d'un déplacement hors-plan non-stationnaire, généré par un impact, sur un disque de 300 mm de diamètre, avec une unique caméra rapide et une amplitude de déplacement de l'ordre de $\approx 0{,}5$ mm au maximum.

La méthode de vision mono-axiale décrite dans cette section est fondamentalement adaptée à la mesure de vibration selon un seul axe, pour des objets plans dont le déplacement respecte ce critère. La question de son utilité potentielle pour des objets d'étude plans ou non-plans dont les vibrations sont multi-axiales s'est néanmoins posée, dans la logique d'une première approche de leur comportement dynamique 3D. Afin d'évaluer la pertinence et la limite de son utilisation dans ce contexte plus large, deux tests ont été menés, en cohérence avec le travail effectué précédemment : le premier sur une plaque pour l'aspect vibratoire, le second sur le haut-parleur déjà choisi comme objet d'étude pour l'aspect vibro-acoustique. Les protocoles et les résultats sont présentés dans les sections suivantes (respectivement **sections** 9.4 et 9.5).

9.4 Application à la mesure plein champ d'une plaque

Un premier test a été effectué sur une plaque en vibration, afin d'obtenir une validation plein champ de la méthode. Les informations sur une mesure plein champ ont été comparées aux résultats obtenus avec le montage pseudo stéréoscopique classique avec un jeu de quatre miroirs, considéré ici comme référence. Pour rappel, les mesures ayant été effectuées successivement, sur la même plaque, avec les montages présentés dans ce manuscrit, les données liées au matériel utilisé pour le montage, l'agencement des éléments ainsi qu'au calibrage, ont été données de façon transversale en **section 8.4.1**. Dans cette section, ne seront présentés que les résultats obtenus.

Notons que la plaque est suspendue, *i.e.* non encastrée, et que le point d'excitation est excentré au niveau de la plaque. De fortes rotations et translations sont générées en basses fréquences et forment un mouvement de corps rigide global qui ne peut être pris en compte par la méthode mono-axiale. Afin de comparer des données comparables, les résultats sont affichés avec ceux obtenus pour l'axe $\overrightarrow{z}$ du montage pseudo stéréoscopique classique. La **figure 9.11** présente les déformées modales opérationnelles des deux montages à 222 Hz, 594 Hz, 888 Hz et 1147 Hz. Les coefficients de corrélation étant compris entre 99,2% et 99,9%, la méthode est validée pour la mesure plein champ. Toutefois, à 222 Hz, un écart de niveau, avec un facteur proche de deux, est observé. Cet écart est fort probablement dû à une rotation de la plaque. Pour les fréquences supérieures, à 594 Hz, 888 Hz et 1147 Hz, de par l'inertie de la plaque, la rotation devient négligeable, les niveaux sont donc semblables. Ceci confirme donc la validité de la méthode, et la limite à des mouvement strictement mono-axiaux si l'on souhaite des valeurs chiffrées correctes.

9.5 Application à la mesure plein champ d'un haut-parleur

La méthode sans miroir est également testée sur un haut-parleur, soit un objet dont les déplacements après excitation ne suivent pas un axe unique en fonction des fréquences (*cf.* **chapitre 6**), afin d'étudier sa pertinence et ses limites pour l'obtention de données initiales sur le comportement dynamique correspondant.

Notons que cette phase expérimentale a aussi servi à tester l'utilisation d'autres capteurs, dits grand public, à savoir la caméra GoPro Hero 5 Black et la caméra d'un iPhone 6, dont les cadences d'acquisition de 240 images par seconde sont très nettement inférieures à celle d'une caméra ultra-rapide standard. Les résultats ne sont pas rapportés dans le corps de ce manuscrit mais sont inclus en **Annexe B**. Ils attestent du potentiel de ces caméras *low-cost*, surtout dans la logique actuelle de lancement sur le marché de caméras très performantes, affichant environ 1000 fps pour moins de mille euros, même si, pour l'instant, le traitement d'images automatique et/ou la compression au format .mp4 amènent un bruit assez conséquent sur la mesure.

Les mesures effectuées avec le montage à deux miroirs sont ré-utilisées dans ce contexte, puisque le point de vue de la caméra ulra-rapide, placée latéralement et sensible au déplacement, est dissocié des miroirs et peut ainsi être considéré comme similaire au point de vue de la caméra du montage sans miroir, à un facteur d'échelle près. Les informations obtenues précédemment avec le système pseudo

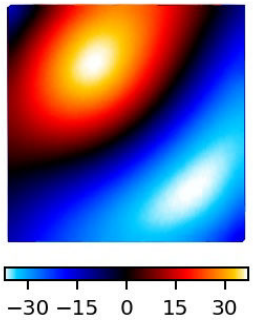
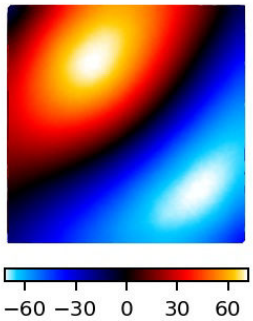
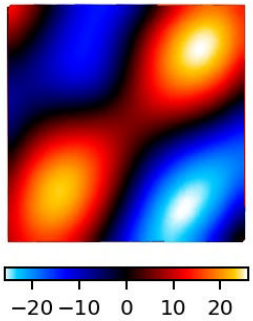
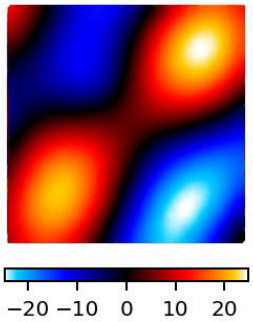
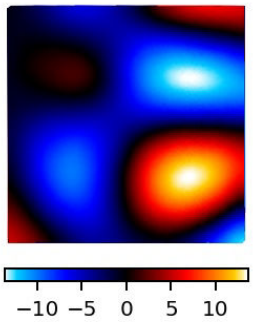
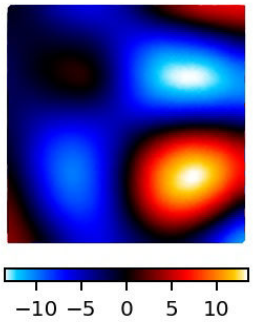
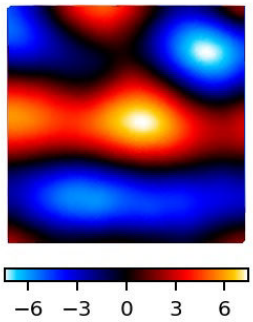
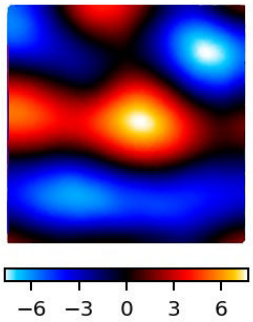
	méthode sans miroir mono-axiale	montage à quatre miroirs (axe $\vec{z}$)	corrélation
222 Hz			99,7%
594 Hz			99,9%
888 Hz			99,4%
1147 Hz			99,2%

FIGURE 9.11 – Déformées modales opérationnelles, en µm/N, calculées avec la méthode de sans miroir mono-axiale (à *gauche*) et le montage pseudo stéréoscopique classique (axe $\vec{z}$) (à *droite*) ; de *haut* en *bas* à 222 Hz, 594 Hz, 888 Hz et 1147 Hz

stéréoscopique avec un jeu de deux miroirs sont ainsi traitées à nouveau, en tenant compte uniquement de la vue réelle. On peut arguer que le montage réglé spécifiquement pour la mesure sans miroir aurait permis de conserver la totalité de la matrice de pixels pour le point de vue sensible au déplacement et non pas la moitié, comme c'est le cas ici en conservant uniquement la vue réelle du montage à deux miroirs. Cet argument est indéniable. Néanmoins, les résultats sont assurément comparables puisque la vibration mesurée et les images d'entrée sont identiques. La comparaison est donc au niveau du traitement et permet bien de mettre en parallèle les résultats obtenus avec les deux montages (à deux miroirs et sans miroir).

Les courbes de l'écart-type des auto-spectres de tous les points de mesure pour les deux cas sont affichées sur la **figure 9.12** pour l'axe $\overrightarrow{z}$, en *rouge* pour le calcul avec les miroirs et en *noir* pour le traitement mono-axial. Notons que, pour cette mesure plein champ, des erreurs sont nécessairement

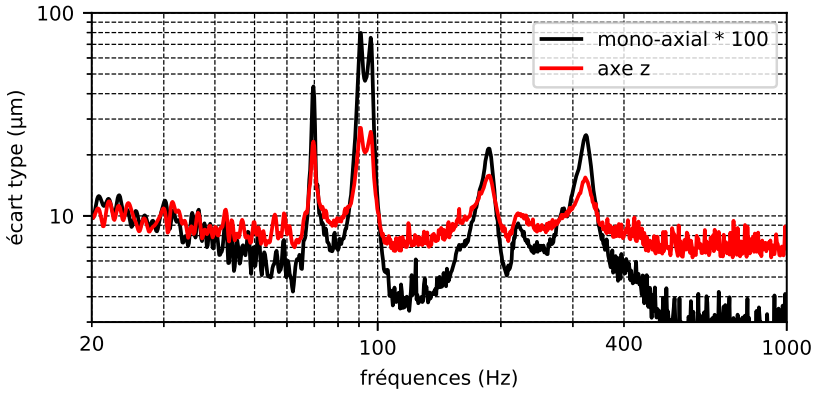


FIGURE 9.12 – Écart-type des transformées de Fourier calculées avec le traitement à deux miroirs et le traitement sans miroir

attendues, puisque les modes mesurés correspondent à des déplacements 3D selon un axe différent de celui de la bobine du haut-parleur (*cf.* **chapitre 6**), que la technique mono-axiale ne peut donc pas totalement imager. Une différence conséquente de niveau est ainsi visible, la courbe *noire* étant multipliée par un facteur 100, mais les pics sont nettement identifiables avec le traitement mono-axial. Des fréquences de résonance sont donc détectables, avec des différences de niveau attendues. Les déformées modales opérationnelles correspondantes sont affichées sur la **figure 9.13**. En regardant les résultats pour un *rocking mode* (à *gauche*), les formes estimées sont visuellement similaires pour les deux traitements. Pour un mode de membrane (à *droite*), la forme calculée sans les miroirs est en revanche clairement dégradée, même si le motif global est reconnaissable.

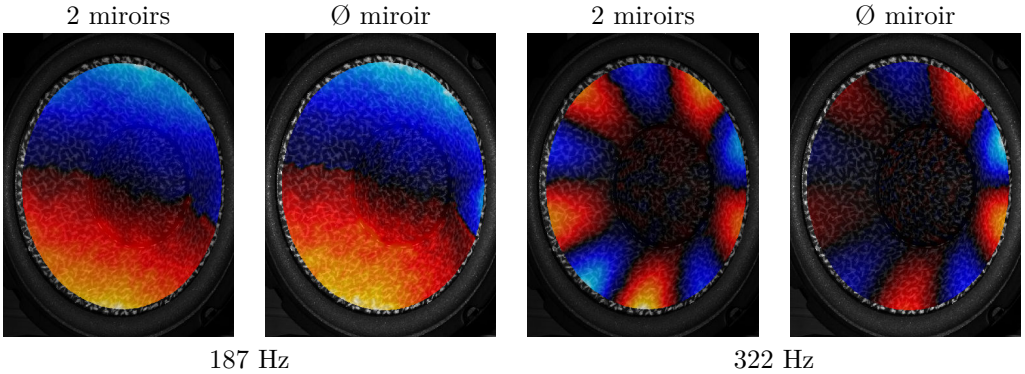


FIGURE 9.13 – Déformées modales opérationnelles calculées avec le traitement à deux miroirs et le traitement sans miroir

9.6 Conclusion

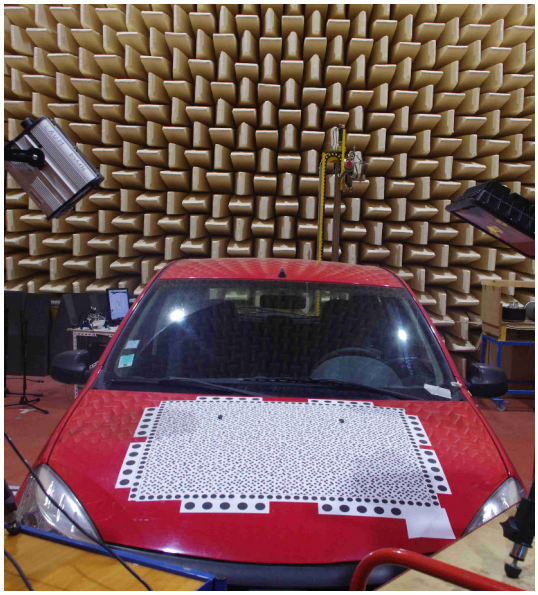
En conclusion, ce chapitre a présenté la méthode de vision plein champ conçue afin de mesurer les vibrations mono-axiales d'objets globalement plans, affichant des déplacements répétables ou non-répétables selon la normale de la surface ou selon un axe connu. Aucun miroir n'est requis : une des vues nécessaires à la triangulation est remplacée par une ligne mathématique. Les résultats obtenus sur un tambourin ont été validés par comparaison avec ceux obtenus pour un point avec un vibromètre laser. Les déformées modales opérationnelles ont aussi pu être calculées et affichées.

De plus, la technique a été testée sur une plaque suspendue, dont le comportement vibratoire est globalement mono-axial après excitation, au delà de 300 Hz environ. Les résultats valident la méthode en tant que technique de mesure plein champ : les coefficients de corrélation calculés à partir des déformées modales opérationnelles obtenues d'une part avec ce montage et d'autre part avec le montage pseudo stéréoscopique classique (axe $\overrightarrow{z}$), considéré ici comme référence, sont supérieurs à 99%. Un second test a été mené sur un haut-parleur, dont les déplacements après excitation ne suivent pas un axe unique en fonction des fréquences, afin d'étudier sa pertinence et ses limites pour l'obtention de données initiales cette fois sur le comportement dynamique multi-axial correspondant. Les résultats obtenus pour cette mesure montrent qu'il est possible de retrouver les fréquences de résonance de l'objet, mais que les déformées modales opérationnelles sont partiellement mal imagées si elles ne sont pas parallèles à l'axe de triangulation. La méthode donne alors accès à un aperçu initial du comportement dynamique de l'objet et permet de circonscrire un protocole d'étude ultérieur. La méthode a aussi pu être utilisée avec succès pour mesurer le déplacement du point central d'un haut-parleur dont le mouvement est mono-axial, *cf.* **Annexe B**, et ainsi d'estimer les paramètres linéaires et un paramètre non-linéaire du haut-parleur.

Finalement, la méthode de vision sans miroir présentée montrant une limite logique pour la restitution des informations correspondant aux déplacements multi-axiaux, le raisonnement a été poussé davantage afin de concevoir une méthode de vision, ne requérant aucun miroir, adaptée cette fois à la mesure de vibrations multi-axiales. Cette technique et le protocole associés sont détaillés dans le chapitre suivant.

Chapitre 10

Méthode à caméras asynchrones



« Our greatest weakness lies in giving up. The most certain way to succeed is always to try just one more time. »
Thomas A. Edison

Sommaire

10.1 Introduction **159**
10.2 Principe **160**
10.3 Application à une plaque **162**
10.3.1 Fonctions de réponse en fréquence 163
10.3.2 Déformées modales opérationnelles 163
10.3.3 En bref... 163
10.4 Application industrielle : cas du capot d'une voiture **165**
10.4.1 Montage et protocole 165
10.4.2 Déformées modales opérationnelles 166
10.4.3 En bref... 166
10.5 Application industrielle : cas du test d'un haut-parleur **168**
10.5.1 Montage et protocole 168
10.5.2 Fonctions de réponse en fréquence du haut-parleur avant dégradation . . . 169
10.5.3 Déformées modales opérationnelles du haut-parleur avant dégradation . . . 169
10.5.4 Résultats après dégradation du haut-parleur 172
10.5.5 En bref... 173
10.6 Conclusion **174**

10.1 Introduction

Les **chapitres 6**, **8** et **9** ont présenté trois méthodes conçues pour mesurer des vibrations avec une unique caméra ultra-rapide et ainsi réduire le coût des montages associés. Les montages pseudo stéréoscopiques décrits (avec jeu de quatre et de deux miroirs) permettent, avec plus ou moins de contraintes en termes de maniabilité, de mesurer avec deux points de vue des déplacements selon les trois axes de l'espace, au détriment du nombre de pixels par vue et donc de la précision de la mesure puisque la surface photosensible est divisée par deux. La méthode sans miroir, quant à elle, fournit avec un montage simplifié et très maniable, des mesures précises de vibrations hors-plan mais elle reste limitée à des déplacements mono-axiaux. L'objectif de cette dernière phase de recherche a consisté à conserver la simplification du montage pseudo stéréoscopique classique obtenue avec le montage sans miroir, tout en élargissant la mesure à des déplacements multi-axiaux.

D'un point de vue métrologique, l'idéal est indubitablement d'utiliser deux caméras ultra-rapides synchronisées. Deux caméras différentes ont été utilisées pour ce travail de recherche, mais leurs plages de disponibilité n'ont jamais coïncidé et n'ont donc pas permis une utilisation simultanée. Dans les phases de recherche décrites précédemment, le travail s'est ainsi logiquement porté sur la conception de méthodes requérant une unique caméra ultra-rapide. Toutefois, le LAUM possède, en plus de la caméra Photron SA-X2, une caméra Phantom v5.1. Notons que, dans ce manuscrit, cette dernière sera dite rapide, puisqu'elle peut enregistrer jusqu'à 1200 images par seconde avec une définition de 1024 × 1024 pixels, tandis que la caméra Photron sera dite ultra-rapide, avec une cadence d'acquisition de 12 500 images par seconde pour la même définition. L'astuce consiste à utiliser ces deux caméras avec des cadences d'acquisition différentes : l'amplitude des vibrations diminuant lorsque la fréquence augmente (*cf.* **chapitre 1**), la caméra rapide permet d'obtenir des informations 3D pour les vibrations allant jusqu'à environ 300 Hz, tandis que la caméra ultra-rapide est utilisée pour acquérir des données allant jusqu'à 1000 Hz environ. La méthode conçue autour de ce principe utilise donc fondamentalement des caméras asynchrones et permet d'effectuer des mesures de vibrations multi-axiales avec une seule caméra ultra-rapide et un montage sans miroir.

Ce chapitre décrit tout d'abord le principe de la méthode (**section 10.2**), ainsi que sa validation, effectuée par le biais d'une mesure plein champ des vibrations d'une plaque (**section 10.3**). Les **sections** (**10.4**) et (**10.5**) présentent finalement les résultats de l'étude de faisabilité menée afin d'évaluer d'une part l'applicabilité de la technique à la mesure de vibrations de plus grands objets (ici un capot d'une voiture) et d'autre part son applicabilité à la détection d'un défaut de suspension d'un haut-parleur, l'objet d'étude étant dans ce cas un haut-parleur volontairement dégradé.

10.2 Principe

A l'issue d'une campagne de mesure et du traitement des données, les informations à disposition sont les positions 2D des subsets dans chaque image des séries temporelles des deux caméras utilisées (*cf.* **chapitre 4**). Les deux caméras constituent un système stéréoscopique, calibré au préalable, ce qui permet de trianguler des positions 3D, à un instant donné, à partir d'une position 2D dans chaque repère image (*cf.* **chapitre 3**). Toutefois, les cadences d'acquisition étant différentes, les deux caméras ne sont nécessairement pas synchronisées et les déplacements de l'objet aux fréquences supérieures à la valeur correspondant à la moitié de la cadence d'acquisition de la caméra rapide engendrent un repliement spectral.

Afin de pallier ces problèmes, deux astuces sont utilisées. Premièrement, les positions 2D des subsets de mesure dans les séries d'images constituent des signaux qui permettent de mesurer le délai entre les déclencheurs des deux caméras. Les signaux de la caméra rapide peuvent ainsi être interpolés afin d'effectuer la triangulation des positions 3D pour chaque image de la caméra ultra-rapide. Deuxièmement, puisque l'amplitude des vibrations décroit lorsque la fréquence augmente et que l'amplitude du déplacement des subsets dans les images est liée à l'angle entre la caméra et l'axe de déplacement (*cf.* **chapitre 6**), la valeur choisie pour l'angle entre la caméra rapide et l'axe principal de déplacement est délibérément faible, environ 10° afin d'éviter le repliement spectral. La caméra utra-rapide est positionnée de telle sorte que l'angle entre la caméra ultra-rapide et l'axe principal de déplacement soit d'environ 40° [1] afin d'avoir une bonne sensibilité pour la mesure de vibrations.

La triangulation est finalement effectuée en prenant comme référence les rayons perspectifs de la caméra ultra-rapide. Un exemple d'agencement du montage est présenté sur la **figure 10.1**. La caméra ultra-rapide, en *noir*, est placée avec un angle de ≈ 40° par rapport à la normale du plan moyen de l'objet, et la caméra rapide, en *rouge*, avec un angle de -10°.

1. Comme dans les chapitres précédents, cette valeur dépend de l'objet d'étude et des possibilités d'agencement.

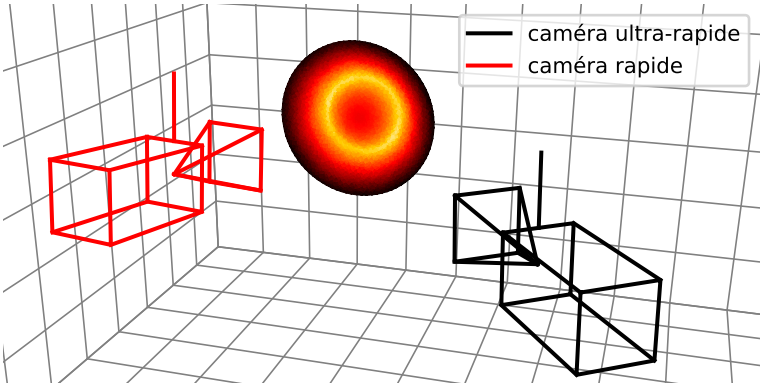


FIGURE 10.1 – Exemple de disposition des deux caméras asynchrones pour la mesure d'un haut-parleur

En illustration, un signal de déplacement de haut-parleur est simulé avec un fréquence d'échantillonnage égale à 80 000 Hz. Ce signal est moyenné dans le temps afin d'obtenir deux signaux, l'un à 8000 Hz et l'autre à 800 Hz, pour des temps d'exposition des photosites de $\approx 1/7200$ ms. Ces signaux sont ensuite projetés dans les repères images des deux caméras, virtuellement positionnées selon l'agencement présenté ci-dessus. Les valeurs des positions sont tronquées avec une précision de 4 décimales, et un bruit d'écart-type égal à 2/100 pixel est ajouté sur les positions 2D.

