

N° d'ordre :

École Doctorale Mathématiques, Sciences de l'Information
et de l'Ingénieur

ULP – INSA – ENGEES

THÈSE

présentée pour obtenir le grade de

Docteur de l'Université Louis Pasteur – Strasbourg I
Domaine : Sciences pour l'Ingénieur - EEA
Spécialité : Traitement d'Images et Vision par Ordinateur

par

Aicha Beya FAR

Analyse multi-images.
Application à l'extraction contrôlée d'indices images
et à la détermination de descriptions scéniques.

Soutenue publiquement le 12 décembre 2005

Membres du jury

M. Jack-Gérard Postaire, Professeur, USTL, Lille
M. Alain Trémeau, Professeur, UJM, Saint-Etienne
M. Jean-Michel Dischler, Professeur, ULP
M. Ernest Hirsch, Professeur, ULP
Mme Sophie Kohler, MCF, ULP

Rapporteur externe
Rapporteur externe
Rapporteur interne
Directeur de thèse
Examineur

Table des matières

Introduction générale	1
1 Vision stéréoscopique	9
1.1 Introduction	11
1.2 Modélisation géométrique d'une caméra	13
1.3 Étalonnage du capteur	14
1.3.1 Méthodes d'étalonnage	15
1.3.1.1 Étalonnage selon Tsai	15
1.3.1.2 Étalonnage selon Faugeras et Toscani	16
1.3.1.3 Étalonnage à base d'ellipses	16
1.3.1.4 Étalonnage selon Zhang	16
1.3.2 Auto-étalonnage	17
1.3.3 Étalonnage faible	17
1.4 Géométrie épipolaire et matrice fondamentale	18
1.4.1 La géométrie épipolaire	19
1.4.2 Expression matricielle de la matrice fondamentale	19
1.4.3 Méthodes d'estimation de la matrice fondamentale	20
1.4.3.1 Méthodes linéaires	21
1.4.3.2 Méthodes itératives	22
1.4.3.3 Méthodes robustes	24
1.5 Extraction des indices visuels	24
1.5.1 Points d'intérêts	25
1.5.2 Régions	25
1.5.3 Contours	26
1.5.3.1 Types de contours	27
1.5.3.2 Caractéristiques des contours	27
1.5.4 Segments	28
1.6 Mise en correspondance	28
1.6.1 Contraintes de mise en correspondance	29
1.6.2 Méthodes de mise en correspondance	32
1.6.2.1 Mise en correspondance par corrélation	33
1.6.2.2 Mise en correspondance par programmation dynamique	35
1.6.2.3 Mise en correspondance par relaxation	35
1.6.2.4 Mise en correspondance par invariants et triangulation de Delaunay	36
1.6.2.5 Mise en correspondance hiérarchique	36

1.6.2.6	Approches basées sur des primitives	37
1.7	Reconstruction 3D	38
1.7.1	Le problème de la triangulation	39
1.7.2	Reconstruction euclidienne	41
1.8	Conclusion	42
2	Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques	45
2.1	Introduction	47
2.2	Représentation des primitives contours utilisées	47
2.3	Détection, construction et classification des chaînes de points de contour	49
2.4	Procédure d'appariement basée sur le calcul bidirectionnel de la contrainte épipolaire	57
2.4.1	Principe d'appariement	57
2.4.2	Élimination des faux appariements	58
2.4.3	Appariement point à point	61
2.5	Comparaison de données images réelles et conceptuelles	64
2.6	Conclusion	71
3	Modélisation CAO et modélisation de l'illumination	73
3.1	Introduction	75
3.2	Modélisation CAO	75
3.2.1	Méthodes de modélisation d'un objet	76
3.2.1.1	Modélisation filaire	76
3.2.1.2	Modélisation surfacique	77
3.2.1.3	Modélisation volumique	77
3.2.2	Modélisation de pièces industrielles et système CATIA V5	79
3.2.2.1	Modélisation de la scène (conditions de prise de vue)	80
3.2.2.2	Génération des images conceptuelles et images filaires	82
3.2.3	Modélisation du capteur et changement de repère	84
3.2.4	Exploitation de l'information <i>a priori</i>	89
3.2.4.1	Projection des primitives obtenues par CATIA V5	90
3.2.4.2	Superposition et comparaison	92
3.2.5	Conclusion	93
3.3	Modélisation de l'illumination	93
3.3.1	Importance de l'éclairage	94
3.3.2	Différents types d'éclairages	94
3.3.3	Modélisation de la réflexion de la lumière	95
3.3.3.1	Modèle de Beckmann et Spizzichino	97
3.3.3.2	Modèle de Torrance et Sparrow	97
3.3.4	Premiers essais de modélisation de l'illumination	98
3.3.5	Conclusion de la modélisation de l'illumination	105
3.4	Conclusion	105