Le délai entre les signaux est tout d'abord calculé par une inter-corrélation (*cf.* **section 1.3.1**) entre le signal de détection d'un subset de la caméra-ultra rapide et celui d'un subset de la caméra rapide ré-échantillonné afin d'avoir la même fréquence d'échantillonnage pour les deux signaux. La méthode basée sur le gradient du traitement d'images (*cf.* équation (4.37) en **section 4.6.1**) est ensuite utilisée pour calculer le décalage entre les deux signaux, donnant accès à des valeurs inférieures à la période d'échantillonnage. Le signal de déplacement du point de mesure, ici selon l'axe $\overrightarrow{u}$ du repère image (*cf.* **section 3.2.1**), de la caméra ultra-rapide (en *noir*), et celui, interpolé, de la caméra rapide (en *rouge*) sont affichés sur la **figure 10.2**.

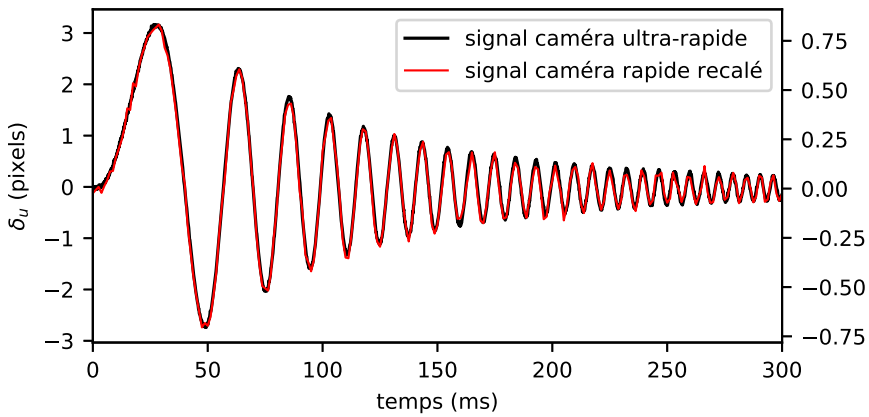


FIGURE 10.2 – Signaux de détection selon l'axe $\overrightarrow{u}$ des deux caméras

L'erreur sur le délai estimé, dans le cas présenté, est inférieure à 3% du temps d'échantillonnage de la caméra rapide. En choisissant un angle relativement faible, $\approx 10°$, entre l'axe de vibration et la caméra rapide, ce décalage n'entrave pas la triangulation des signaux temporels. La fonction de réponse en fréquence simulée (en *noir*), et celle estimée à partir des signaux de détection simulés (en *rouge*) sont affichées sur la **figure 10.3**. L'erreur résultante est globalement inférieure à 5% jusqu'à

200 Hz et augmente graduellement avec la fréquence. Malgré le bruit visible, sur la figure, on constate que la courbe estimée suit la courbe simulée.

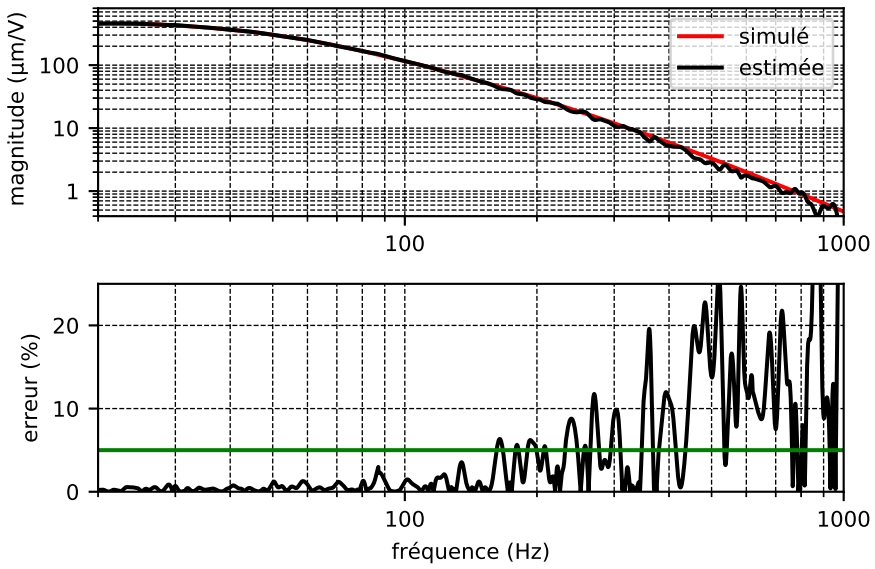


FIGURE 10.3 – Fonctions de réponse en fréquence du signal simulé et du signal estimé

Même si le principe inclut des approximations, le fait de conserver les rayons perspectifs de la caméra ultra-rapide comme ligne de triangulation permet *de facto* de mesurer des déplacements 3D, avec des potentiels écarts d'amplitude en hautes fréquences. De plus, afin de limiter les erreurs, un filtrage spatial des signaux de déplacement est réalisé à chaque image : pour chaque point de mesure, les points de mesure situés à une distance R, dans l'espace à trois dimensions, sont utilisés pour moyenner le déplacement à chaque image. Les poids associés sont alors calculés en fonction de la distance, avec l'équation (4.6) du filtre convolutif Gaussien, la somme des coefficients étant réglée à 1.

Les différentes étapes du protocole appliqué sont ainsi les suivantes :

1. ré-échantillonner le signal de détection d'un subset de mesure de la caméra rapide ;
2. estimer le délai entre le signal ré-échantillonné de la caméra rapide et le signal homologue de la caméra ultra-rapide par une inter-corrélation puis avec la méthode du gradient ;
3. interpoler les signaux de détection de la caméra rapide à partir de la valeur estimée du délai et trianguler les positions 3D correspondantes ;
4. effectuer un filtrage des signaux de déplacement à chaque image en calculant des poids à partir d'une part de l'équation du filtre convolutif Gaussien et d'autre part de la distance entre chaque point de mesure et les points qui lui sont adjacents dans l'espace à trois dimensions.

10.3 Application à une plaque

L'approche a été en premier lieu testée sur une plaque en vibration, excitée par un pot-vibrant, afin d'obtenir une validation sur une mesure plein champ par comparaison avec les résultats obtenus avec le montage pseudo stéréoscopique classique avec un jeu de quatre miroirs, considéré ici comme référence. Rappelons que les mesures ont été été effectuées successivement, sur la même plaque, avec les montages présentés dans ce manuscrit et que les informations liées au matériel utilisé pour le

montage, à l'agencement des éléments ainsi qu'au calibrage, ont été données de façon transversale en **section 8.4.1**. Dans cette section, ne seront présentés que les résultats obtenus.

10.3.1 Fonctions de réponse en fréquence

Afin de vérifier les résultats, la fonction de réponse en fréquence, accélération sur force, mesurée avec la méthode asynchrone, en *noir*, est comparée avec celle obtenue avec un accéléromètre, en *rouge* sur la **figure 10.4**. Malgré des différences pour les fréquences inférieures à 70 Hz environ, dues principalement à la plus faible sensibilité des accéléromètres en basses fréquences, une bonne concordance visuelle est notable. Une comparaison chiffrée ne semble pas pertinente puisqu'il est impossible de parfaitement aligner la position du subset de mesure et de l'accéléromètre situé à l'arrière de la plaque. De plus, le subset de mesure correspond à un disque de 6,6 mm de diamètre sur la plaque, alors que l'accéléromètre est vissé sur une pastille de 20 mm de diamètre : l'intégration spatiale est donc différente.

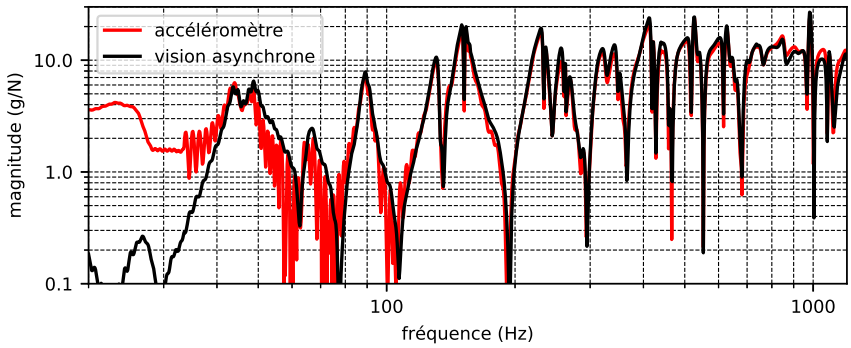


FIGURE 10.4 – Fonctions de réponse en fréquence mesurées pour un point avec un accéléromètre et la méthode à caméras asynchrones

10.3.2 Déformées modales opérationnelles

Afin de valider les résultats obtenus sur toute la surface de mesure, des déformées modales opérationnelles mesurées avec la méthode à caméras asynchrones sont comparées avec celles obtenues avec le montage pseudo stéréoscopique classique, la rotation de la plaque étant calculée et corrigée dans les deux cas. La **figure 10.5** présente les déformées à 222 Hz, 594 Hz, 888 Hz et 1147 Hz, à *gauche* pour la méthode à caméras asynchrones et à *droite* pour le montage pseudo stéréoscopique classique. Les coefficients de corrélation des déformées estimées avec les deux méthodes sont compris entre 99,3% et 99,8%, ce qui valide les résultats plein champ de la méthode à caméras asynchrones par rapport à la méthode à quatre miroirs.

10.3.3 En bref...

Les résultats obtenus avec la méthode à caméras asynchrones sont donc validés pour un point en utilisant un accéléromètre comme référence (*cf.* **figure 10.4**) et en plein champ en comparant les résultats avec ceux obtenus avec le montage pseudo stéréoscopique avec jeu de quatre miroirs (*cf.* **figure 10.5**). Même si l'objet vibre principalement selon un seul axe, la méthode prend en compte le mouvement 3D global, potentiellement multi-axial, et permet de corriger le mouvement de corps rigide. Elle est donc appliquée à deux cas industriels, dans la logique du travail effectué précédemment : à savoir un capot de voiture pour l'aspect vibratoire, et un haut-parleur pour l'aspect vibro-acoustique. L'objectif, à ce stade, est d'effectuer une étude de faisabilité, volontairement limitée à des contextes très ciblés, respectivement :

— l'applicabilité à un objet de grande taille, afin d'envisager à terme la caractérisation de panneaux de voitures, de trains ou autres, par exemple ;

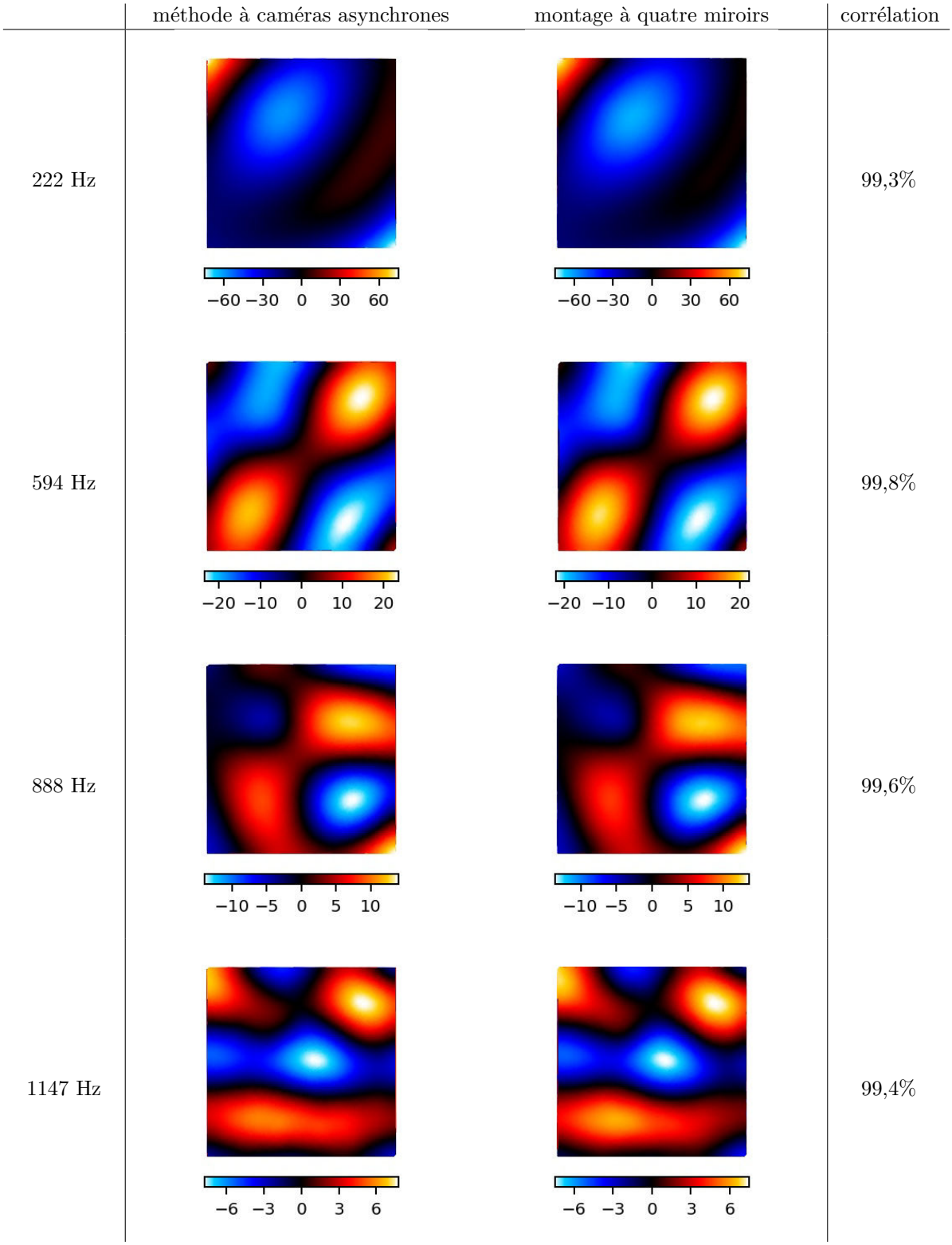


FIGURE 10.5 – Exemples de déformées modales opérationnelles, en µm/N, calculées avec la méthode à caméras asynchrones (à *gauche*) et le montage pseudo stéréoscopique classique (à *droite*) ; de *haut* en *bas* à 222 Hz, 594 Hz, 888 Hz et 1147 Hz

— l'applicabilité à la détection de défauts sur la suspension d'un haut-parleur, dans une logique de test-qualité de fin de chaine d'assemblage par exemple.

Les résultats présentés dans les sections suivantes correspondent donc uniquement à des informations initiales, potentiellement préliminaires à des études plus poussées.

10.4 Application industrielle : cas du capot d'une voiture

La méthode à caméras asynchrones ayant été validée, elle a été appliquée à la mesure de vibration d'un capot de voiture, afin d'évaluer l'applicabilité de l'approche à des objets plus grands.

10.4.1 Montage et protocole

La voiture est placée à l'intérieur de la salle semi-anéchoïque du CTTM. La zone de mesure correspond à une surface d'environ 600 mm × 800 mm, au centre du capot, sur laquelle est collé un motif aléatoire constitué d'ellipses, imprimé sur des feuilles auto-collantes, comme le montre la **figure 10.6** à *gauche*. La caméra rapide (Phantom) est placée en face du capot, tandis que la caméra ultra-rapide (Photron) est placée latéralement. Afin de positionner les caméras le plus haut possible, des trépieds eux-même posés sur des tables sont utilisés. L'agencement des caméras est présenté sur la **figure 10.6** à *droite*.

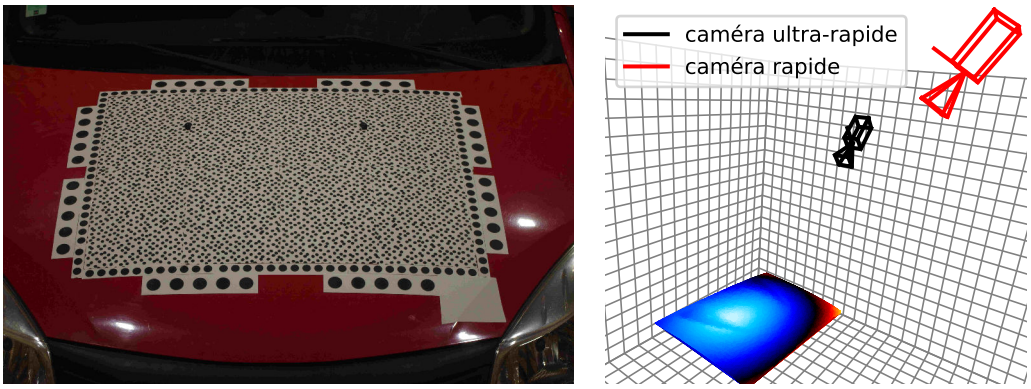


FIGURE 10.6 – (*gauche*) Photographie du motif aléatoire et des cibles collées sur le capot ; (*droite*) représentation 3D de la forme mesurée et de l'agencement des caméras

La mire utilisée est carrée, constituée de 19 × 19 cibles, collées sur un plaque en aluminium d'environ 300 mm × 300 mm. Le repère lié au capot est défini tel que l'axe $\overrightarrow{x}$ soit vers la gauche de la voiture, l'axe $\overrightarrow{y}$ soit vers l'avant de la voiture, et l'axe $\overrightarrow{z}$ soit dirigé vers le sol. Les angles de rotation autour des axes $\overrightarrow{y}$, $\overrightarrow{x}$ puis $\overrightarrow{z}$, par rapport au plan moyen de la surface de mesure, sont respectivement de ≈ 5°, ≈ -44° et ≈ 88° pour la caméra ultra-rapide, et de ≈ 4°, ≈ -29° et ≈ 0° pour la caméra rapide. Les distances épipolaires obtenues au niveau de la surface de mesure sont de ≈ 0,12 pixel en moyenne avec un écart-type de ≈ 0,17 pixel.

Lorsque les caméras sont positionnées et que le montage stéréoscopique est calibré, la mesure est effectuée en claquant le capot, *i.e.* en le soulevant au maximum, puis en le relâchant. Les cadences d'acquisition sont réglées à 12 500 images par seconde pour la caméra ultra-rapide et à 1200 images par seconde pour la caméra rapide. Les séquences d'images durent une seconde chacune.

La voiture subissant une légère rotation lorsque le capot est rabattu, des cibles de surface plus grande sont positionnées autour du motif aléatoire afin de mesurer et corriger cette rotation. Ainsi, à chaque image, le plan moyen de ces cibles est calculé, donnant accès à la normale (*cf.* **section 3.3.4**).

La rotation est corrigée par rapport au plan moyen des cibles lorsque le capot est au repos. Pour appliquer la correction à une image donnée, la forme mesurée est déplacée tel que son barycentre soit centré en zéro, avant d'appliquer la rotation et de replacer la forme autour de son barycentre initial. Les signaux obtenus pour les rotations autour des axes $\overrightarrow{x}$ et $\overrightarrow{y}$ sont affichés sur la **figure 10.7**. La rotation autour de l'axe $\overrightarrow{z}$ est ici négligeable.

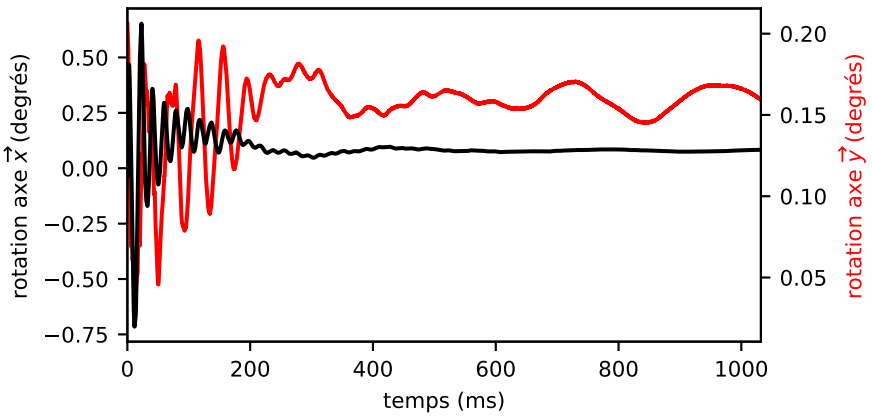


FIGURE 10.7 – Signaux de rotation autour des axes $\overrightarrow{x}$ et $\overrightarrow{y}$ mesurés pour le claquement du capot

10.4.2 Déformées modales opérationnelles

En calculant la transformée de Fourier des signaux mesurés, aux fréquences obtenues à partir de l'écart-type de l'auto-spectre (*cf.* **section 1.3.2**), des cartes de déformées modales avec des formes plus ou moins complexes sont observables. Des exemples sont affichés sur la **figure 10.8**, à 288 Hz, 488 Hz, 696 Hz, 826 Hz, 975 Hz et 1178 Hz. La symétrie de la forme du capot ainsi que l'excitation uniforme selon l'axe $\overrightarrow{x}$ génèrent une symétrie notable des cartes de déformées modales observées.

10.4.3 En bref...

La principale difficulté de cette expérience est la taille de la surface de mesure, couplée au fait que le capot est presque à l'horizontale. Le positionnement des caméras et le calibrage du système stéréoscopique sont alors délicats. De plus, le capot est claqué, les signaux vibratoires sont donc transitoires et non répétables. Notons aussi que la voiture subit une rotation relativement basses fréquences qui est ici corrigée à l'aide du déplacement mesuré au niveau de cibles situées autour de la zone de mesure (*cf.* **figure 10.7**).

En dépit de ces contraintes, les déformées modales opérationnelles mesurées sont très cohérentes, ce qui valide la faisabilité de la mesure et montre que la méthode à caméras asynchrones proposée est tout à fait applicable à la mesure de larges objets d'étude.

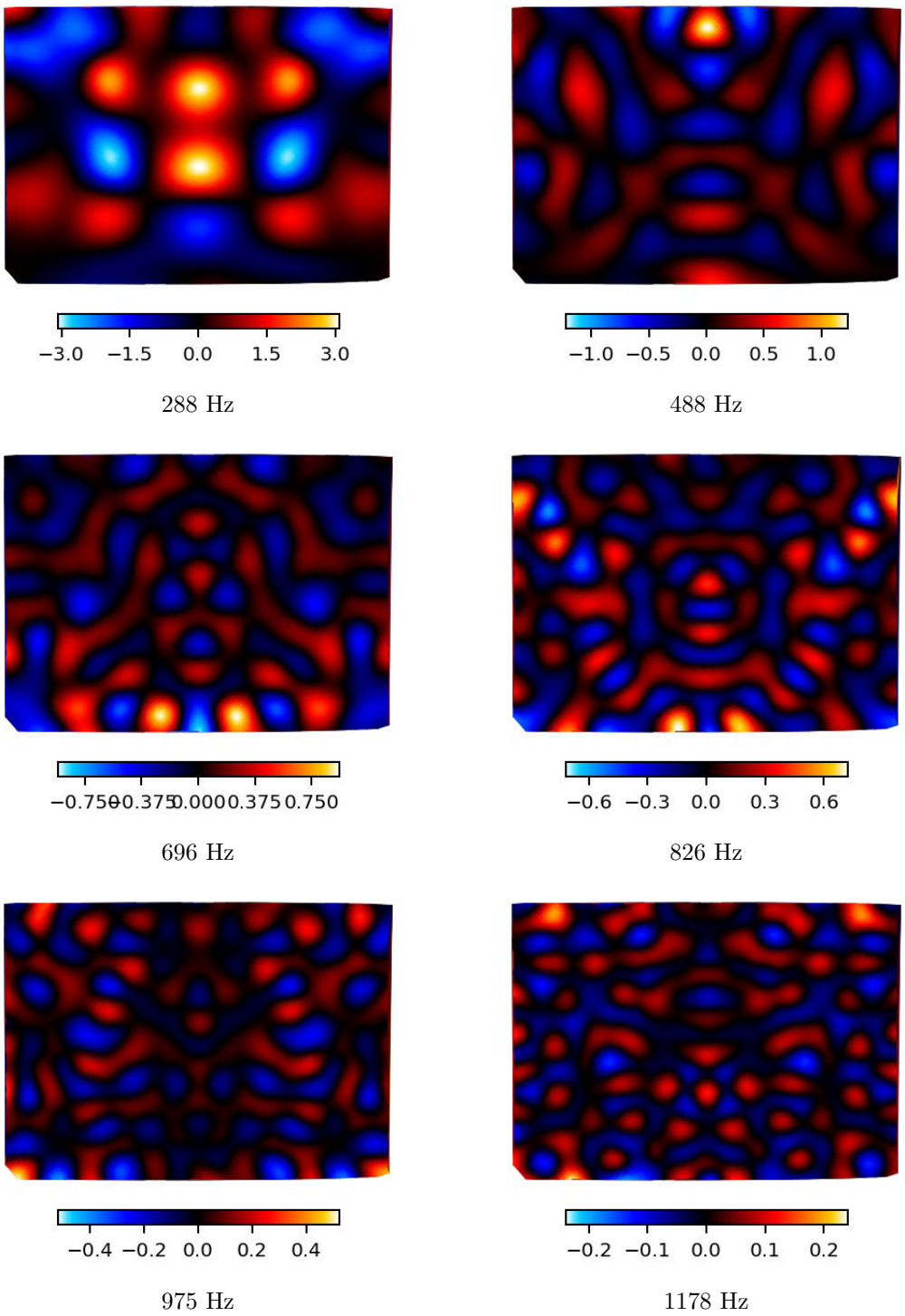


FIGURE 10.8 – Exemples de déformées modales mesurées avec la méthode à caméras asynchrones sur un capot de voiture, en µm ; de *gauche* à *droite* et de *haut* en *bas* : à 288 Hz, 488 Hz, 696 Hz, 826 Hz, 975 Hz et 1178 Hz

10.5 Application industrielle : cas du test d'un haut-parleur dégradé

Cette section présente les résultats obtenus pour la mesure d'un haut-parleur, volontairement dégradé, afin d'évaluer l'applicabilité de la méthode à la détection d'un défaut de suspension. Deux mesures ont ainsi été réalisées. Entre la première et la seconde mesure, la suspension externe a été entaillée sur environ deux centimètres (*cf.* **figure 10.9**), le but étant de simuler un défaut de symétrie de la suspension, lors du collage par exemple [2].

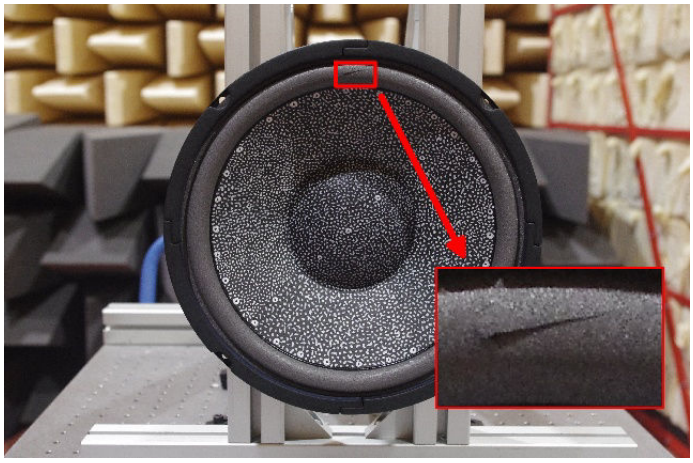


FIGURE 10.9 – Photographie du haut-parleur après l'entaille réalisée sur la suspension externe

10.5.1 Montage et protocole

L'objet d'étude est un haut-parleur AB Sound TW-1041, sur lequel un motif aléatoire constitué de taches a été dessiné à l'aide d'un feutre blanc. À la main, l'opération prend environ trois heures. Cependant, pour une application fréquente, un masque imprimé en 3D et une bombe de peinture pourraient très probablement être utilisés avec succès pour une opération réalisée en un temps bien plus court. La question de la commercialisation d'un haut-parleur présentant un mouchetis sur une partie de la membrane se pose alors. Un motif intégré sur toute la surface du haut-parleur ou simplement sur la membrane, ou encore un processus spécifique pour repeindre le haut-parleur pourraient être imaginés. Alternativement, un motif aléatoire non pas peint mais généré par l'illumination pourrait être utilisé. Le motif suivi ne serait alors pas fixe sur la surface mesurée : des algorithmes spécifiques seraient à développer pour prendre en compte cet aspect.

Les caméras utilisées sont :
- la caméra ultra-rapide Photron SA-X2 du LAUM avec un objectif de 50 mm ;
- la caméra rapide Phantom v5.1 du LAUM avec un objectif de 25 mm ;
- la caméra lente Mako U-503B avec un objectif de 17 mm.

La caméra Mako est utilisée comme troisième vue afin de simplifier l'appariement initial mené avec la méthode IRIs, tel que présenté au **chapitre 7**. Les données de calibrage et la mesure de forme associées ont été présentées en **section 7.6**. Les cadences d'acquisition sont de 8000 images par seconde pour la caméra Photron et de 800 images par seconde pour la caméra Phantom. Le signal d'excitation appliqué aux bornes du haut-parleur est un sinus glissant, de 2 V d'amplitude, allant de 5 Hz à 1300 Hz en 2,25 secondes. Un vibromètre laser Polytec a été utilisé afin d'avoir un point de contrôle, depuis l'arrière de la membrane. Une photographie du montage est affichée sur la **figure 10.10**.

2. La suspension externe est relativement souple comparativement à la suspension principale (*spider*), la dégradation ne modifie donc pas profondément le comportement vibratoire du haut-parleur.

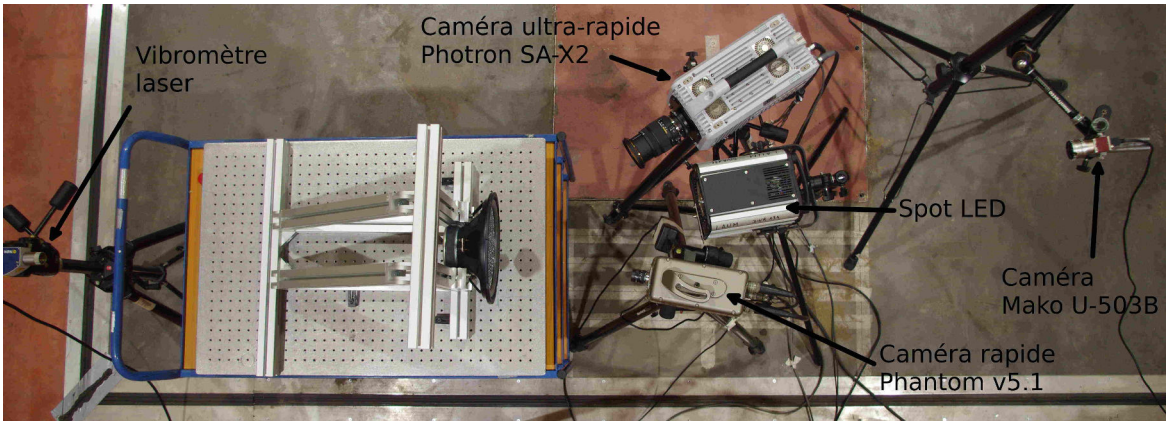


FIGURE 10.10 – Photographie du montage réalisé

10.5.2 Fonctions de réponse en fréquence du haut-parleur avant dégradation

La **figure 10.11** présente l'image acquise avec la caméra Phantom, ainsi que la position du subset (en *rouge*) utilisé pour comparer la fonction de réponse en fréquence obtenue avec celle calculée à partir du signal mesuré avec le vibromètre laser. Les réponses en fréquence résultantes sont affichées

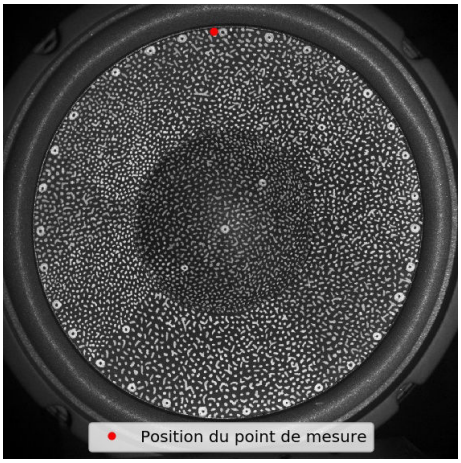


FIGURE 10.11 – Image du haut-parleur acquise avec la caméra Phantom et position du point de mesure

sur la **figure 10.12**, en *rouge* pour le vibromètre et en *noir* pour la méthode de vision. La différence entre les deux courbes est globalement inférieure à 5% (*cf.* ligne *verte*) jusqu'à $\approx$ 230 Hz. Au-delà de cette fréquence, l'allure des courbes est visuellement similaire, malgré un niveau moins élevé pour la méthode de vision. Il est à noter qu'il est très difficile de parfaitement recaler la position du point de mesure des deux méthodes, puisque le laser est situé à l'arrière du haut-parleur. Les écarts en hautes fréquences pourraient donc être en partie dus à une différence de positionnement des points de mesure affichés.

10.5.3 Déformées modales opérationnelles du haut-parleur avant dégradation

Cette section présente quelques exemples de déformées modales opérationnelles, mesurées jusqu'à $\approx$ 1000 Hz, pour le haut-parleur avant l'entaille, selon les trois axes de l'espace, à savoir : à 289 Hz (**figure 10.13**), 410 Hz (**figure 10.14**), 600 Hz (**figure 10.15**) et 978 Hz (**figure 10.16**). Le calcul de

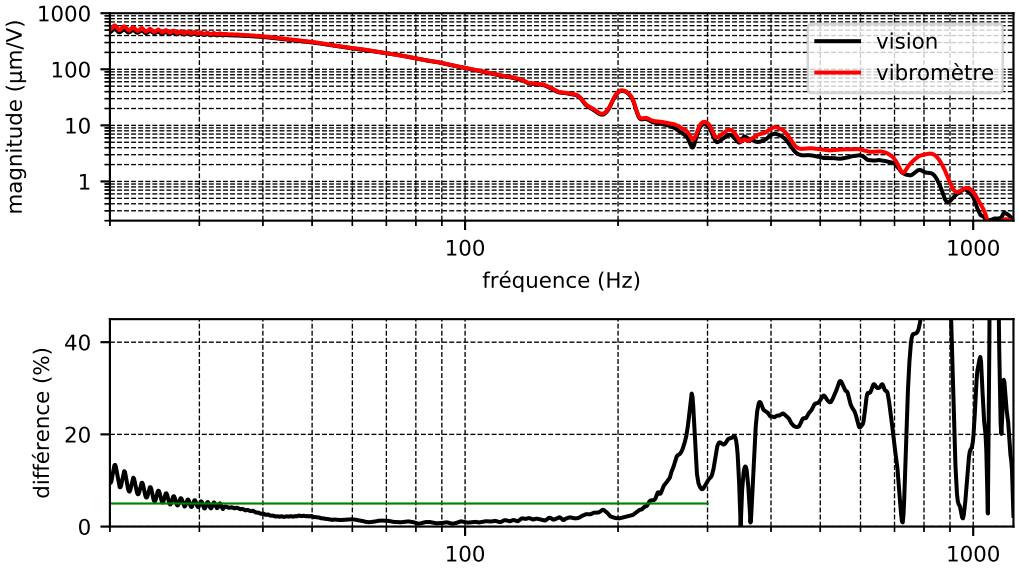


FIGURE 10.12 – Fonctions de réponse en fréquence mesurées avec un vibromètre laser et la méthode de vision

la normale indique que $\approx 88\%$ du mouvement se fait selon l'axe $\overrightarrow{z}$. On constate ainsi que davantage d'information est fournie sur la carte de droite à chaque fréquence affichée. Il est néanmoins possible de retrouver l'information correspondant au mouvement selon les deux autres axes, avec parfois des erreurs plus visibles sur la gauche de la membrane. Celles-ci sont très vraisemblablement dues à l'angle relativement faible entre la caméra et cette zone du haut-parleur, la caméra étant alors à cet endroit beaucoup moins sensible au déplacement normal. La mesure n'est certes pas parfaite mais les déformées semblent visuellement très cohérentes et donnent des informations pertinentes sur le comportement vibratoire de la membrane.

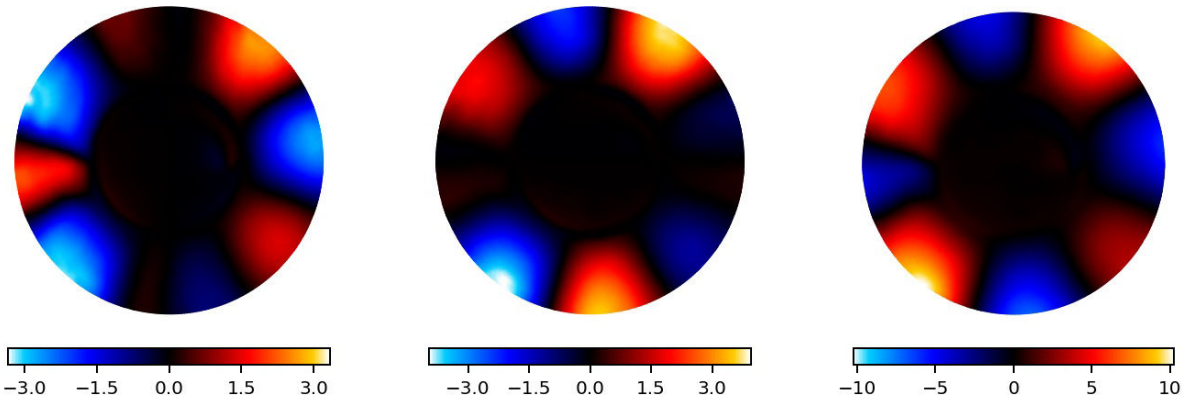


FIGURE 10.13 – Exemples de déformée modale mesurée à 289 Hz ; les données sont exprimées en µm/V, (*gauche*) selon l'axe $\overrightarrow{x}$, (*milieu*) selon l'axe $\overrightarrow{y}$, (*droite*) selon l'axe $\overrightarrow{z}$

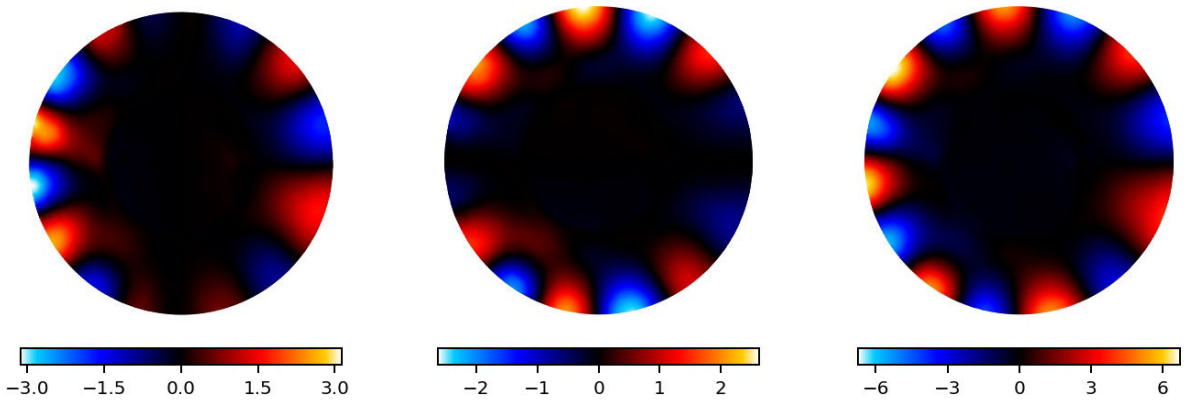


FIGURE 10.14 – Exemple de déformée modale mesurée à 410 Hz ; les données sont exprimées en µm/V, (*gauche*) selon l'axe $\overrightarrow{x}$, (*milieu*) selon l'axe $\overrightarrow{y}$, (*droite*) selon l'axe $\overrightarrow{z}$

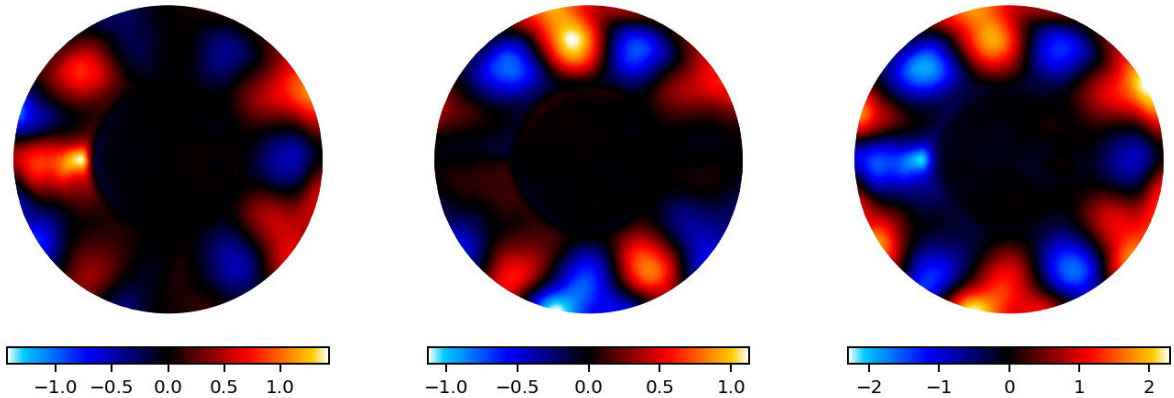


FIGURE 10.15 – Exemple de déformée modale mesurée à 600 Hz ; les données sont exprimées en µm/V, (*gauche*) selon l'axe $\overrightarrow{x}$, (*milieu*) selon l'axe $\overrightarrow{y}$, (*droite*) selon l'axe $\overrightarrow{z}$

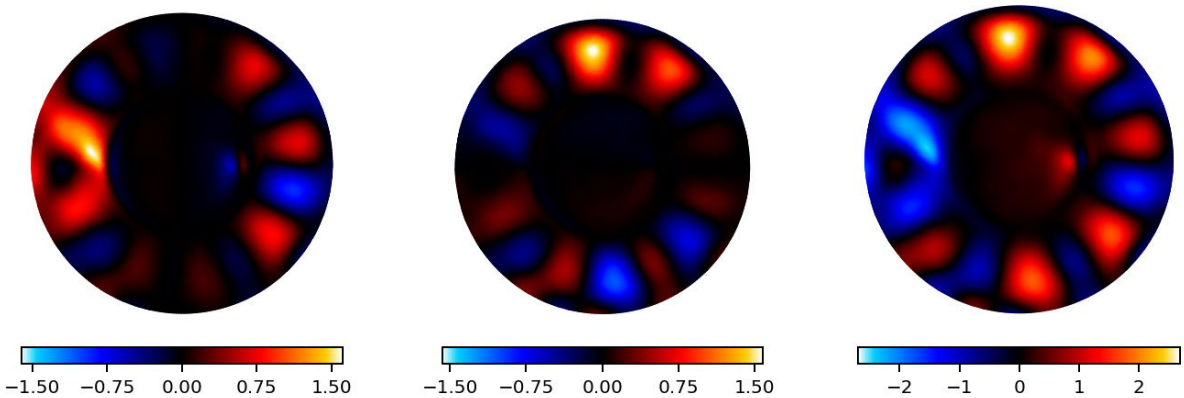


FIGURE 10.16 – Exemple de déformée modale mesurée à 978 Hz ; les données sont exprimées en µm/V, (*gauche*) selon l'axe $\overrightarrow{x}$, (*milieu*) selon l'axe $\overrightarrow{y}$, (*droite*) selon l'axe $\overrightarrow{z}$

10.5.4 Résultats après dégradation du haut-parleur

Écart-type des auto-spectres

Les courbes de l'écart-type des auto-spectres des signaux mesurés selon l'axe $\vec{z}$, axe de déplacement de la bobine, sont affichées sur la **figure 10.17**. La courbe *noire* correspond à la première mesure

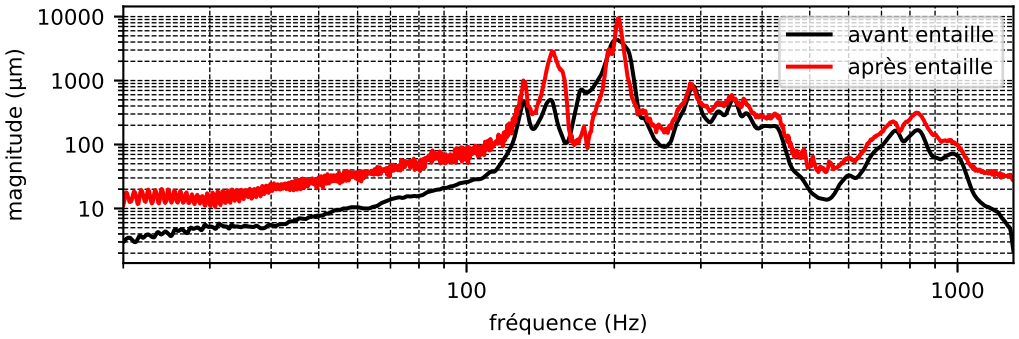


FIGURE 10.17 – Écart-type des auto-spectres des signaux de mesure selon l'axe $\vec{z}$ avant et après l'entaille

avant l'entaille de la suspension externe, et la courbe *rouge* à la seconde mesure, après l'entaille. Les principales différences se situent à $\approx$ 150 Hz et $\approx$ 170 Hz, avec un augmentation d'un facteur $\approx$ 4 du niveau du premier pic et une diminution d'un facteur $\approx$ 10 du niveau du deuxième pic. Ces fréquences correspondent à des *rocking modes*, sensibles à l'asymétrie de la raideur de la suspension engendrée par l'entaille.