TABLE DES MATIÈRES

4	Automatisation et mise en place d'un système d'inspection	107
4.1	Introduction	109
4.2	Inspection automatique	110
4.3	Graphes de situations	111
4.3.1	Éléments de base constituant un SGT	112
4.3.2	Contexte du développement et d'utilisation classique des SGT	112
4.3.3	SGT adapté à notre application : reconstruction tridimensionnelle	113
4.4	Modélisation du comportement du capteur par les SGT	114
4.5	Reconstruction 3D guidée par les SGT : tête de mesure stéréoscopique	115
4.5.1	Traitements effectués hors ligne	116
4.5.2	Traitements effectués en ligne	117
4.6	Automatisation d'une reconstruction 3D par stéréovision et replanification dynamique utilisant les SGT	118
4.7	Description du dispositif expérimental	123
4.8	Correction de l'éclairage	123
4.9	Implémentation	125
4.9.1	Génération du SGT dédié à la reconstruction 3D partielle par stéréovision	127
4.9.2	Système d'acquisition et d'évaluation	127
4.10	Validation de la replanification par SGT	129
4.11	Conclusion	133
	Conclusion générale et perspectives	135
	Bibliographie	139
	Liste de publications	149
	Annexes	151
	A. Étalonnage d'une caméra	154
	B. Présentation CATIA V5	159

Introduction générale

La vision est l'un de nos sens les plus élaborés. L'œil humain nous donne l'opportunité de percevoir et d'interagir avec le monde qui nous entoure, pour reconnaître des personnes ou des objets par exemple. Depuis sa naissance dans les années 60, la vision artificielle a pour vocation de permettre à des systèmes automatiques ou robotiques d'acquérir et de traiter des informations visuelles en reproduisant le fonctionnement complexe de la vision humaine.

Depuis plusieurs années, le Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT) de l'Université Louis Pasteur (ULP) est un laboratoire de recherche interdisciplinaire fédéré par l'imagerie. Les grandes disciplines qui y sont représentées sont l'informatique, le traitement du signal, l'automatique, la télédétection. Parmi les disciplines, certaines sont plus particulièrement centrées sur l'imagerie, le traitement et l'analyse des images, mais également sur les interactions de celles-ci avec la synthèse. Depuis l'installation à l'ULP, en février 2002, d'une station de réalité virtuelle à base de projections sur grands écrans (workbench), le programme Réalité Virtuelle et Augmentée, simulation et calcul intensif (RéVA) mobilise trois équipes du LSIIT : l'équipe Informatique Géométrique et Graphique (IGG), l'équipe Image et Calcul Parallèle et Scientifique (ICPS) et l'équipe Modèles Images et Vision (MIV). Ces trois équipes s'impliquent dans ce projet fédérateur en développant les thèmes suivants : Réalité virtuelle et simulation (équipe IGG), Calcul parallèle et intensif (équipe ICPS), Réalité augmentée (équipe MIV).