Affichage des déformées modales opérationnelles

Pour le premier *rocking mode* à 131 Hz, affiché sur la **figure 10.18**, la seule différence visible est une légère augmentation du niveau, de $\approx$ 5 µm/V au maximum avant dégradation à $\approx$ 6 µm/V au maximum après dégradation. Pour le *rocking mode* à 147 Hz (*cf.* **figure 10.19**), l'amplitude augmente

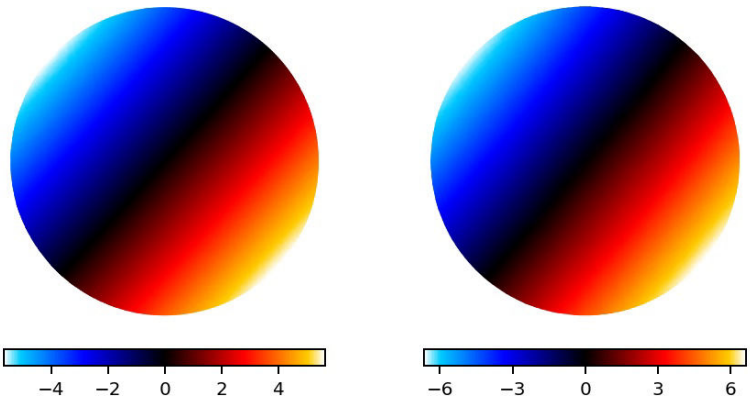


FIGURE 10.18 – *Rocking mode* observé à 131 Hz avant (*gauche*) et après (*droite*) dégradation

après dégradation d'un facteur proche de deux, et l'axe de la rotation est décalé d'environ 20°. Pour le *rocking mode* à 171 Hz (*cf.* **figure 10.20**), cette fois l'amplitude diminue après dégradation d'un facteur de $\approx$ 4, et l'axe de la rotation est décalé d'environ 40°.

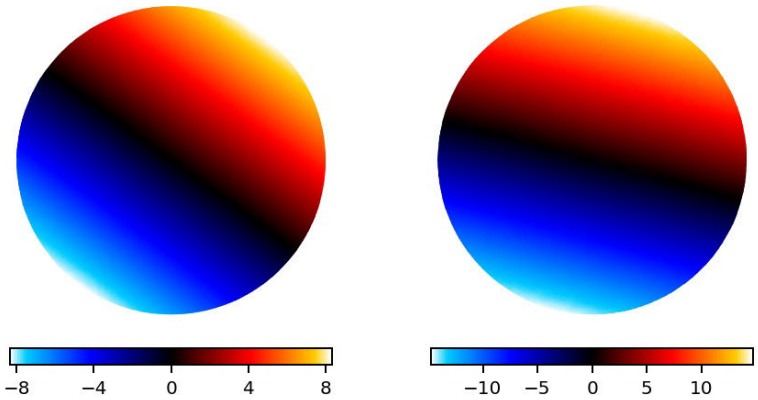


FIGURE 10.19 – *Rocking mode* observé à 147 Hz avant (*gauche*) et après (*droite*) dégradation

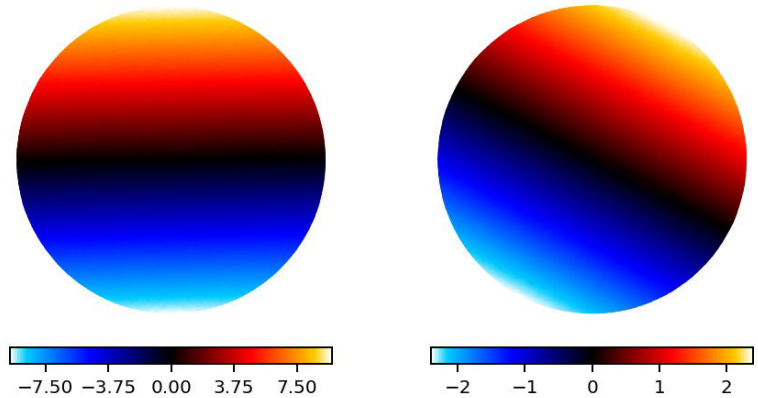


FIGURE 10.20 – *Rocking mode* observé à 171 Hz avant (*gauche*) et après (*droite*) dégradation

10.5.5 En bref...

Rappelons que cette section visait à évaluer l'applicabilité de la méthode à caméras asynchrones à la détection d'un défaut de suspension d'un haut-parleur. Les déformées modales opérationnelles obtenues à partir des signaux mesurés peuvent être calculées et affichées jusqu'à ≈ 1000 Hz. Les résultats attestent de la pertinence de l'approche puisque la mesure plein champ effectuée du déplacement de la membrane avant et après dégradation du haut-parleur permet de mettre à jour des différences, dues en toute logique à l'impact de l'entaille sur le comportement dynamique du haut-parleur. On constate en particulier une altération des *rocking modes* de la membrane. En ne mesurant que quelques points sur le bord de la membrane, la symétrie de la raideur des suspensions pourrait très probablement être vérifiée en temps réel. Dans l'optique par exemple d'une vérification d'un haut-parleur en fin de chaine d'assemblage, des études ultérieures pourraient s'intéresser notamment à la définition d'un critère évaluant, de façon chiffrée, la qualité des suspensions ainsi qu'à une simplification de la phase de montage et de réglage des éléments requis, dans la logique d'une semi-automatisation du processus de vérification. Notons qu'il est aussi possible de mesurer des paramètres non-linéaires d'un haut-parleur à partir de la mesure de son déplacement. Cette application, présentée avec l'utilisation de la méthode sans miroir (*cf.* **chapitre 9**) en **Annexe B** est aussi possible avec la méthode à caméras asynchrones.

10.6 Conclusion

En conclusion, ce chapitre a présenté la méthode à caméras asynchrones conçue afin de conserver les avantages du montage sans miroir, en termes de maniabilité et de précision, tout en ciblant les phénomènes multi-axiaux et non plus simplement mono-axiaux. Le principe consiste à utiliser une caméra ultra-rapide et une caméra rapide, aux cadences d'acquisition différentes, mesurant alors simultanément des vibrations selon les trois axes de l'espace de façon synchrone avec des caméras asynchrones, la caméra rapide étant mobilisée pour la mesure des fréquences jusqu'à environ 300 Hz et la caméra ultra-rapide pour la mesure des fréquences allant jusqu'à environ 1200 Hz en fonction de l'objet d'étude. L'approche a été validée sur un point en comparant les résultats avec ceux obtenus avec un accéléromètre et en plein champ avec une mesure des vibrations d'une plaque, les résultats étant alors comparés avec ceux obtenus avec le montage pseudo stéréoscopique classique avec un jeu de quatre miroirs, considéré dans ce cas comme référence.

Le traitement des signaux obtenus expérimentalement prenant en compte la composante multi-axiale du déplacement, la méthode est applicable à des objets d'étude dont les vibrations suivent les trois axes de l'espace. Afin de tester son applicabilité à des cas industriels, une première mesure a été effectuée sur le capot d'une voiture. Les résultats de l'étude de faisabilité menée montrent que la technique est applicable à la mesure de vibration de grands objets d'étude. Elle pourrait donc être utilisée pour la caractérisation de panneaux par exemple. Une seconde mesure a été menée sur un haut-parleur volontairement dégradé afin d'évaluer la pertinence de son application à la détection d'un défaut de suspension. Les résultats montrent que la technique à caméras asynchrones permet de mettre à jour des variations du comportement dynamique du haut-parleur avant et après dégradation, ce qui laisse envisager, après des études plus approfondies, une application pour le test-qualité en fin de chaine d'assemblage par exemple.

Conclusion générale

« *L'imagination est plus importante que le savoir.* »
Albert Einstein

Les méthodes de vision 3D permettent d'effectuer des mesures de forme plein champ. Couplées à une ou deux caméras (ultra-)rapides, elles donnent accès à la dimension temporelle du déplacement de la structure étudiée. Elles constituent ainsi une alternative potentielle intéressante aux méthodes de mesure de vibrations utilisées actuellement, avec la possibilité d'une mesure plein champ et synchrone de phénomènes vibratoires transitoires, non répétables, possiblement non-linéaires. Le travail de recherche qui a été présenté dans ce manuscrit visait ainsi à évaluer la pertinence et les limites de la vision 3D couplée à une caméra ultra-rapide pour la mesure 3D de champs vibratoires, sans contact et de manière synchrone, dans le domaine des fréquences associées à la vibro-acoustique, sur des objets d'étude en adéquation avec les champs disciplinaires concernés (*i.e.* une plaque, un capot de voiture, un tambourin et un haut-parleur). La démarche s'inscrit dans le prolonguement de la mesure de forme déjà très étudiée avec ces techniques puisque l'intérêt se porte cette fois sur la mesure de petits déplacements, allant de moins d'un micromètre jusqu'au millimètre. Ces ordres de grandeur ont dû être pris en compte lors de la conception des méthodes de vision proposées.

La première étape du travail a ciblé la mise en œuvre du montage pseudo stéréoscopique classique avec une seule caméra ultra-rapide et un jeu de quatre miroirs, issu de la robotique, pour la mesure de vibrations avec l'utilisation de cibles sur une plaque et un haut-parleur. Ce dernier objet d'étude, de par sa géométrie 3D complexe, a mis a jour l'inadéquation du suivi de cibles dans ce contexte. Le passage à l'utilisation d'un mouchetis et à des outils de corrélation d'images a mené à une deuxième phase de travail : la conception d'une méthode itérative de rectification d'images (méthode IRIs) adaptée aux besoins spécifiques identifiés. Le montage à quatre miroirs s'avérant particulièrement délicat à manipuler, un troisième axe de recherche s'est intéressé en parallèle à la simplification des éléments optiques et a abouti à la conception d'un montage pseudo stéréoscopique à deux miroirs uniquement. Une quatrième phase de travail s'est attachée à simplifier le protocole global et a mené à la conception de deux méthodes sans miroir. La première est adaptée à la mesure de vibrations localement mono-axiales, dans la mesure où la reconstruction 3D que permettent ces deux montages pseudo stéréoscopiques ne semble pas systématiquement nécessaire, certains objets vibrant majoritairement selon un seul axe, celui de la normale locale à la surface. La seconde est en revanche adaptée à la mesure de vibrations multi-axiales d'objets vibrant en 3D et utilise deux caméras asynchrones (une caméra rapide est alors requise en plus de la caméra ultra-rapide).

Méthode itérative de rectification d'images - IRIs

La méthode IRIs a été développée afin de permettre, avec une implémentation relativement simple, d'appairer, à partir d'une mesure de forme initiale, des points de mesure avec un angle supérieur à 60° entre les caméras d'un système stéréoscopique, pour des objets présentant d'importantes variations de profondeur.

Un des principaux avantages de la méthode réside dans le fait que les calculs sont effectués sur des valeurs de niveaux de gris directement interpolées dans les images initiales, ce qui réduit les imprécisions liées aux interpolations successives. Autre atout : la forme globale de la surface mesurée (et non la déformation locale de chaque subset) est estimée itérativement par le biais d'ortho-images, ce qui permet d'utiliser une mesure de forme initiale très grossière avec des points appairés avec un précision relativement faible, éventuellement en partie avec des entrées graphiques.

La méthode, validée expérimentalement et numériquement avec un haut-parleur comme objet d'étude, a ensuite été optimisée pour automatiser l'appariement pour la mesure de forme initiale, grâce à l'utilisation d'un troisième point de vue situé stratégiquement et obtenu avec une caméra standard. Des subsets sont alors calculés dans les images initiales à partir d'un cercle défini dans le plan moyen de la surface mesurée. L'algorithme final, intégrant cette automatisation de la phase de mesure de forme initiale, permet d'initialiser la mesure de vibrations et a été validé par les résultats obtenus. Il a été appliqué avec succès, en tant que pré-traitement, à toutes les données de mesure obtenues à partir d'un motif aléatoire et l'un ou l'autre des montages conçus.

Méthode de vision pour la mesure de vibrations multi-axiales avec le montage pseudo stéréoscopique avec jeu de quatre miroirs

En préliminaire à la mise en œuvre de ce montage, une étude de sensibilité, menée par le biais de simulations, a permis de démontrer que, pour la mesure de vibrations avec une méthode de suivi de cibles, il était préférable d'utiliser un motif point-Gaussien, avec l'estimateur d'ellipses duales ainsi qu'un angle entre les caméras virtuelles et la surface à mesurer de l'ordre de 50°. Les résultats de l'étude ont aussi permis d'attester que, dans la configuration sélectionnée, la mesure de petits déplacements était plus précise que la mesure de forme, et que la mesure de vibrations l'était encore davantage en basculant dans le domaine fréquentiel.

Le principe fondamental sous-jacent à l'élaboration du montage est celui de la pseudo stéréovision : deux points de vue virtuels sont générés à partir d'une unique caméra réelle ultra-rapide et d'un jeu de quatre miroirs, au détriment du nombre de pixels par vue puisque la matrice de photosites est divisée par deux. Les images 2D issues de ces deux points de vue sont ensuite traitées afin de reconstruire, par triangulation, une image 3D de l'objet d'étude à chaque instant. La méthode a été testée avec succès sur une plaque, les résultats étant validés par comparaison avec ceux obtenus avec un accéléromètre pour un point et un vibromètre laser à balayage pour une mesure plein champ. Elle a également été appliquée à la mesure de vibrations d'un haut-parleur, dont les caractéristiques géométriques requièrent une adaptation des paramètres. Même si des déformées modales opérationnelles ont pu être calculées à partir des réponses en fréquence et affichées, l'utilisation de cibles dans ce contexte s'est avérée inadaptée et a abouti à l'utilisation d'un mouchetis et d'outils de corrélation d'images pour les phases expérimentales ultérieures.

La limite fréquentielle de la méthode, 1000 Hz pour l'expérimentation menée, est plus faible que pour un vibromètre mais offre une mesure 3D de phénomènes vibratoires sans limitation quant à la non-stationnarité et la non-répétabilité des phénomènes étudiés, de par la mesure synchrone de plusieurs centaines de points de mesure. Ces résultats attestent de la validité du montage pseudo stéréoscopique pour mesurer des phénomènes vibratoires multi-axiaux de surfaces relativement planes, tel qu'une caisse de guitare, des capots ou des portières de voitures, des ailes d'avions, etc. L'utilisation des quatre miroirs permet aussi de réduire drastiquement les coûts du montage puisqu'une seule caméra est requise. Cet atout est néanmoins contrebalancé par un encombrement fort du montage, des réglages complexes et chronophages des éléments optiques, notamment pour tenter de limiter la zone d'ombre centrale présente dans les images. Peu maniable, ce capteur stéréoscopique est difficilement utilisable

hors du laboratoire et limité à la mesure de surfaces d'environ 300 mm × 300 mm, pour un angle entre la normale de la surface et les points de vue d'environ 50°, de par les contraintes spatiales imposées par l'agencement des différents éléments et la réduction de la taille de la matrice de pixels pour chaque point de vue.

Méthode de vision pour la mesure de vibrations multi-axiales avec le montage pseudo stéréoscopique simplifié avec jeu de deux miroirs

La simplification du montage pseudo stéréoscopique classique à quatre miroirs a abouti à la suppression du prisme central. Les deux miroirs latéraux sont conservés et réorientés afin de générer un seul point de vue virtuel, approximativement en face de l'objet d'étude. Le second point de vue, sensible au déplacement hors plan, est obtenu directement à travers la caméra ultra-rapide, placée latéralement. Ce second montage pseudo stéréoscopique est donc une alternative simplifiée spécifiquement conçue pour la mesure de vibrations multi-axiales. Les résultats obtenus sur une plaque et un haut-parleur, comparés avec ceux obtenus avec un accéléromètre pour un point de mesure et/ou avec ceux obtenus avec le montage classique pour une mesure plein champ, attestent de la validité de la méthode proposée.

Le montage à deux miroirs est beaucoup plus maniable et moins encombrant et la suppression du prisme central n'impacte pas visiblement, dans les tests effectués, la précision de la mesure de vibration. L'angle panoramique entre les caméras est *de facto* réduit de moitié, ce qui simplifie la phase d'appariement initiale. La zone d'ombre centrale présente dans les images obtenues avec le montage précédent est réduite. Les deux vues sont néanmoins toujours obtenues à l'aide d'une seule caméra. Le nombre de pixels par vue est ainsi égal à environ la moitié de la matrice disponible, avec une surface mesurable réduite d'environ 200 mm × 200 mm. Ce montage reste difficilement utilisable hors du laboratoire, de par l'utilisation des miroirs et d'une table optique.

Méthode de vision sans miroir pour la mesure de vibrations mono-axiales

Cette méthode repose sur l'hypothèse selon laquelle les objets vibrent majoritairement selon l'axe de la normale locale à la surface (ou selon un seul axe connu) et a été spécifiquement conçue pour la mesure plein champ de vibrations mono-axiales, avec une seule caméra ultra-rapide, pour des objets globalement plans affichant des déplacements répétables ou non-répétables. Les miroirs latéraux présents pour le montage à deux miroirs sont supprimés : le système n'est donc plus pseudo stéréoscopique et une seule vue, réelle, est utilisée. Placée latéralement, la caméra ultra-rapide est sensible au déplacement hors-plan alors que la seconde vue requise pour la triangulation des points est remplacée par une ligne mathématique, calculée par le biais de la mesure de forme initiale (obtenue à l'aide de l'utilisation ponctuelle d'une seconde caméra, simple ou grand public cette fois).

L'approche est validée par les résultats obtenus pour la mesure de vibrations d'un tambourin et d'une plaque par comparaison avec ceux obtenus avec un vibromètre laser pour un point et avec ceux obtenus avec le montage à quatre miroirs (selon l'axe $\overrightarrow{z}$) pour une mesure plein champ. Elle a aussi pu être utilisée avec succès pour mesurer le déplacement du point central d'un haut-parleur dont le mouvement est mono-axial et ainsi pour estimer les paramètres linéaires et un paramètre non-linéaire du haut-parleur. Un dernier test, mené sur la mesure plein champ d'un haut-parleur, dont le comportement dynamique ne suit pas un seul axe, a montré que la méthode proposée, bien qu'elle soit par nature fondamentalement inadaptée à la mesure multi-axiale, permet néanmoins d'obtenir des données initiales viables sur le comportement dynamique du haut-parleur et de circonscrire un protocole d'étude ultérieur.

L'avantage principal de la technique réside dans l'utilisation de toute la matrice de pixels pour une seule vue, avec un montage simple, peu encombrant et très maniable, pour une surface de mesure potentiellement grande. Elle peut facilement être utilisée hors du laboratoire, pour la mesure de vibrations d'objets dont le comportement dynamique respecte la contrainte mono-axiale supposée.

Méthode de vision à caméras asynchrones pour la mesure de vibrations multi-axiales

Cette technique a été conçue afin de conserver les avantages du montage sans miroir tout en mesurant des phénomènes multi-axiaux et non plus simplement mono-axiaux. Le principe repose sur l'utilisation de deux caméras, l'une ultra-rapide et l'autre rapide, aux cadences d'acquisition différentes. Les vibrations selon les trois axes de l'espace sont ainsi mesurées de façon synchrone avec des caméras asynchrones. La caméra rapide, située en face de l'objet d'étude, est mobilisée pour la mesure des fréquences allant jusqu'à environ 300 Hz. La caméra ultra-rapide, positionnée latéralement afin d'être sensible aux déplacements hors-plan, est affectée à la mesure des fréquences allant jusqu'à environ 1200 Hz, en fonction de l'objet d'étude. Les résultats valident l'approche, par la comparaison avec les données obtenues pour un point avec un accéléromètre et en plein champ avec le montage pseudo stéréoscopique à quatre miroirs.

La méthode à caméras asynchrone possède un certain nombre d'avantages : adaptée à la mesure de vibrations multi-axiales, potentiellement hors laboratoire, elle est très maniable avec un montage d'encombrement moyen, dû à la présence des deux caméras et des deux trépieds, ne nécessitant en revanche aucun élément opto-mécanique. Les caméras sont donc assez mobiles et permettent alors de mesurer des objets de plus grande taille. L'intégralité de la matrice de pixels est utilisée pour chaque point de vue et amène une meilleure précision. Le principal inconvénient est financier et vient de l'ajout d'une caméra rapide. Notons néanmoins qu'une caméra industrielle d'un coût modéré (quelques milliers d'euros) est tout à fait adaptée.

Le protocole est ainsi applicable à des objets d'étude dont les vibrations suivent les trois axes de l'espace. L'étude de faisabilité effectuée a validé une première application de la méthode à la mesure d'objets de taille plus importante (ici un capot de voiture d'environ 600 mm $\times$ 800 mm). Elle pourrait donc être utilisée pour la caractérisation de divers panneaux par exemple. Une seconde étude, menée sur un haut-parleur volontairement dégradé, a montré que la technique à caméras asynchrones permettait de mettre à jour des variations du comportement dynamique du haut-parleur avant et après dégradation. L'approche est donc validée pour la détection d'un défaut de suspension et pourrait, après des études complémentaires plus poussées, être à terme utilisée pour le test-qualité de fin de ligne de haut-parleurs par exemple.

Comparaison générale des quatre méthodes de mesure de vibrations

Le **tableau 10.1** reprend tout d'abord, sous forme synthétique et comparée, une partie des éléments généraux présentés ci-dessus.

Les sections suivantes proposent une comparaison des déformées modales opérationnelles obtenues avec les quatre montages présentés, lors d'une même phase expérimentale visant à mesurer les vibrations d'une plaque suspendue excitée par un pot vibrant [3]. Les coefficients de corrélation sont donnés pour chaque fréquence et chaque montage, en prenant pour référence les résultats obtenus avec le montage pseudo stéréoscopique classique à quatre miroirs.

Méthode mono-axiale et axe $\vec{z}$ du montage pseudo stéréoscopique à quatre miroirs

Afin de comparer des données comparables, seuls les résultats de l'axe $\vec{z}$ du montage à quatre miroirs sont pris en compte, la rotation de la plaque n'étant pas corrigée. Des exemples de déformées modales opérationnelles sont affichées sur la **figure 10.21**, à *gauche* pour la méthode mono-axiale sans miroir, à *droite* pour le montage classique à quatre miroirs, à 339 Hz, 490 Hz, 934 Hz, 1036 Hz et 1245 Hz. On constate une excellente corrélation, supérieure à 98%, malgré de légères différences d'amplitude occasionnelles, ce qui valide les résultats obtenus pour la mesure plein champ de vibrations mono-axiales avec la méthode sans miroir proposée.

3. Pour rappel, le protocole, les descriptions des montages et les données de calibrage sont donnés **section 8.4.1**.

Tableau 10.1 – Comparaison générale des quatre méthodes

Méthode	montage à 4 miroirs	montage à 2 miroirs	montage sans miroir	montage à caméras asynchrones
Nature du déplacement mesurable	multi-axial	multi-axial	mono-axial	multi-axial
Surface de mesure maximale	$\approx 300 \times 300$ mm^2	$\approx 200 \times 200$ mm^2	grande	grande
Nombre de pixels par vue*	$\approx 1024 \times 500$	$\approx 1024 \times 500$	1024×1024	1024×1024
Maniabilité	$--$	$-$	$+$	$+$
Encombrement du montage	fort	moyen	faible	moyen
Utilisation hors laboratoire / coût surajouté**	difficile / cher	difficile / cher	facile / faible	facile / modéré

* pour un matrice de 1024×1024 pixels par caméra
** estimation du coût surajouté à celui de la caméra ultra-rapide

Méthode à caméras asynchrones, montages pseudo stéréoscopiques à quatre et deux miroirs

Les trois méthodes permettent de mesurer des vibrations multi-axiales et dans les trois cas la rotation de la plaque a été calculée et corrigée. Des exemples de déformées modales opérationnelles sont affichés sur la **figure 10.22**, à *gauche* pour le montage à deux miroirs, au *milieu* pour la méthode à caméras asynchrones et à *droite* pour le montage classique à quatre miroirs à 339 Hz, 490 Hz, 934 Hz, 1036 Hz et 1245 Hz. Les coefficients de corrélation, indiqués dans le **tableau 10.2**, sont supérieurs à 98,8% et valident les mesures plein champ effectuées avec les méthodes présentées.

Tableau 10.2 – Coefficients de corrélation calculés (en %) pour les déformées modales opérationnelles obtenues avec le montage à deux miroirs et la méthode à caméras asynchrones

Fréquence (Hz)	339	490	934	1036	1245
Montage à deux miroirs	100	99,7	99,4	99,8	99,1
Méthode à caméras asynchrones	99,9	98,8	99,0	99,8	99,0

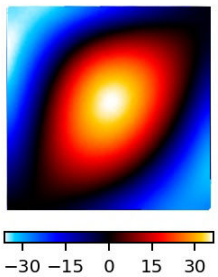
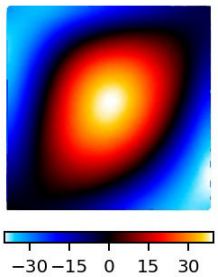
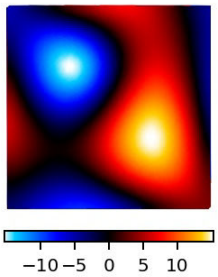
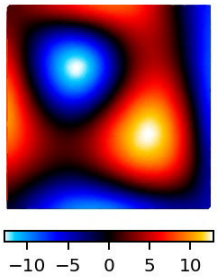
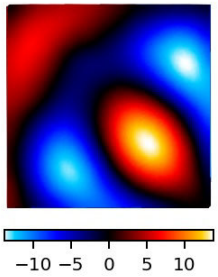
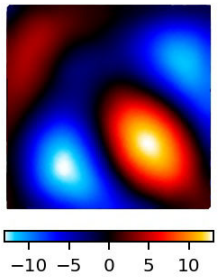
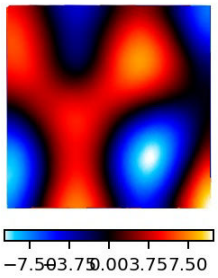
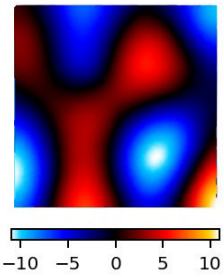
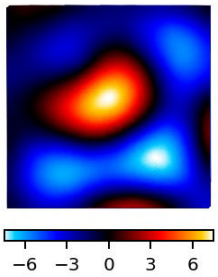
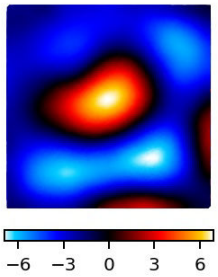
	méthode sans miroir mono-axiale	montage à quatre miroirs (axe $\overrightarrow{z}$)	corrélation
339 Hz			99,9%
490 Hz			99,6%
934 Hz			98,1%
1036 Hz			99,8%
1245 Hz			99,2%

FIGURE 10.21 – Exemples de déformées modales opérationnelles, en µm/N, calculées avec (*gauche*) la méthode sans miroir mono-axiale et (*droite*) le montage pseudo stéréoscopique classique (axe $\overrightarrow{z}$) ; de *haut* en *bas* à 339 Hz, 490 Hz, 934 Hz, 1036 Hz et 1245 Hz

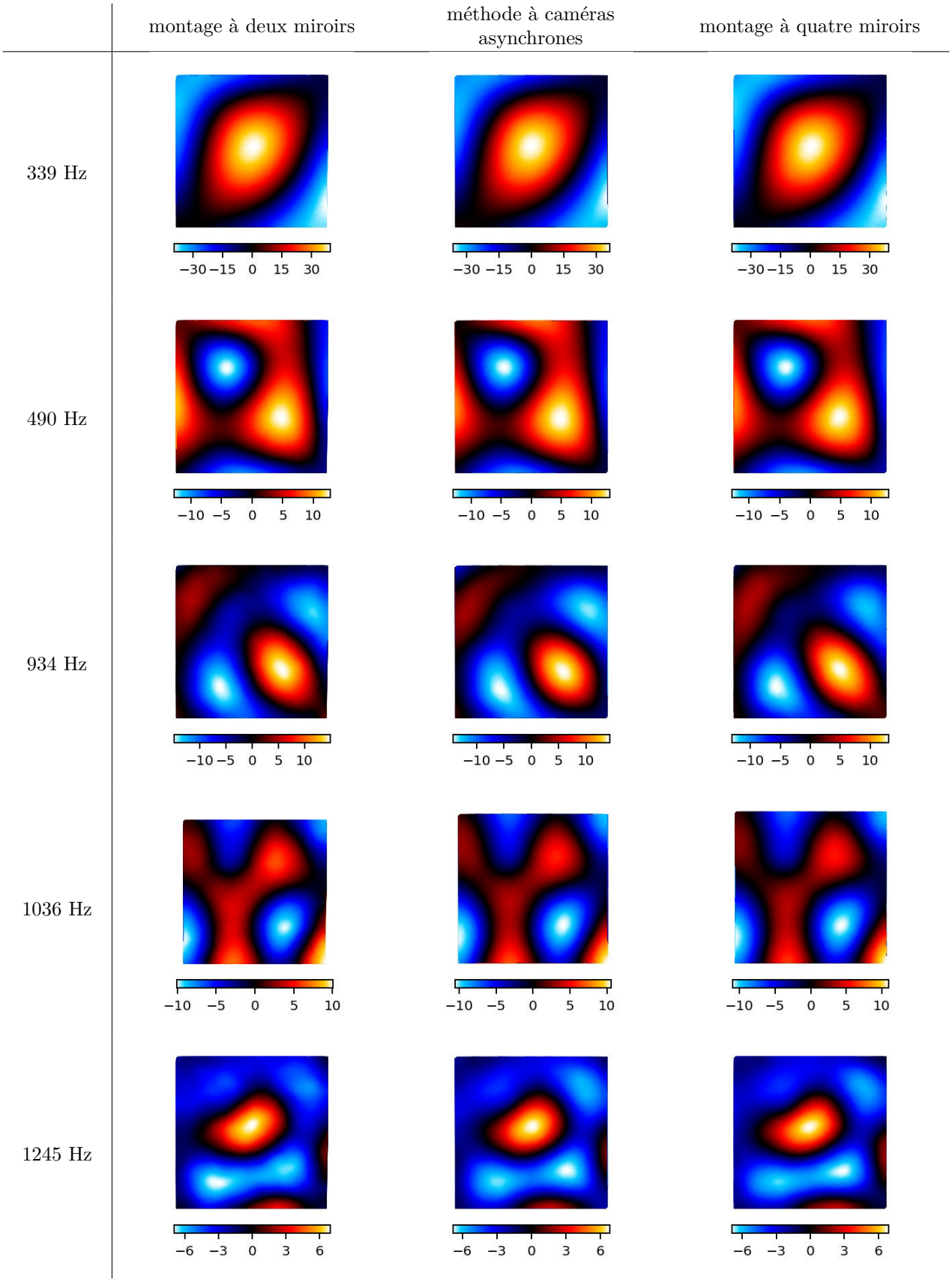


FIGURE 10.22 – Exemples de déformées modales opérationnelles, en µm/N, calculées avec (*gauche*) le montage à deux miroirs, (*milieu*) la méthode à caméras asynchrones et (*droite*) le montage à quatre miroirs ; de *haut* en *bas* à 339 Hz, 490 Hz, 934 Hz, 1036 Hz et 1245 Hz

Perspectives

Dans le prolongement de ce manuscrit, un certain nombre de travaux sont envisageables. Tout d'abord un projet pourrait cibler l'automatisation, couplée éventuellement à la simplification, de la création du motif aléatoire utilisé. On pourrait ainsi envisager la création de pochoirs avec une imprimante 3D par exemple ou encore la projection d'un mouchetis fait de lumière moyennant une adaptation des codes, puisque le mouchetis serait mobile par rapport à la surface de mesure. Ensuite, une phase de développement pourrait s'intéresser à l'automatisation de toutes les étapes du protocole de mesure, avec notamment l'écriture de codes pour la détection automatique des cibles des mires de calibrage. À terme, une interface ou un logiciel pourraient être conçus, automatisant alors complètement l'utilisation des méthodes.

Rappelons que dans le cadre du travail présenté, l'étape finale a volontairement été fixée à la visualisation des déformées modales, le sujet de cette thèse n'étant pas l'analyse modale *per se*. L'exploitation des données est néanmoins tout à fait possible pour l'analyse, les mesures de vibrations étant souvent utilisées pour remonter aux paramètres des objets de mesure. Dans cette logique d'extraction de paramètres, un projet, en cours au moment de la rédaction de ce manuscrit, dans le cadre du Master IMDEA, s'intéresse à la caractérisation et l'estimation de paramètres linéaires et non-linéaires d'un haut-parleur avec la méthode mono-axiale sans miroir. De même, un projet d'étudiants au sein de l'ESGT (en cours en 2018-2019) cible la mesure de vibration d'une passerelle piétonne, un des objectifs étant, à terme, d'effectuer la mesure de fréquences propres de ponts et de structures élancées. On pourrait aussi s'intéresser au calcul du champ acoustique rayonné par la surface à partir du champ de déplacement et de la forme mesurés, ou encore coupler les méthodes vison de présentées avec d'autres approches, la méthode RIFF [119] par exemple, afin d'extraire diverses informations pour des poutres ou des plaques.

En conclusion, les travaux effectués ont finalement montré tout le potentiel des méthodes de vision 3D conçues ainsi que leur applicabilité aux problématiques industrielles. Dans le cadre de l'électro-acoustique, diverses applications R&D pourraient être testées sur des haut-parleurs. Dans le domaine de la vibro-acoustique, d'autres objets d'étude pourraient être choisis : soit de grands objets pour la caractérisation de panneaux, soit des tables d'harmonie ou des anches d'instruments à vent, ou encore des trous noirs acoustiques. Enfin, couplée par exemple avec une caméra thermique, la méthode à caméras asynchrones permettrait très probablement d'obtenir une vue très complète du comportement dynamique de poutres, de plaques ou potentiellement de panneaux.

Liste des images

Remerciements : http://2.bp.blogspot.com/-O4xT6CjCZ7o/TbJmXaDoJhI/AAAAAAAAAg4/M_-lwDHX_Ck/s1600/Mount-Everest-Wallpapers.jpg

Avant Propos : http://www.beliefnet.com/columnists/haveamagnificentday/files/2014/11/caldron-1.jpg

Introduction : https://www.itespresso.fr/wp-content/uploads/2014/03/marketing-cible-internautes.jpg

Chapitre 1 : https://altered-states.net/barry/newsletter420/tuningforkresonance.jpg

Chapitre 2 : https://en.wikipedia.org/wiki/File:Vitruvian.jpg

Chapitre 5 : http://www.ucciani-dessins.com/procedure-complexe/ avec l'autorisation de Jean-Michel Ucciani

Chapitre 7 : https://commons.wikimedia.org/wiki/File:Purple-Iris_pn.jpg

Chapitre 8 : http://www.photos-public-domain.com/2011/09/14/simplify/

Chapitre 9 : http://etcp.fr/wp-content/uploads/2016/03/iStock_000016521178Large.jpg

Conclusion générale :
http://ms-mci.audencia.com/wp-content/uploads/2016/11/Audencia-ecole-management-mastere-specialise-master-MCI-cible-2.png http://www.lumieresdaltitude.com/banque-images.php?gallery_id=61&image_id=4248#.W-wYQhzjKV4

Annexe A

Liste des publications

Revue internationale à comité de lecture

1. T. Durand-Texte, E. Simonetto, S. Durand, M. Melon & M.-H. Moulet, "Vibration measurement using a pseudo-stereo system, target tracking and vision methods", Mech. Syst. Sig. Process., 118C, 30–40 (2019) doi :10.1016/j.ymssp.2018.08.049.

Revue nationale à comité de lecture

1. T. Durand-Texte, E. Simonetto, S. Durand, M. Melon & M.-H. Moulet, "Estimation des incertitudes d'une méthode de mesure de déformées vibratoires par vision 3D", Instrumentation, Mesure, Métrologie, Vol. 16 (1-4), 71-98 (2017)

Conférences internationales

1. T. Durand-Texte, M. Melon , E. Simonetto, S. Durand & M.-H. Moulet, "Vision methods applied to measure the displacement of loudspeaker membranes", 144th AES Convention, Milan, Italy, May 2018.
2. T. Durand-Texte, M. Melon , E. Simonetto, S. Durand & M.-H. Moulet, "3D vision method applied to measure vibrations of non-flat items with a two-mirror adapter", 13th international conference on vibration measurements by laser and non-contact technique, Ancona, June 2018.

Conférence nationale

1. T. Durand-Texte, E. Simonetto, S. Durand, P. Picart, M.-H. Moulet & M. Melon, "Mesure de déformations vibratoires par vision 3D", 13e Congrès Français d'Acoustique/VISHNO, Le Mans (2016).

Colloques nationaux

1. T. Durand-Texte, E. Simonetto, S. Durand, M.H. Moulet & M. Melon, "La vision 3D appliquée à la mesure de vibrations basses fréquences", Workshop TEOBF de la SFA, Marseille, 17-19 janvier 2018.
2. T. Durand-Texte, M. Melon, E. Simonetto, S. Durand et M.-H. Moulet, "Mesure de vibrations par caméra ultra-rapide", Journée mesures de champs en dynamique des structures (JMDC 2018), Supmeca Saint-Ouen, 20 juin 2018.

Annexe B

Acte de la 144e convention de l'AES (Audio Engineering Society)

Audio Engineering Society

Convention Paper

Presented at the 144th Convention
2018 May 23 – 26, Milan, Italy

This convention paper was selected based on a submitted abstract and 750-word precis that have been peer reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This convention paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Vision methods applied to measure the displacement of loudspeaker membranes

Thomas Durand-Texte[1], Manuel Melon[1], Elisabeth Simonetto[2], Stéphane Durand[2], and Marie-Hélène Moulet[3]

[1]*LAUM UMR CNRS 6613, Le Mans Université, Avenue olivier Messiaen, 72 085 Le Mans cedex 9, France*
[2]*Laboratoire Géomatique et Foncier, CNAM, 1 Boulevard Pythagore, 72000 Le Mans, France*
[3]*Centre de Transfert de technologie du Mans, 20 Rue de Thalès de Milet, 72000 Le Mans, France*

Correspondence should be addressed to Manuel Melon (`manuel.melon@univ-lemans.fr`)

ABSTRACT

Increased interest has been witnessed over the past decades for techniques measuring the vibration of a loudspeaker membrane. In this proceeding, vision methods have been adapted to measure the displacement of the cones of low-frequency drivers. The movement of the membrane is recorded with industrial or consumer market high-speed cameras, with frame rates greater than or equal to 240 fps. The measured displacement shows acceptable to very good agreement with the one measured by a laser vibrometer, depending on the camera model. The displacement of the membrane, coupled to the measured electrical impedance, can be used to retrieve the small signal parameters or some non-linear parameters of the speaker. Finally, the vision methods are used to retrieve the vibration patterns of the membrane.

0 Introduction

Since the work by N. Thiele [1] on electro-dynamic loudspeakers, many measurement methods have been developed to retrieve the standard Thiele & Small parameters. Some are based on the reciprocity method [2] or on the measurement of the electrical impedance of the speaker [1]. The latter technique has been extensively used, as anechoic conditions are not required in that case and as voltage measurement can be easily performed [3, 4, 5].

In parallel with the development of measurement methods for loudspeakers, optics-based methods have been increasingly used in structural acoustics [6]. For electro-acoustics, the use of a laser sensor to measure the displacement of the membrane gave insight into the non-linear behaviour of loudspeakers [7]. Optical holography [8] or laser measurements [9, 10, 11, 12] have also been implemented to measure the vibration patterns of membranes. More recently, Wang *et al.* [13] proposed to use fringe projection coupled to a high-speed camera to measure the vibration modes of membranes up to around $1,300$ Hz.

High-speed cameras coupled to 3D vision methods have also been applied to measure vibrations. One of the most widely used techniques for full-field 3D displacement measurement is the Stereo Digital Image Correlation method (SDIC), which requires two points of view that can be obtained from two high-speed cam-

eras to perform vibration measurement. In order to reduce the cost and to avoid problems due to camera synchronisation, a four-mirror set-up can be used and leads to a single camera pseudo-stereo system [14] for which the camera sensor is split into two halves, each half corresponding to a different point of view [15, 16, 17, 18]. Recently, using the four-mirror set-up, a single high-speed camera and a target tracking method, the vibrations of a speaker membrane have been measured and the first results reported [19]. The 3D vision data showed good agreement with measurement data obtained with a Doppler laser vibrometer. Along with these encouraging results, some questions arose: is it possible to simplify the measurement set-up by removing the mirrors and still using a single camera? Is it possible to replace an expensive professional high-speed camera by a low-cost product from the consumer market? These questions will be addressed in this paper.

The first section describes the vision methods used to measure the displacement of the membrane of an electro-dynamic speaker. The second section is dedicated to the validation of the proposed single-axis vision method, by comparing the results to those obtained with a single-point laser. The third section applies the vision methods to measure the characteristics of loudspeakers. Finally, the results are discussed in the last section.

1 Method

3D vision methods can be used to measure the 3D locations of points in space. With at least two different points of view of these points, recorded by cameras with known positions and orientations, the triangulation process can be performed and results in establishing their 3D locations. These methods are nowadays widely used in the industry to measure the shape and the deformations of mechanical parts, either to assess the quality of a part or to measure stress [20]. Recently, with high-speed cameras, videos showing deformations have been recorded at high frame rates, which has paved the way for vibration measurement.

1.1 Principle with two cameras

In order to perform triangulation, the first information required is to know how a 3D point from the real world is projected onto the pixel matrix of a camera. If the location of the 3D point $(X,Y,Z)^T$ is expressed in the coordinate system of the camera, the 2D projected point $(u,v)^T$ in the image, expressed in pixels, is:

$$\begin{pmatrix} uZ \\ vZ \\ Z \end{pmatrix} = K \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \tag{1}$$

with

$$K = \begin{bmatrix} f_u & 0 & p_u \\ 0 & f_v & p_v \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

f_u and f_v are the focal lengths along the $\vec{u}$ and $\vec{v}$ axes and $(p_u,\ p_v)$ is the optical centre of the objective of the camera in the image. These values, coupled with the distortion coefficients, correspond to the 'intrinsic parameters' of a camera.

In equation (1), Z is unknown, but the 3D location $(X,Y,Z)^T$ is necessarily on the line going through the camera centre and the image of the point $(X,Y,Z)^T$. Figure 1 presents the principle in the $(\vec{y},\vec{z})$ plane.

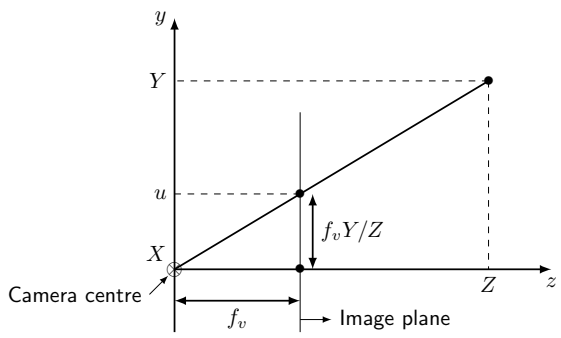


Fig. 1: Pinhole camera model (adapted from [21])

However, the 3D coordinate systems of two different cameras are inherently not aligned. A rotation matrix R_i (3x3) and a translation vector T_i (3x1) (called the 'extrinsic parameters' of a camera) are defined for each camera to switch from the world coordinates to the camera coordinates[1]. Thus, if the coordinate system shift is included, equation (1) becomes:

$$(ku_i, kv_i, k)^T = K_i R_i((X,Y,Z)^T + T_i). \tag{3}$$

Note that k is the distance from the point $(X,Y,Z)^T$ to the camera centre.

The 2D location of a point in the image of the camera defines a line in the 3D coordinate system. With two

[1]Here K_1, R_1 and T_1 correspond to the left image and K_2, R_2 and T_2 to the right one in figure 3.

cameras and two corresponding 2D locations, two 3D lines are defined, which allows triangulating the 3D location of the point $(X,Y,Z)^T$.

The triangulation method used here first remaps the point (u_1,v_1), in the pixel matrix of the first camera, to the world 3D coordinate system by reversing equation (3):

$$\begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} = R_1^{-1}K_1^{-1}\begin{pmatrix} z_{cam}u_1 \\ z_{cam}v_1 \\ z_{cam} \end{pmatrix} - T_1, \qquad (4)$$

with z_{cam} the distance from the camera centre.

This 3D point is then reprojected onto the pixel matrix of the second camera with equation (3). The triangulation process aims at finding the value of z_{cam} which minimises the distance between the point reprojected from the first matrix and the corresponding point in the second matrix, thus allowing retrieving the 3D location $(X,Y,Z)^T$. Figure 2 shows the triangulation principle[2].

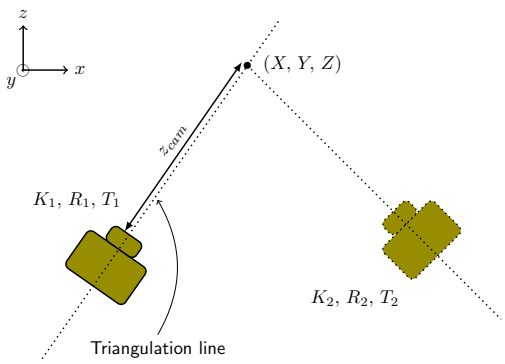


Fig. 2: 3D measurement set-up with two cameras

The intrinsic and extrinsic parameters are obtained by a calibration process, using several images of a flat calibration target (see figure 3). The calibration of the camera should be performed first, to obtain the intrinsic parameters for each camera, and the calibration of the stereo system subsequently, while setting the intrinsic parameters in the optimisation process. For more details, see [20, 22]. Here, the measurement points are Gaussian point targets (see figure 3) and are paired manually.

[2]Other methods exist to perform triangulation, see [21, 20].

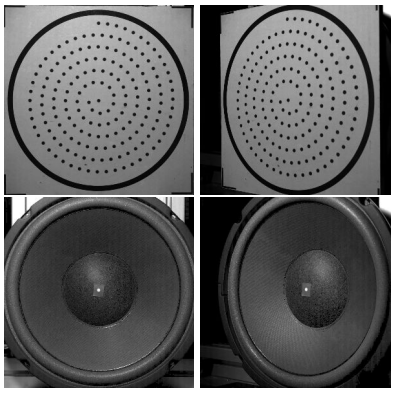


Fig. 3: Images used for stereo calibration and initial triangulation. Left: camera 1 (front); right: camera 2 (side); upper: calibration target; lower: speaker

1.2 Single camera method

High-speed cameras are currently very expensive and difficult to synchronise neatly in order to measure the 3D displacement of a point. The usual solution to go round this difficulty consists in using a four-mirror adapter in order to generate two virtual cameras from a real one, at the expense of the number of pixels available for each view. Nevertheless, in principle, loudspeakers are designed to move along a single axis. Therefore, is it possible to measure the displacement of a loudspeaker membrane with only one camera and no mirrors? If the initial 3D position of a point is known, and if the displacement is assumed to be along a single axis, a known triangulation line is sufficient to perform triangulation with only one camera (see figure 4), which leads to design a single-camera single-axis vision method.

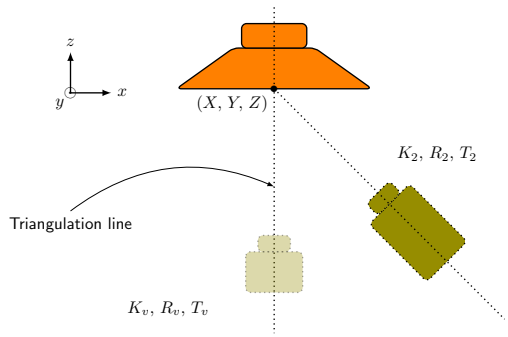


Fig. 4: 1D measurement set-up with one camera

So, the method proposed here consists in defining a virtual camera and the 2D location of a point in order to perform, with only one real camera, the triangulation of the single-axis displacement signal of a measurement point. This measurement point is chosen at the centre of the cone in order to assess the relevance of the technique. Note that, if the displacement occurs in several directions, errors will be necessarily induced.

In order to define the position of the virtual camera, the 3D position of a point is required, when the membrane of the loudspeaker is at rest. This position is obtained with two images (four with calibration, see figure 3), acquired before performing vibration measurement. The line in 3D space, usually obtained with the first camera and the 2D location of a point, is assumed to go through the measured 3D position. For simplicity, the first position of the calibration target defines the world 3D coordinate system and is set so that the $\overrightarrow{z}$ axis may be as aligned as possible with the displacement axis of the speaker (figures 3 & 4).

So, K_v and R_v are set as (3x3) identity matrices, the 2D point (u_v, v_v) is set to $(0,0)$ and the translation vector T_v to:

$$T_v = -(X,Y,Z)^T. \tag{5}$$

The triangulation line in space is thus parallel to the $\overrightarrow{z}$ axis and goes through the initial 3D world coordinate system of the measurement point, as shown in figure 4.

The targets are first detected with the dual conic detector [23]. The estimated ellipses are then used to define an area of pixels to perform pixel and sub-pixel [24] 2D correlation in order to track the displacement of each target in the images.

Finally, the triangulation process is performed with the intrinsic and extrinsic parameters of the virtual and real cameras, the 2D locations (u_v, v_v) and (u_2^i, v_2^i) of the point at each time sample i (for each image), which results in a single-axis displacement signal.

2 Validation

2.1 Professional high-speed camera

In order to assess the validity of the single-axis vision method, the displacement of a loudspeaker membrane is measured for a single point, both with a triangulation laser vibrometer (Polytec OFV 302 laser with a controller OFV3000) and a professional high-speed camera (Photron SA-X2). The Frequency Response Function (FRF), displacement over voltage X/U, is measured for a woofer (AB Sound TW1041). The excitation signal is a swept sine (1 Hz-480 Hz over 10 seconds, $\pm$700 mV). The voltage at the terminals of the speaker, and the electrical current going through, are measured during the acquisition of the images. The frame rate of the camera is set to 1,000 fps and the sampling frequency of the laser to 1,000 Hz.

A picture of the measurement set-up is shown in figure 5. The speaker is fastened to a stand, and placed on a plate with dampers to avoid the transmission of vibrations. A 185 W LED spot is used to illuminate the speaker and to provide optimal lighting for each image at 1,000 fps. The high-speed camera is placed aside with an angle of approximately 40° from the $\overrightarrow{z}$ axis to minimise the impact of the noise in the image and increase the visibility of the displacement [19].

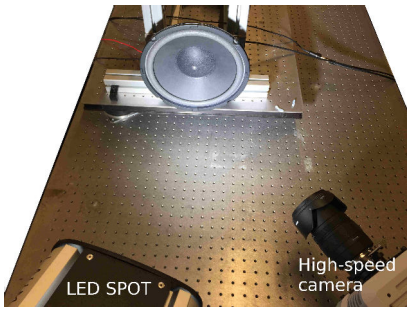


Fig. 5: Picture of the measurement set-up

Electrical and vision measurements have to be synchronised. Here, this synchronisation is achieved with signal processing, by adjusting the phase of the FRF between the displacement signal (X) and the electrical current signal (I). The resulting FRF X/I is shown in figure 6.

The curves representing the measured FRF X/U with the single-axis vision and laser methods are drawn in figure 7 and show good agreement up to 200 Hz.

2.2 Low-cost high-speed cameras

Professional high-speed cameras give good results; what about low-cost ones? The FRF X/U is thus measured with the camera of an Iphone 6 and a GoPro Hero 5 Black, with a frame rate of 240 fps. The swept sine is adjusted to avoid aliasing (1 Hz - 120 Hz over 5 seconds). The curves, displayed in figure 8, show more noise.

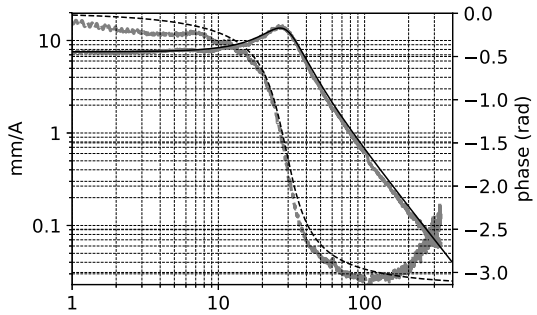


Fig. 6: FRF X/I (plain lines: magnitude; dashed lines: phase; **gray**: measured; **black**: estimated)

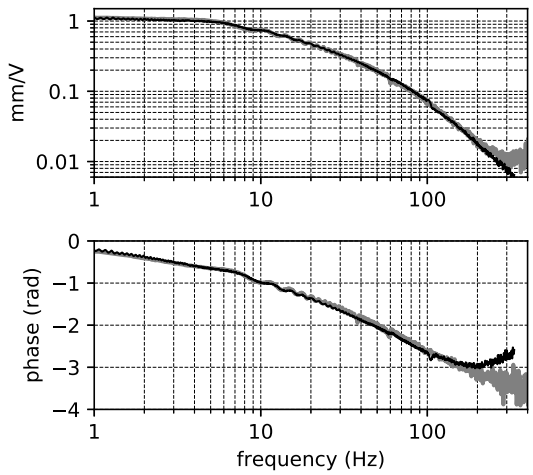


Fig. 7: Comparison of the measured FRFs with a professional high-speed camera and a triangulation laser vibrometer (**gray**: laser; **black**: vision)

2.3 Discussion

The results obtained with the single-axis vision method match well those obtained, for a single measurement point, with the triangulation laser vibrometer, which validates the method. It may be argued that a displacement laser sensor is easier to set up, does not require such heavy processing and is *per se* designed for single-point measurements. Indeed, this is true. Yet, the single-axis vision method can also provide full-field measurements with several targets or a speckle pattern, unlike the triangulation laser. It can thus be fruitfully used to measure relatively easily some vibration patterns of many items, such as plates, the body of a guitar

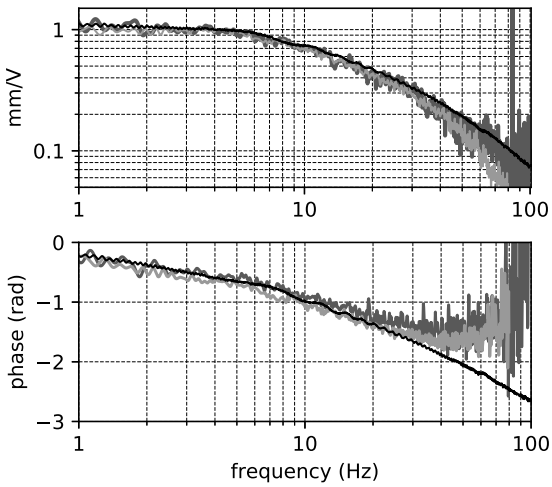


Fig. 8: Comparison of the measured FRF X/U with a professional high-speed camera and mass market cameras (**black**: high-speed Photron SAX-2; **dark gray**: Iphone6; **light gray**: GoPro Hero 5 Black)

or tambourines, whose displacements are locally along a single axis.

As regards the cameras used, the professional high-speed one logically gives better results: indeed, the measurements with the mass market ones are visibly more noisy, which is most probably due to the automatic image-processing of these devices. Yet, the corresponding FRF curves remain close enough in the low frequency band. These low-cost cameras may thus be sufficient for some applications. Please note that mass market cameras with 1,000 fps, for instance the Sony RX100, have recently been launched on the market. These devices should be relevant low-cost options for vibration measurements.

3 Application to a loudspeaker

The Thiele & Small parameters are usually calculated with at least two electrical measurements. However, with a displacement measured with a triangulation laser vibrometer, these parameters can be retrieved with only one electrical measurement, and some non-linear parameters can also be estimated. The single-axis vision method allows doing the same, but presents the advantage of being a full-field measurement technique, which can thus be used to recover some vibration patterns.

3.1 Estimation of the linear parameters

The single-axis method is tested on a low-frequency loudspeaker Visaton 300 WS, for a single point. The excitation signal is a swept sine, ±700 mV, ranging from 1 Hz to 480 Hz in 10 seconds. A professional high-speed camera Photron SAX-2 is used.

The model used for the impedance of the speaker is:

$$Z_{spkr} = \frac{Bl^2}{Z_m} + Z_e, \tag{6}$$

with Z_{spkr} the impedance of the speaker, Z_m the mechanical impedance and Z_e the electrical impedance of the voice coil. The radiation impedance of the membrane is included in Z_m. Using the effective inductance model from Leach [25, 26], the measured electrical impedance of the speaker allows estimating Z_e:

$$Z_e = Re + j\omega L_{eff}(f). \tag{7}$$

Moreover, from the mechanical coupling equation of the electro-dynamic speaker, we get:

$$\frac{j\omega X}{I} = \frac{Bl}{Z_m}. \tag{8}$$

So, from equations (6) and (8), the force factor Bl can be obtained with:

$$Bl = \frac{Z_{spkr} - Ze}{j\omega X/I}. \tag{9}$$

The curves $Z_{spkr} - Z_e$ and $Bl * j\omega X/I$ (corresponding to Bl^2/Z_m) displayed in figure 9 show good agreement.

When multiplying the curve $I/j\omega X$ by Bl, an estimation of the mechanical impedance of the speaker Z_m is obtained. From the imaginary part of this impedance, the value of the mechanical compliance C_{ms} is then estimated with the Least Mean Square method (figure 10).

Finally, with the values of Bl, C_{ms} and the electrical impedance curve, all the linear parameters can be estimated, which allows simulating the full impedance of the speaker (figure 11).

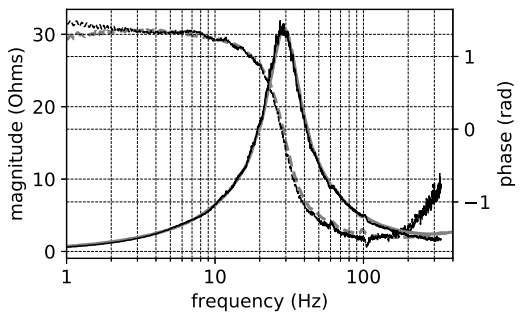


Fig. 9: Curves used to measure the force factor Bl (plain lines: magnitude; dashed lines: phase; **gray**: $z_{spkr} - Z_e$; **black**: $Bl * j\omega X/I$)

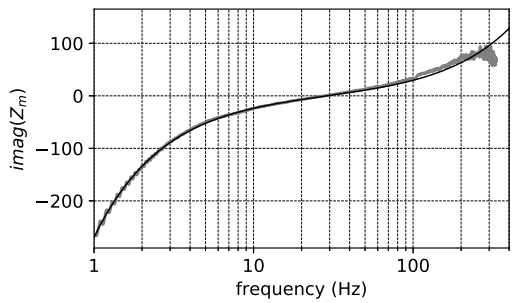


Fig. 10: Imaginary part of the mechanical impedance (**gray**: measured; **black**: estimated)

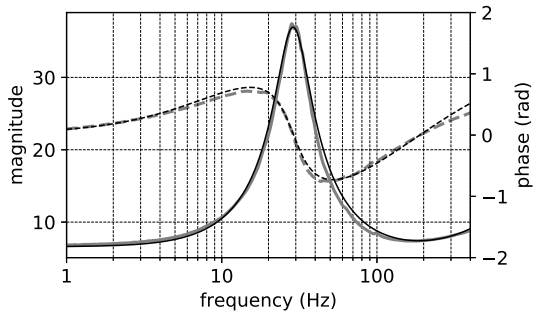


Fig. 11: Electrical impedance of the speaker (plain lines: magnitude; dashed lines: phase; **gray**: measured; **black**: estimated by the single-axis vision method)

3.2 Estimation of the non-linear parameters

When moving the membrane mechanically, the voltage measured at the terminals of the speaker is equal to

the force factor multiplied by the velocity: $U_{spkr} = Blv$. The position of the membrane is measured over time, and the corresponding velocity is calculated. The value of the force factor as a function of excursion $Bl(x)$ is then estimated.

The results obtained with the single-axis vision method, for a single point, are compared with those obtained with the R&D Klippel System in figure 12. However different these methods may be (regarding the sensors, excitation signals or models used), the results are similar even if a slight discrepancy can be observed, mainly in the positive direction. Additional tests should be performed to investigate the accuracy of the proposed method.

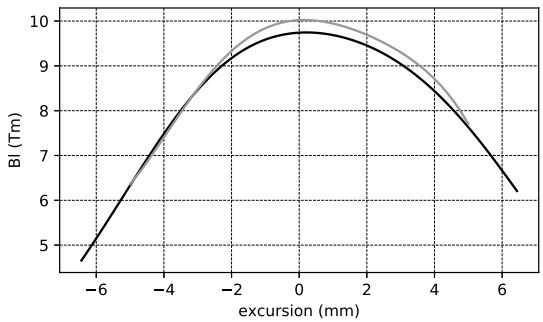


Fig. 12: Estimation of the non-linearity of the force factor (**black**: estimated with the R&D Klippel System; **gray**: estimated with the single-axis vision method)

3.3 Multipoint measurements

The previous section validates the single-axis vision method applied to measure the 1D displacement of a single point. However, the technique allows tracking several points simultaneously, resulting in full-field measurement. The object of study is an electrodynamic speaker. Its main displacement is along the axis of the voice coil (*i.e.* in a single direction); but vibration modes occur in 3D along the shape normals and its rocking modes induce rotations.

Although a 3D method is logically required to get accurate data on the displacement of the membrane, one may wonder whether the single-axis vision method can give a relevant overview or relevant initial information on the vibration patterns of a low-frequency loudspeaker membrane. This section is thus dedicated to a quick comparison of the results obtained with the single-axis and the 3D vision methods.

The 3D measurement is performed with the pseudo-stereo system, designed with a single professional high-speed camera and a four-mirror adapter [14]. In both cases, the loudspeaker used is a AB Sound TW1041, excited with a sinusoidal burst (1 period, 3 ms long, with 19 V of amplitude). A speckle pattern drawn on the membrane is used to perform image correlation with SDIC tools [20], leading to full-field shape and vibration measurements. Figure 13 shows the set-up of the pseudo-stereo system and the image obtained.

Fig. 13: Picture of the set-up (upper) and image obtained (lower) with the pseudo-stereo system

The processing of the measured data allows calculating the standard deviation of the FFT along the corresponding axes of space and results in retrieving the frequencies of the operational modal shapes, which gives insight into the 1D/3D displacement of the membrane of the speaker. Figures 14 and 15 show the standard deviations of the FFTs and two operational modal shapes are shown in figures 16 and 17 for a rocking mode and a membrane mode. The displayed data correspond to the absolute value of the displacement, in the frequency domain, signed with the phase in the $\overrightarrow{z}$ direction.

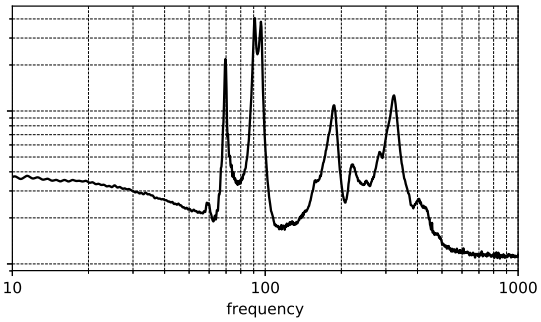


Fig. 14: Standard deviation of the displacement FFT for each frequency with the single-axis vision method

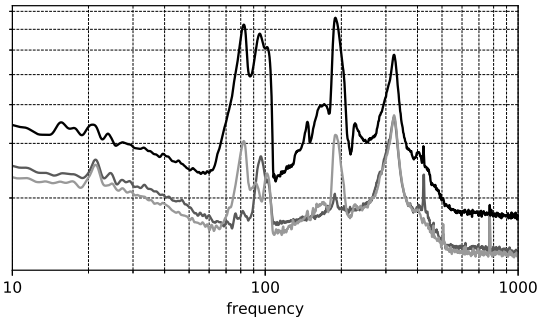


Fig. 15: Standard deviation of the displacement FFT for each frequency with the 3D vision method (**dark gray**: X displacement; **light gray**: Y displacement; **black**: Z displacement)

3.4 Discussion

With electrical and displacement measurements, all the linear and some non-linear parameters of loudspeakers can be retrieved. The results presented demonstrate the validity of the single-axis vision method, even if, for single-point measurements, lasers are easier to set up and seem better suited.

Regarding multipoint measurements, the results obtained with the 3D vision method clearly indicate that the displacement of the membrane occurs along the three axes of space, which implies that the single-axis method necessarily gives incomplete information. However, the standard deviation of the FFTs calculated from the data obtained with the single-axis vision method also allows retrieving the frequencies of

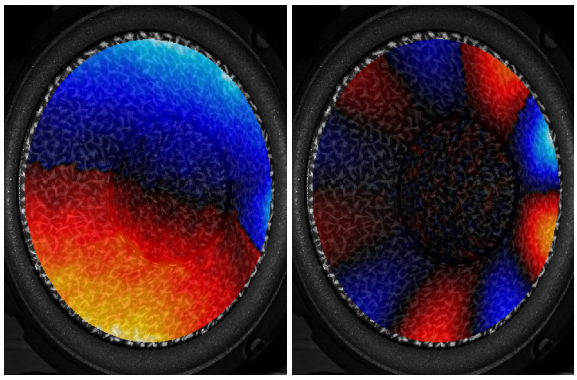


Fig. 16: Operational modal shapes with the single-axis vision method, modulus signed with the phase: at 188 Hz (left) and 322 Hz (right)

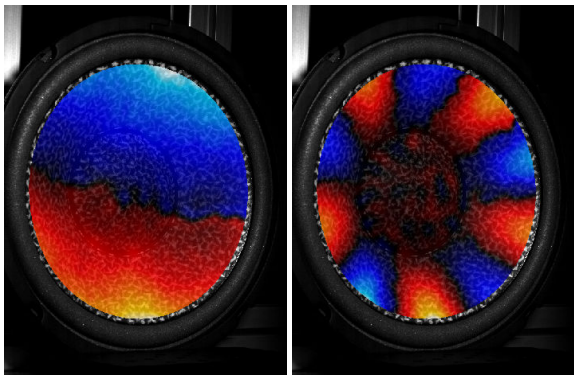


Fig. 17: Operational modal shapes with the 3D vision method, |X,Y,Z| signed with the phase in the $\vec{z}$ direction: at 188 Hz (left) and 323 Hz (right)

the vibration patterns. It can thus fruitfully be used to easily get initial information on the behaviour of a loudspeaker membrane. Note that, even if the results between both vision methods are globally in agreement, errors in amplitude are visible for the curves and shapes from the data from the single-axis method. Nevertheless, for many items whose displacement occurs exclusively in one direction, the proposed method can be relevantly applied.

4 Summary

In this paper, a single-camera single-axis vision method, with no mirrors, has been proposed in an attempt to simplify the now classical single-camera pseudo-stereo

system, which requires a four-mirror adapter and is more complex to operate. It is designed to measure 1D displacements and was tested on the single-point measurement of a loudspeaker membrane. The results were compared to those obtained with a single-point triangulation laser and show good agreement, which validates the method.

The displacement measured with the single-axis vision method and a professional high-speed camera coupled to an electrical impedance measurement was used to retrieve the Thiele & Small parameters of a woofer. The variation of the *Bl* factor was also calculated and showed good agreement with the one measured with the R&D Klippel system. These results highlight the potential of the vision method to obtain the linear and non-linear parameters of a loudspeaker. It should be borne in mind that the vision method requires a calibration procedure of both the camera and the set-up ($\simeq$ 10 minutes), that can be tedious for people familiar with laser sensor measurements.

Now, vision methods provide full-field data, *i.e.* allow measuring several points simultaneously, and are thus inherently well-suited for (transient) vibration measurements. The single-axis vision method may be fruitfully used on many items, whose displacements are locally along a single axis, for example plates, tambourines... In the case of a loudspeaker, the complexity of the vibration patterns requires a 3D measurement technique to get accurate results. For example, the single-camera pseudo-stereo system, with its four-mirror adapter, allows recovering 3D vibration patterns, such as the rocking modes or the operational modal shapes of the loudspeaker membrane. The results obtained with the single-axis method remain close to those obtained with the 3D method even if errors occur in the amplitude of the data, which is due to the very nature of the multiaxial displacement of the speaker membrane. Nevertheless, this single-axis method offers a relevant overview or relevant initial information on the behaviour of the object of study, and may even be sufficient for some applications.

The camera used for the vision methods is generally a professional high-speed one, whose expensive price remains a major drawback of the approach. When a mass market high-speed camera is used, a much lower signal-to-noise ratio can be observed, most probably due to the automatic and not user-controllable image-processing performed in such devices. Yet, the results obtained with these low-cost cameras are in agreement with the data obtained with a triangulation laser vibrometer. One may expect that, in the very near future, new references will be available on the market, at a reasonable price ($\leq$ 1000 \$) with higher frame rates and more controllable parameters. These devices will thus be potentially effective alternatives.

Another limitation is the minimum displacement z_{min} that can be measured by the vision method. With $z_{min} \simeq 10\ \mu$m, subwoofers and woofers seem good candidates for the proposed method, whose high frequency limit is estimated to be around 500 Hz for an input voltage around a few volts.

Finally, the challenge is to print the random speckle pattern on the membrane. A possible solution to avoid this step could be to project a pattern using a beamer. Future work will consist in adapting the algorithm to this configuration and perform the first tests, on items whose displacement is along a single axis, in order to explore the potentialities of the approach.

Acknowledgement

The authors wish to thank Antonin Novak for fruitful discussions about the non-linear characterization.

References

[1] Thiele, N., "Loudspeakers in Vented Boxes: Part 2," *J. Audio Eng. Soc*, 19(6), pp. 471–483, 1971.

[2] Merhaut, J., "Low Frequency Measurement of Loudspeakers by the Reciprocity Method," *J. Audio Eng. Soc*, 30(12), pp. 882–888, 1982.

[3] Cabot, R. C., "Automated Measurements of Loudspeaker Small Signal Parameters," in *Audio Engineering Society Convention 81*, 1986.

[4] Struck, C. J., "Determination of the Thiele-Small Parameters Using Two-Channel FFT Analysis," in *Audio Engineering Society Convention 82*, 1987.

[5] d'Appolito, J., *Testing Loudspeakers*, Audio Amateur Press, 1998.

[6] Baqersad, J., Poozesh, P., Niezrecki, C., and Avitabile, P., "Photogrammetry and optical methods in structural dynamics - A review," *Mechanical Systems and Signal Processing*, 86, pp. 17–34, 2017.

[7] Seidel, U. and Klippel, W., "Fast and Accurate Measurement of the Linear Transducer Parameters," in *Audio Engineering Society Convention 110*, 2001.

[8] Frankort, F. J. M., "Vibration Patterns and Radiation Behavior of Loudspeaker Cones," *J. Audio Eng. Soc*, 26(9), pp. 609–622, 1978.

[9] Struck, C. J., "Analysis of the Nonrigid Behavior of a Loudspeaker Diaphram Using Modal Analysis," in *Audio Engineering Society Convention 86*, 1989.

[10] Klippel, W. and Schlechter, J., "Measurement and Visualization of Loudspeaker Cone Vibration," in *Audio Engineering Society Convention 121*, 2006.

[11] Klippel, W. and Schlechter, J., "Distributed Mechanical Parameters Describing Vibration and Sound Radiation of Loudspeaker Drive Units," in *Audio Engineering Society Convention 125*, 2008.

[12] Bellini, M. C., Collini, L., Farina, A., Pinardi, D., and Riabova, K., "Measurement of Loudspeakers with a Laser Doppler Vibrometer and the Exponential Sine Sweep Excitation Technique," *J. Audio Eng. Soc*, 65(7/8), pp. 600–612, 2017.

[13] Jian, W., Yong, L., Zhang, Z., and Gao, P., "Cost-effective and full-field method for measuring vibration of loudspeaker membrane using fringe projection," in *Proc. SPIE 9276, Optical Metrology and Inspection for Industrial Applications III*, volume 9276, pp. 9276 – 9276 – 8, 2014.

[14] Inaba, M., Hara, T., and Inoue, H., "A stereo viewer based on a single camera with view-control mechanisms," in *Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 1993, Tokyo, Japan, July 26 - 30, 1993*, pp. 1857–1865, 1993.

[15] Li, J., Dan, X., Xu, W., Wang, Y., Yang, G., and Yang, L., "3D digital image correlation using single color camera pseudo-stereo system," *Optics & Laser Technology*, 95, pp. 1 – 7, 2017.

[16] Yu, L. and Pan, B., "Single-camera high-speed stereo-digital image correlation for full-field vibration measurement," *Mechanical Systems and Signal Processing*, 94, pp. 374–383, 2017.

[17] White, T. G., Patten, J. R. W., Wan, K.-H., Pullen, A. D., Chapman, D. J., and Eakins, D. E., "A single camera three-dimensional digital image correlation system for the study of adiabatic shear bands," *Strain*, 53(3), 2017.

[18] Durand-Texte, T., Simonetto, E., Durand, S., Melon, M., and Moulet, M.-H., "Estimation des incertitudes d'une méthode de mesure de déformées vibratoires par vision 3D," *Instrumentation, Mesure, Métrologie*, 16(1-4), pp. 71–95, 2017.

[19] Durand-Texte, T., Simonetto, E., Durand, S., Melon, M., and Moulet, M.-H., "Vibration measurement using a pseudo-stereo system, target tracking and vision methods," *Submitted to Mechanical Systems and Signal Processing*, 2018.

[20] Sutton, M. A., Orteu, J. J., and Schreier, H., "Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications," 2009.

[21] Hartley, R. I. and Zisserman, A., *Multiple View Geometry in Computer Vision*, Cambridge University Press, second edition, 2004, ISBN 0521540518.

[22] Bradski, G., "Open Computer Vision, OpenCV," *Dr. Dobb's Journal of Software Tools*, 2000.

[23] Ouellet, J.-N. and Hébert, P., "A Simple Operator for Very Precise Estimation of Ellipses," in *CRV '07: Proceedings of the Fourth Canadian Conference on Computer and Robot Vision*, pp. 21–28, IEEE Computer Society, Montréal, Qc, Canada, 2007, ISBN 0-7695-2786-8.

[24] C. Quentin Davis, D. M. F., "Statistics of subpixel registration algorithms based on spatiotemporal gradients or block matching," *Optical Engineering*, 37, pp. 37 – 37 – 9, 1998, doi:10.1117/1.601966.

[25] Leach, W. M., Jr., "Loudspeaker Voice-Coil Inductance Losses: Circuit Models, Parameter Estimation, and Effect on Frequency Response," *J. Audio Eng. Soc*, 50(6), pp. 442–450, 2002.

[26] Dodd, M., Klippel, W., and Oclee-Brown, J., "Voice Coil Impedance as a Function of Frequency and Displacement," in *Audio Engineering Society Convention 117*, 2004.

Bibliographie

[1] R. Anchini, G. d. Leo, C. Liguori, and A. Paolillo. Metrological characterization of a vision-based measurement system for the online inspection of automotive rubber profile. In *2007 IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement*, pages 121–126, July 2007.

[2] W. B. Ng and Y. Zhang. Stereoscopic imaging and reconstruction of the 3d geometry of flame surfaces. 34 :484–493, 04 2003.

[3] M. Bai and J.-H. Lin. Source identification system based on the time-domain near field equivalence source imaging : Fundamental theory and implementation. *J. Sound Vib.*, 307(1-2) :202–225, 2007.

[4] R. Balcaen, P. Reu, P. Lava, and D. Debruyne. Stereo-dic uncertainty quantification based on simulated images. *Experimental Mechanics*, 57(6) :939–951, Jul 2017.

[5] J. Baqersad, T. Lundstrom, C. Niezrecki, and P. Avitabile. Measuring the dynamics of operating helicopter rotors and wind turbines using 3d digital stereophotogrammetry. In *Annual Forum Proceedings - AHS International*, volume 3, pages 2250–2256, 01 2013.

[6] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile. Photogrammetry and optical methods in structural dynamics - A review. *MSSP*, 86 :17–34, Mar. 2017.

[7] V. Bavigadda, E. Mihaylova, R. Jallapuram, and V. Toal. Vibration phase mapping using holographic optical element-based electronic speckle pattern interferometry. *Optics and Lasers in Engineering*, 50(8) :1161 – 1167, 2012. Fringe Analysis Methods & Applications.

[8] E. Bavu and A. Berry. High-resolution imaging of sound sources in free field using a numerical time-reversal sink. *Acta Acustica united with Acustica*, 95(4) :595–606, 2009.

[9] T. Beberniss, T. Eason, and S. Spottswood. High-speed 3d digital image correlation measurements of long-duration random vibration ; recent advancements and noted limitations. In *Proceedings, International Conference on Noise and Vibration Engineering (ISMA), KatholiekeUniversiteit Leuven, Belgium*, 2012.

[10] T. J. Beberniss and D. A. Ehrhardt. High-speed 3d digital image correlation vibration measurement : Recent advancements and noted limitations. *MSSP*, 86 :35 – 48, 2017. Full-field, non-contact vibration measurement methods : comparisons and applications.

[11] R. Bellet, B. Cochelin, P. Herzog, and P.-O. Mattei. Experimental study of targeted energy transfer from an acoustic system to a nonlinear membrane absorber. *Journal of Sound and Vibration*, 329(14) :2768–2791, 2010.

[12] B. Benoît. *Étude physique et modélisation numérique de procédés d'assemblage par soudo-brasage et sous-ensembles en carrosserie automobile.* PhD thesis, Université Pierre et Marie Curie - Paris VI, 2014.

[13] M. Berny, C. Jailin, A. Boutef, F. Hild, and S. Roux. Mode-enhanced space-time DIC : Applications to ultra high-speed imaging. *Measurement Science and Technology*, 29(12) :125008, 2018.

[14] A. Berry and O. Robin. Identification des excitations sur des panneaux par la mesure de leur réponse vibratoire et la méthode des champs virtuels. In *CFA 2016*, Le Mans, France, 2016.

[15] G. Besnard. Caractérisation et quantification de surfaces par stéréocorrélation pour des essais mécaniques du quasi statique à la dynamique ultra-rapide, 2010.

[16] M. Biggio. Mesure de vibrations acoustiques par vision par ordinateur. Rapport de stage, École supérieure des géomètres et topographes - CNAM, 2013.

[17] P. Bing, X. Hui-min, X. Bo-qin, and D. Fu-long. Performance of sub-pixel registration algorithms in digital image correlation. *Measurement Science and Technology*, 17 :1615, 05 2006.

[18] Blender Foundation. Blender a 3D modelling and rendering package. `http://www.blender.org`.

[19] J. Bouguet. Matlab calibration tool.

[20] G. Bradski. Open Computer Vision, OpenCV. *Dr. Dobb's Journal of Software Tools*, 2000.

[21] A. Bright. Vibration behaviour of single-suspension electrodynamic loudspeakers. In *AES 109th Convention*, Los Angeles, California, USA, Septembre.

[22] D. C. Brown. Close-range camera calibration. *Photogramm. Eng.*, 37, 12 2002.

[23] H. A. Bruck, S. R. McNeill, M. A. Sutton, and W. H. Peters. Digital image correlation using newton-raphson method of partial differential correction. *Experimental Mechanics*, 29(3) :261–267, Sep 1989.

[24] S. Bruinsma, D. Di Maio, and T. Tinga. Diagnostics based on continuous scanning ldv methods : numerical study case. In *13th Intl conference on vibration measurements by laser and noncontact techniques*, 2018.

[25] A.-M. Bruneau. Amplitude and phase measurements of vibrations of radiating surfaces in order to determine the emitted sound field. *J. Audio Eng. Soc*, 31(12) :907–913, 1983.

[26] D. M. F. C. Quentin Davis. Statistics of subpixel registration algorithms based on spatiotemporal gradients or block matching. *Optical Engineering*, 37 :37 – 37 – 9, 1998.

[27] C. Cameron, J. Sirohi, and D. Uehara. Transient hub loads and blade deformation of a mach-scale coaxial rotor in hover. In *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 01 2015.

[28] P. Castellini, P. Chiariotti, M. Martarelli, E. Zappa, and A. Lavatelli. Experimental modal analysis on vibration data measured by digital image correlation. In J. M. Harvie and J. Baqersad, editors, *Shock & Vibration, Aircraft/Aerospace, Energy Harvesting, Acoustics & Optics, Volume 9*, pages 285–291, Cham, 2017. Springer International Publishing.

[29] P. Castellini, G. M. Revel, and E. P. Tomasini. Laser Doppler vibrometry : A review of advances and applications. *The Shock and vibration digest*, 30(6) :443–456, 1998.

[30] Y. Chang, K. Hinnen, and C. Varekamp. Accuracy in distance measurements of stereo-camera based head-localization for auto-stereoscopic 3d displays. In *SID Mid-Europe Chapter Spring Meeting on*, pages 78–81, 4 2013.

[31] D. J. Chen, F. P. Chiang, Y. S. Tan, and H. S. Don. Digital speckle-displacement measurement using a complex spectrum method. *Appl. Opt.*, 32 :1839–1849, 1993.

[32] J. G. Chen, N. Wadhwa, Y.-J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk. Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration*, 345 :58–71, 2015.

[33] B. Cochelin, P. Herzog, and P.-O. Mattei. Experimental evidence of energy pumping in acoustics. *Comptes Rendus MÃ©canique*, 334(11) :639 – 644, 2006.

[34] K. Creath and G. Å. Slettemoen. Vibration-observation techniques for digital speckle-pattern interferometry. *J. Opt. Soc. Am. A*, 2(10) :1629–1636, Oct 1985.

[35] J.-M. Desse, P. Picart, and P. Tankam. Digital color holography applied to fluid and structural mechanics. 50 :18–28, 01 2012.

[36] J. Dold. Influence of target size on the results of photogrammetric bundle adjustment. In *ISPR Archives, Technical Comission V*, volume XXXI, Part B5, Vienna, 1996.

[37] M. Dou, J. Taylor, H. Fuchs, A. Fitzgibbon, and S. Izadi. 3d scanning deformable objects with a single rgbd sensor. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 493–501, June 2015.

[38] J.-E. Dufour. *Mesures de forme, de déplacement, et de paramètres mécaniques par stéréo-corrélation d'images isogéométrique.* PhD thesis, Université Paris-Saclay, 2015.

[39] T. Durand-Texte, E. Simonetto, S. Durand, P. Picart, M.-H. Moulet, and M. Melon. Mesure de déformations vibratoires par vision 3d. In *CFA - VISHNO*, Le Mans, France, 2016.

[40] L. Duvieubourg, S. Ambellouis, and F. Cabestaing. *SINGLE-CAMERA STEREOVISION SETUP*, pages 173–178. Springer Netherlands, Dordrecht, 2006.

[41] W. Faig. Calibration of close-range photogrammetry systems : Mathematical formulation. *Photogramm Eng Remote Sens*, 41, 12 1975.

[42] C. R. Farrar, T. W. Darling, A. Migliori, and W. E. Baker. Microwave interferometers for non-contact vibration measurements on large structures. *MSSP*, 13(2) :241 – 253, 1999.

[43] O. Faugeras. *Three-dimensional computer vision : a geometric viewpoint.* MIT press, 01 1993.

[44] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least squares fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(5) :476–480, May 1999.

[45] F. Fooladgar, S. Samavi, S. M. R. Soroushmehr, and S. Shirani. Geometrical analysis of localization error in stereo vision systems. *IEEE Sensors Journal*, 13(11) :4236–4246, Nov 2013.

[46] P. F. Forman. The zygo interferometer system. volume 0192, pages 41–48, 1979.

[47] W. Forstner and E. Gulch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Intercommission Conf. on Fast Processing of Photogrammetric Data*, pages 281–305, 1987.

[48] F. Frank and J. Walker. *Advanced Applications in Acoustics, Noise and Vibration.* CRC Press, 2004.

[49] D. Garcia. *Mesure de formes et de champs de déplacements tridimensionnels par stéréo-corrélation d'images.* PhD thesis, Institut National Polytechnique de Toulouse - INPT, 2001.

[50] D. Garcia, J.-J. Orteu, L. Robert, B. Wattrisse, and F. Bugarin. A generic synthetic image generator package for the evaluation of 3d digital image correlation and other computer vision-based measurement techniques. 05 2013.

[51] O. Gendelman, L. Manevitch, A. F. Vakakis, and R. M'closkey. Energy pumping in nonlinear mechanical oscillators : Part i—dynamics of the underlying hamiltonian systems. *Journal of Applied Mechanics*, 68(1) :34–41, 2001.

[52] K. Genovese. Three-dimensional microscopic deformation measurements on cellular solids, 2016.

[53] K. Genovese, L. Casaletto, J. Rayas, V. Flores, and A. Martinez. Stereo-digital image correlation (dic) measurements with a single camera using a biprism. *Optics and Lasers in Engineering*, 51(3) :278 – 285, 2013.

[54] H. Ghazouani, M. Tagina, and R. Zapata. Robot Navigation Map Building Using Stereo Vision Based 3d Occupancy Grid. *JAITA*, 1(3) :63–72, 2011.

[55] A. Goeller, R. Le Breton, J.-L. Ddion, and T. Soriano. Real time modal identification of vibrating beam using Kalman filters and high speed cameras. In *SEM IMACS XXXVAt : Garden Grove*, California, USA, February 2017.

[56] W. Guo-Qing and D. M. Song. Implicit and explicit camera calibration : theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5) :469–480, May 1994.

[57] R. A. Haddad and A. N. Akansu. A class of fast gaussian binomial filters for speech and image processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 39 :723–727, 1991.

[58] H. Haddadi and S. Belhabib. Use of rigid-body motion for the investigation and estimation of the measurement errors related to digital image correlation technique. *Optics and Lasers in Engineering*, 46 :185–196, feb 2008.

[59] J. Hald. Time domain acoustical holography and its applications. *Sound and Vib.*, 35(2) :16–25, 2001.

[60] J. Hald. Basic theory and properties of statistically optimized near-field acoustical holography. *The Journal of the Acoustical Society of America*, 125(4) :2105–2120, 2009.

[61] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50. Citeseer, 1988.

[62] R. I. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6) :580–593, June 1997.

[63] R. I. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2) :146 – 157, 1997.

[64] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

[65] S. W. Hasinoff. *Photon, Poisson Noise*, pages 608–610. Springer US, Boston, MA, 2014.

[66] M. N. Helfrick, C. Niezrecki, P. Avitabile, and T. Schmidt. 3d digital image correlation methods for full-field vibration measurement. *MSSP*, 25(3) :917 – 927, 2011.

[67] F. Hild, A. Bouterf, P. Forquin, and S. Roux. *On the Use of Digital Image Correlation for the Analysis of the Dynamic Behavior of Materials*, pages 185–206. Springer International Publishing, Cham, 2018.

[68] H. Horaud and O. Monga. *Vision par ordinateur : outils fondamentaux*. Hermès, second edition, 1995.

[69] S. Hu, Y. Matsumoto, T. Takaki, and I. Ishii. Monocular stereo measurement using high-speed catadioptric tracking. *Sensors*, 17(8).

[70] R. Huñady and M. Hagara. A new procedure of modal parameter estimation for high-speed digital image correlation. *MSSP*, 93 :66 – 79, 2017.

[71] T. Huang, G. Yang, and G. Tang. A fast two-dimensional median filtering algorithm. *IEEE Transactions on Signal Processing*, 27(1) :13–18, 1 1979.

[72] P.-C. Hung and A. S. Voloshin. In-plane strain measurement by digital image correlation. *J. Braz. Soc. Mech. Sci. Eng.*, 25 :215–221, 2003.

[73] D. Hyun Lee, S.-O. Kweon, and R. Cipolla. A biprism-stereo camera system. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, page 87 Vol. 1, 02 1999.

[74] M. Inaba, T. Hara, and H. Inoue. A stereo viewer based on a single camera with view-control mechanisms. In *Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 1993, Tokyo, Japan, July 26 - 30, 1993*, pages 1857–1865, 1993.

[75] H. Inoue, T. Tachikawa, and M. Inaba. Robot Vision System with a Correlation Chip for Real-time Tracking, Optical Flow and Depth Map Generation. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, pages 1621–1626, 1992.

[76] J. Javh and M. Slaviqc, Boltežar. The subpixel resolution of optical-flow-based modal analysis. *MSSP*, 88 :89–99, 2017.

[77] D. Jean-Luc, A. Goeller, and M. Rosatello. Mesures de champs et assimilation de données en dynamique des structures. In *2ème Journée Mesures de Champs en Dynamique des structures (JMDC 2018)*, Saint-Ouen, France, Juin 2018.

[78] W. Jian, L. Yong, Z. Zhang, and P. Gao. Cost-effective and full-field method for measuring vibration of loudspeaker membrane using fringe projection. In *Proc. SPIE 9276, Optical Metrology and Inspection for Industrial Applications III*, volume 9276, pages 9276 – 9276 – 8, 2014.

[79] X. Jiang, D. M. McFarland, L. A. Bergman, and A. F. Vakakis. Steady state passive nonlinear energy pumping in coupled oscillators : theoretical and experimental results. *Nonlinear Dynamics*, 33(1) :87–102, 2003.

[80] C. V. Karsen, T. Bouman, and G. Gwaltney. Operating Deflection Shapes of a Violin String via High Speed/High Resolution Videography. In R. Allemang, J. D. Clerck, C. Niezrecki, and A. Wicks, editors, *Topics in Modal Analysis, Volume 7*, Conference Proceedings of the Society for Experimental Mechanics Series, pages 637–644. Springer New York, 2014.

[81] K. Khoshelham and S. O. Elberink. Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors*, 12(2) :1437–1454, 2012.

[82] M. Kleiner. *Electroacoustics*. Taylor & Francis Inc., Bosa Roca, United States, 02 2013.

[83] W. Klippel and J. Schlechter. Measurement and visualization of loudspeaker cone vibration. In *Audio Engineering Society Convention 121*, Oct 2006.

[84] K. Kraus and P. Waldhausl. *Manuel de Photogrammétrie, Principes et procédés fondamentaux*. 1998.

[85] M. Kyto, M. Nuutinen, and P. Oittinen. Method for measuring stereo camera depth accuracy based on stereoscopic vision. volume 7864, pages 7864 – 7864 – 9, 2011.

[86] J.-M. Lavest, M. Viala, and M. Dhome. Quelle précision pour une mire d'étalonnage ? *I-Revues, Traitement du Signal*, 16(3), 1999.

[87] J. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(2) :165–168, March 1980.

[88] J. J. Lee and M. Shinozuka. A vision-based system for remote sensing of bridge displacement. *NDT & E International*, 39(5) :425 – 431, 2006.

[89] G. D. Leo and A. Paolillo. Uncertainty evaluation of camera model parameters. In *2011 IEEE International Instrumentation and Measurement Technology Conference*, pages 1–6, May 2011.

[90] Y. Liping and P. Bing. Single-camera stereo-digital image correlation with a four-mirror adapter : optimized design and validation. *Optics and Lasers in Engineering*, 87 :120 – 128, 2016.

[91] S. Lobreau, E. Bavu, and M. Melon. Hemispherical double-layer time reversal imaging in reverberant and noisy environments at audible frequencies. *The J. of the Acoust. Soc. Am.*, 137-2 :785–796, Fev 2015.

[92] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60 :91–, 11 2004.

[93] H. Lu and P. D. Cary. Deformation measurement by digital image correlation : implementation of a second-order displacement gradient. *Exp. Mech.*, 40 :393–400, 2000.

[94] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of Imaging Understanding Workshop*, pages 121–130, 1981.

[95] T. Lundstrom, J. Baqersad, and C. Niezrecki. Monitoring the dynamics of a helicopter main rotor with high-speed stereophotogrammetry. *Exp. Tech.*, 40(3) :907–919, Jun 2016.

[96] M. Malesa and M. Kujawinska. Modified two-dimensional digital image correlation method with capability of merging of data distributed in time. *Appl. opt.*, 51 :8641–55, 12 2012.

[97] S. G. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7) :674–693, July 1989.

[98] R. Marik, J. Kittler, and M. Petrou. Error sensitivity assessment of vision algorithms based on direct error-propagation. In *In Proceedings of the ECCV Workshop on Performance Characteristics of Vision Algorithms*, 1996.

[99] H. Mathieu and F. Devernay. Système de miroirs pour la stéréoscopie. Technical report.

[100] J. D. Maynard, E. G. Williams, and Y. Lee. Nearfield acoustic holography : I. theory of generalized holography and the development of NAH. *J. Acoust. Soc. Am. 784*, pages 1395–1413, 1985.

[101] S. Mguil-Touchal, F. Morestin, and M. Brunet. Various experimental applications of digital image correlation method. *Rhodes. CMEM*, pages 45–58, 2018.

[102] B. V. Mieghem, J. Ivens, and A. V. Bael. Benchmarking of depth of field for large out-of-plane deformations with single camera digital image correlation. *Optics and Lasers in Engineering*, 91 :134–143, 2017.

[103] P. Monasse, J.-M. Morel, and Z. Tang. Epipolar rectification. 10 2018.

[104] S. Montresor. Filtrage d'images : Fondamentaux. `http://www.optique-ingenieur.org`, Mis en ligne septembre 2007.

[105] M. Moonen, P. Van Dooren, and J. Vandewalle. A singular value decomposition updating algorithm for subspace tracking. *SIAM Journal on Matrix Analysis and Applications*, 13(4) :1015–1038, 1992.

[106] V. Morozov. Pseudo-solutions. *USSR Computational Mathematics and Mathematical Physics*, 9(6) :196 – 203, 1969.

[107] F. K. Musgrave. *Methods for Realistic Landscape Imaging.* PhD thesis, Yale University, 1993.

[108] S. A. Nene and S. K. Nayar. Stereo with mirrors. In *Sixth International Conference on Computer Vision, 1998*, pages 1087–1094, 1998.

[109] H. Nguyen, Z. Wang, P. Jones, and B. Zhao. 3d shape, deformation, and vibration measurements using infrared kinect sensors and digital image correlation. *Appl. Opt.*, 56 :9030, 11 2017.

[110] S. Nijdam, J. S. Moerman, T. M. P. Briels, E. M. Veldhuizen, and U. Ebert. Stereo-photography of streamers in air. 92, 03 2008.

[111] P. Olaszek. Investigation of the dynamic characteristic of bridge structures using a computer vision method. *Measurement*, 25(3) :227 – 236, 1999.

[112] L. Oriat and E. Lantz. Subpixel detection of the centre of an object using a spectral phase algorithm on the image. *Pattern Recognit.*, 31 :761–71, 1998.

[113] J.-J. Orteu. *Mesure de forme et de déformations par stéréo-corrélation d'images : applications et mécanique expérimentale des solides*, pages 186–218. 11 2009.

[114] J.-N. Ouellet and P. Hébert. A simple operator for very precise estimation of ellipses. In *CRV '07 : Proceedings of the Fourth Canadian Conference on Computer and Robot Vision*, pages 21–28, Montréal, Qc, Canada, 28 - 30 May 2007. IEEE Computer Society.

[115] M. Ozbek, D. J. Rixen, O. Erne, and G. Sanow. Feasibility of monitoring large wind turbines using photogrammetry. *Energy*, 35(12) :4802 – 4811, 2010. The 3rd International Conference on Sustainable Energy and Environmental Protection, SEEP 2009.

[116] B. Pan, L. Yu, and Q. Zhang. Review of single-camera stereo-digital image correlation techniques for full-field 3D shape and deformation measurement. *SCIENCE CHINA Technological Sciences*, 61(1) :2–20, 2018.

[117] M. Pankow, B. Justusson, and A. Waas. Three-dimensional digital image correlation technique using single high-speed camera for measuring large out-of-plane displacements at high framing rates. 49 :3418–27, 06 2010.

[118] A. Pelat, J. Cuenca, F. Gautier, and L. Barguet. Potentialities of the acoustic black hole effect for damping plate vibrations. In *Acoustics*, Nantes, 04 2012.

[119] C. Pezerat and J.-L. Guyader. Identification of vibration sources. *Appl. Acoust.*, 61 :309–324, 11 2000.

[120] A. G. Piersol and T. L. Paes. *Harris' Shock and Vibration Handbook, Sixth edition.* Oct. 2009.

[121] A. Pilch and T. Chu. Measurement of whole-field surface displacements and strain using a genetic algorithm based intelligent image correlation method. *Trans. ASME, J. Dyn. Syst. Meas. Control*, 126 :479–488, 2004.

[122] M. C. Pitter, C. W. See, and M. G. Somekh. Fast subpixel digital image correlation using artificial neural networks. In *Int. Conf. Image Proc.*, volume 2, pages 901–904, 2001.

[123] M. C. Pitter, C. W. See, and M. G. Somekh. Subpixel microscopic deformation analysis using correlation and artificial neural networks. *Opt. Exp.*, 8 :322–327, 2001.

[124] J. Poittevin, P. Picart, C. Faure, F. Gautier, and C. Pézerat. Multi-point vibrometer based on high-speed digital in-line holography. *Appl. Opt.*, 54(11) :3185–3196, Apr 2015.

[125] W. H. Press, S. A. Teukolsky, and W. T. Vetterling. *Numérical Recipies in C, The art of Scientific Computing.* Cambridge University Press, second edition, 1992.

[126] T. Putze, K. Raguse, and H.-G. Maas. Configuration of multi mirror systems for single high-speed camera based 3D motion analysis. In *Proc. SPIE, 6491, Videometrics IX, 64910L*, 2007.

[127] J. Réthoré, A. Gravouil, F. Morestin, and A. Combescure. Estimation of mixed-mode stress intensity factors using digital image correlation and an interaction integral. *International Journal of Fracture*, 132 :65–79, 2005.

[128] J. Réthoré, F. Morestin, L. Lafarge, and P. Valverde. 3d displacement measurements using a single camera. *Optics and Lasers in Engineering*, 57 :20–27, 2014.

[129] J. Sabater. Optique instrumentale, microscopie. `http://www.optique-ingenieur.org`, Mis en ligne juillet 2007.

[130] F. Sakai and S. Takaeda. Tuned liquid column damper — new type device for suppression of building vibrations. In *Proceedings of International Conference on High Rise Buildings*, pages 25–27, Nanjing, China, 1989.

[131] W. Sankowski, M. Włodarczyk, D. Kacperski, and K. Grabowski. Estimation of measurement uncertainty in stereo vision system. *Im. Vis. Comp.*, 61 :70 – 81, 2017.

[132] F. M. Santoyo, M. C. Shellabear, and J. R. Tyrer. Whole field in-plane vibration analysis using pulsed phase-stepped espi. *Appl. Opt.*, 30(7) :717–721, Mar 1991.

[133] A. Sarkissian. Extension of measurement surface in near-field acoustic holography. *The Journal of the Acoustical Society of America*, 115(4) :1593–1596, 2004.

[134] H. W. Schreier, J. R. Braasch, and M. A. Sutton. Systematic errors in digital image correlation caused by intensity interpolation. *Opt. Eng*, 39 :2915–2921, 2000.

[135] K. Schreve. How accurate can a stereovision measurement be ? In *15th International Workshop on Research and Education in Mechatronics (REM)*, pages 1–7, Sept 2014.

[136] J. R. Seal, D. G. Bailey, and G. S. Gupta. Depth perception with a single camera. In *Int. Conf. Sens. Tech.*, pages 96–101, 2005.

[137] U. Seidel and W. Klippel. Fast and accurate measurement of the linear transducer parameters. In *Audio Engineering Society Convention 110*, May 2001.

[138] T. Semenova and S. F. Wu. On the choice of expansion functions in the helmholtz equation least-squares method. *The Journal of the Acoustical Society of America*, 117(2) :701–710, 2005.

[139] J. Shi and C. Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pages 593–600, Seattle, Washington, 1994.

[140] J. Sicard and J. Sirohi. Measurement of the deformation of an extremely flexible rotor blade using digital image correlation. *Meas. Sc. Tech.*, 24 :065203, 05 2013.

[141] T. Siebert, T. Becker, K. Splitthof, I. Neumann, and R. Krupka. High-speed digital image correlation : error estimations and applications. *Optical Engineering*, 46 :46 – 46 – 7, 2007.

[142] E. Simonetto and C. Pezerat. Feasability of the deformation analysis of plate subjected to vibrations using recent computer vision methods. volume 38, Saint Mandé, Septembre 2010.

[143] C. C. Slama. *Manual of Photogrammetry (4th Ed.)*. Am. Soc. Photogramm., Falls Church, VA, 1980.

[144] J. Smisek, M. Jancosek, and T. Pajdla. *3D with Kinect*, pages 3–25. Springer London, London, 2013.

[145] S. Starosielec and D. Hägele. Discrete-time windows with minimal RMS bandwidth for given RMS temporal width. *Signal Processing*, 102 :240 – 246, 2014.

[146] B. Stasicki and F. Boden. In-flight measurements of aircraft propeller deformation by means of an autarkic fast rotating imaging system. In *Proc. SPIE*, volume 9302, pages 9302 – 6, 2015.

[147] C. Steger. Subpixel-precise extraction of lines and edges. *ISPRS*, XXXIII, 01 2000.

[148] W. Steinchen, L. Yang, G. Kupfer, P. Mäckel, and F. Vössing. Vibration analysis by means of digital speckle pattern shearing interferometry (digital shearography). In A. Lagarde, editor, *IUTAM Symposium on Advanced Optical Methods and Applications in Solid Mechanics*, pages 263–274, Dordrecht, 2002. Springer Netherlands.

[149] H. Stewénius, C. Engels, and D. Nistér. Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(4) :284–294, 2006.

[150] B. Stoffregen and A. Felske. Scanning laser doppler vibration analysis system. In *SAE Technical Paper*. SAE International, 02 1985.

[151] M. Sutton, J.-J. Orteu, and H. W. Schreier. *Image Correlation for Shape, Motion and Deformation Measurements. Basic Concepts, Theory and Applications*. Springer US, 01 2009.

[152] M. Sutton, W. Wolters, W. Peters, W. Ranson, and S. McNeill. Determination of displacements using an improved digital correlation method. *Im. Vi. Comp.*, 1(3) :133 – 139, 1983.

[153] M. A. Sutton, S. R. McNeill, J. Jang, and M. Babai. Effects of subpixel image restoration on digital correlation error estimates. *Optical Engineering*, 27 :870–877, 1988.

[154] S. Suzuki and K. Abe. Topological structural analysis of digitized binary images by border following. *j-CVGIP*, 30(1) :32–46, Apr 1985.

[155] U. P. Svensson. Modelling Acoustic Spaces for Audio Virtual Reality. In *Proceedings of the 1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA)*, 11 2002.

[156] R. Szeliski. *Computer Vision : Algorithms and Applications*, volume 5. 01 2011.

[157] N. Thiele. Loudspeakers in vented boxes : Part 2. *J. Audio Eng. Soc*, 19(6) :471–483, 1971.

[158] J.-H. Thomas, V. Grulier, P. Paillasseur, J.-C. Pascal, and J.-C. Le Roux. Real-time near-field acoustic holography for continuously visualizing nonstationary acoustic fields. *J. Acoust. Soc. Am.*, 3554-67, Dec 2010.

[159] J.-H. Thomas, V. Grulier, S. Paillasseur, and J.-C. Pascal. Real-Time Nearfield Acoustic Holography : Implementation of the Direct and Inverse Impulse Responses in the Time-Wavenumber Domain. *Proceeding of ICSV15*, 2008.

[160] V. Tiwari, M. A. Sutton, and S. R. McNeill. Assessment of high speed imaging systems for 2d and 3d deformation measurements : Methodology development and validation. *Experimental Mechanics*, 47(4) :561–579, Aug 2007.

[161] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal on Robotics and Automation*, 3(4) :323–344, August 1987.

[162] A. F. Vakakis and O. Gendelman. Energy pumping in nonlinear mechanical oscillators : part ii—resonance capture. *Journal of Applied Mechanics*, 68(1) :42–48, 2001.

[163] P. L. Walter. The history of the accelerometer. *Sound & vibration*, 41(1) :84–92, January.

[164] J. M. Wanek and C. H. Wu. Automated trinocular stereo imaging system for three-dimensional surface wave measurements. *Ocean Engineering*, 33(5) :723 – 747, 2006.

[165] R. Wang, X. Li, and Y. Zhang. Analysis and optimization of the stereo-system with a four-mirror adapter. *Journal of the European Optical Society - Rapid publications*, 3(0), 2008.

[166] W. Wang, J. E. Mottershead, A. Ihle, T. Siebert, and H. Reinhard Schubach. Finite element model updating from full-field vibration measurement using digital image correlation. *Journal of Sound and Vibration*, 330(8) :1599 – 1620, 2011.

[167] C. Warren, C. Niezrecki, and P. Avitabile. Optical non-contacting vibration measurement of rotating turbine blades ii. In T. Proulx, editor, *Structural Dynamics and Renewable Energy, Volume 1*, pages 39–44, New York, NY, 2011. Springer New York.

[168] B. Wattrisse, A. Chrysochoos, J.-M. Muracciole, and M. Némoz-Gaillard. Analysis of strain localization during tensile tests by digital image correlation. *Experimental Mechanics*, 41(1) :29–39, Mar 2001.

[169] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10) :965–980, Oct 1992.

[170] A. D. White and J. D. Rigden. Continuous gas maser operation in the visible. *Proc IRE*, 50 :1967, July 1962.

[171] T. G. White, J. R. W. Patten, K.-H. Wan, A. D. Pullen, D. J. Chapman, and D. E. Eakins. A single camera three-dimensional digital image correlation system for the study of adiabatic shear bands. *Strain*, 53(3), 2017.

[172] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series :With Engineering Applications*. MIT press, 1949.

[173] J. Wiesel. Digital image processing for orthophoto generation. *Photogrammetria*, 40(2) :69 – 76, 1985. Pattern Recognition in Photogrammetry, 3.

[174] E. G. Williams. *Fourier Acoustics Sound radiation and Near field Acoustical Holography*. San Diego, 1999.

[175] C. Woon and L. Mitchell. Variations in structural dynamic characteristics caused by changes in ambient temperature : I. Experimental. In *Proceedings-spie the international society for optical engineering*, pages 963–671. SPIE International society for optical, 1996.

[176] Y. Xu, Y. Zhao, F. Wu, and K. Yang. Error analysis of calibration parameters estimation for binocular stereo vision system. In *2013 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 317–320, Oct 2013.

[177] T. Xue, L. Qu, Z. Cao, and T. Zhang. Three-dimensional feature parameters measurement of bubbles in gas–liquid two-phase flow based on virtual stereo vision. *Flow Meas. Instrum.*, 27 :29–36, 10 2012.

[178] Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar, and D. Mascarenas. Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. *MSSP*, 85 :567 – 590, 2017.

[179] Y. Yeh and H. Z. Cummins. Localized fluid flow measurement with an he–ne laser spectrometer. *Applied Physics Letters*, 4(10) :176–178, 1964.

[180] L. Yu and B. Pan. Structure parameter analysis and uncertainty evaluation for single-camera stereo-digital image correlation with a four-mirror adapter. *Appl. Opt.*, 55(25) :6936–6946, Sep 2016.

[181] L. Yu and B. Pan. Full-frame, high-speed 3d shape and deformation measurements using stereo-digital image correlation and a single color high-speed camera. *Optics and Lasers in Engineering*, 95 :17 – 25, 2017.

[182] L. Yu and B. Pan. Single-camera high-speed stereo-digital image correlation for full-field vibration measurement. *MSSP*, 94 :374–383, 2017.

[183] D. Zhang, X. Zhang, and G. Cheng. Compression strain measurement by digital speckle correlation. *Experimental Mechanics*, 39(1) :62–65, Mar 1999.

[184] J. Zhang, G. Jin, S. Ma, and L. Meng. Application of an improved subpixel registration algorithm on digital speckle correlation measurement. *Optics & Laser Technology*, 35(7) :533 – 542, 2003.

[185] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11) :1330–1334, Nov 2000.

[186] F. Zhou, Y. Wang, B. Peng, and Y. Cui. A novel way of understanding for calibrating stereo vision sensor constructed by a single camera and mirrors. *Measurement*, 46(3) :1147 – 1160, 2013.

[187] P. Zhou and K. E. Goodson. Subpixel displacement and deformation gradient measurement using digital image/speckle correlation. *Optical Engineering*, 40 :1613 – 1620, 2001.

[188] J. G. Zhu, Y. J. Li, and S. H. Yz. Design and calibration of a single-camera-based stereo vision sensor. *Opt. Eng.*, 08 2006.

UNIVERSITE BRETAGNE LOIRE / SCIENCES POUR L'INGENIEUR

Le Mans Université

Titre : Mesure de vibrations par vision 3D

Mots clés : Vision 3D, Caméra ultra-rapide, Vibro-acoustique, Stéréo corrélation d'images (SDIC).

Résumé : La finalité de cette thèse est d'étudier la pertinence et les limites des méthodes de vision 3D couplées à une caméra ultra-rapide pour la mesure de champs vibratoires, sans contact et de manière synchrone, dans le domaine des fréquences associées à la vibro-acoustique. Un premier montage pseudo stéréoscopique, issu de la robotique, mobilisant un jeu de quatre miroirs afin de générer deux vues virtuelles à partir d'une seule caméra réelle, a été testé sur une plaque et un haut-parleur. Les résultats, validés par comparaison avec ceux obtenus avec un vibromètre laser, attestent de la pertinence de l'approche en dépit des contraintes liées aux éléments optiques. Dans une logique de simplification, trois autres montages ont alors été proposés et testés, permettant de concevoir deux techniques de mesure de vibration plein champ et une méthode itérative de rectification d'images (IRIs), adaptées au contexte.

La méthode sans miroir utilise une ligne mathématique pour la triangulation et est fondamentalement adaptée à la mesure de vibrations mono-axiales d'objets globalement plans, affichant des déplacements non-répétables selon la normale de la surface ou selon un axe connu. La méthode à caméras asynchrones, quant à elle, utilise une caméra ultra-rapide et une caméra rapide, et permet la mesure de vibrations multi-axiales de phénomènes vibratoires 3D. Les résultats obtenus sur le capot d'une voiture et sur un haut-parleur attestent de son potentiel pour la caractérisation de panneaux ou le test qualité de fin de chaine d'assemblage de haut-parleurs par exemple. En conclusion, les trois protocoles de mesure et les résultats associés sont comparés afin de cibler leurs potentialités et limites respectives dans le contexte de la mesure de vibrations.

Title : 3D vision vibration measurement

Keywords : 3D vision, High-speed camera, Vibro-acoustics, Stereo Digital Image Correlation (SDIC)

Abstract: The objective of this Ph.D. is to study the relevance and limits of 3D vision methods coupled to high-speed cameras and applied to non-contact synchronous vibration measurement, in the vibro-acoustic range of frequencies. A first pseudo stereoscopic set-up, taken from robotics, using a four-mirror adapter in order to generate two virtual viewpoints from a single real camera, has been tested on a plate and a loudspeaker. The results, validated by comparison with those obtained with a laser vibrometer, prove the relevance of the approach, despite some constraints related to the optical elements. In a logic of simplification, three other set-ups have then been proposed and tested, allowing designing two full-field vibration measurement techniques and a method for the Iterative Rectification of Images (IRIs), adapted to the context.

The no-mirror method uses a mathematical line to triangulate positions and is basically suited to measure the single-axis vibrations of globally plane objects, displaying non-repeatable displacements along the normal of the surface or along a known axis. The asynchronous cameras technique requires a high-speed and an industrial camera used simultaneously to measure the multi-axis displacements of 3D vibratory phenomena. The results obtained on a car bonnet and a loudspeaker prove its potential to characterise large panels or to carry out end-of-line testing of loud-speakers for example. In conclusion, the three measurement protocols and the associated results are compared in order to assess their respective potentialities and limits in the context of vibration measurement.