L'équipe MIV vise notamment à développer des approches algorithmiques pour la comparaison d'images réelles et conceptuelles (CIRC), en cherchant à modéliser la scène à analyser, son contenu et les conditions d'acquisition, soit *a priori* ou en ligne pendant le déroulement des algorithmes. L'accent est mis sur la capacité d'apprentissage des algorithmes, conduisant à une évolution dynamique de la base de connaissance mise en œuvre. Les objectifs de l'équipe MIV sont donc, à partir de séquences spatiales, de construire des descriptions scéniques (modélisation et reconstruction 3D) et d'évaluer la qualité d'un objet (contrôle qualité, applications, etc.). Pour atteindre ces objectifs, il est nécessaire de faire appel, d'une part, à des connaissances *a priori* telles qu'une connaissance structurelle (capteurs, illumination, modèles, etc.) ou une connaissance procédurale (comportement des « opérateurs », stratégie, plans d'action, etc.) et, d'autre part, à des techniques d'acquisition variées (lumière structurée, stéréovision, etc.). L'objectif est de reconstruire dans sa totalité, à partir de séquences d'images CCD acquises au moyen d'un scanner (robot de mesure), un objet quasi-polyédrique incluant des surfaces gauches en exploitant de manière optimale ces données 3D. Le travail présenté dans ce mémoire s'insère dans le thème fédérateur RéVA, a été développé au sein de l'équipe MIV et porte sur la reconstruction en ligne de pièces manufacturées quasi-polyédriques.

L'approche que nous avons adoptée se base sur la comparaison d'images réelles de la pièce à évaluer avec des images conceptuelles de celles-ci, issues par exemple d'un outil CAO. Pour notre application, un modèle CAO unique est conçu. L'approche proposée fait essentiellement appel à une représentation tridimensionnelle des contours de l'objet à évaluer. Par ailleurs, la stéréovision est une technique puissante et très adaptée à l'obtention d'une information 3D à partir de deux images 2D. Une étape préalable cruciale est

l'appariement de caractéristiques extraites du contenu des images. L'appariement ou mise en correspondance est une tâche difficile et délicate faisant encore l'objet de nombreuses études.

Cependant, la plupart des techniques d'appariement existantes ne sont souvent pas adaptées aux images de pièces industrielles. Dans une première contribution à cette thèse, nous présentons une approche d'appariement adaptée aux caractéristiques extraites de ce type d'images. L'approche proposée est adaptée à l'appariement de contours et se fonde sur l'estimation de la géométrie épipolaire dans la paire d'images stéréoscopiques, ainsi que sur la comparaison de données réelles et synthétiques dans le but de sélectionner l'ensemble de contours à appairer. Cette comparaison permet d'exploiter une connaissance *a priori* (p. ex. le modèle CAO de l'objet à inspecter) comme contrainte pour les traitements ultérieurs, en ne conservant que les contours de l'objet qui sont visibles dans les deux images.

L'appariement se fait d'une façon dite bidirectionnelle en calculant la contrainte épipolaire d'abord de l'image gauche vers l'image droite, puis de l'image droite vers l'image gauche pour éliminer les faux appariements. La technique proposée repose sur une étape primordiale de segmentation et de classification des contours en segments de droites ou segments curvilignes selon un critère géométrique. Cette classification permet de restreindre la recherche du correspondant d'un segment de type donné aux segments du même groupe géométrique. Au lieu d'utiliser des points singuliers, notre méthode utilise des segments comme primitives, en prenant en compte le nombre de points appartenant à chaque primitive, en faisant usage de ce que nous avons appelé « matrices d'appariement ». Cela revient à introduire une contrainte géométrique dans notre méthode d'appariement. L'approche a été testée sur des images réelles et les résultats expérimentaux montrent que notre méthode aboutit à une mise en correspondance de contours précise et à une reconstruction 3D efficace.

En se basant sur la comparaison d'images réelles et d'images conceptuelles de la pièce à inspecter, une étude d'implémentation de techniques organisant la coopération entre images acquises et représentations conceptuelles (modèle CAO) est proposée, en vue d'une meilleure description du contenu des images.

L'élimination des effets indésirables (effet d'éclairage, bruit, etc.) est un préalable à l'exploitation des images. La méthode mise au point pour réaliser l'extraction contrôlée des indices intéressants consiste à comparer les images réelles incluant les nuisances avec leur modélisation issue d'outil CAO et représentée par des images de synthèse ne reproduisant pas ces défauts. Pour ce faire, un modèle complet 3D est établi en utilisant l'outil CAO CATIA (Dassault Systems). L'approche proposée pour sélectionner les primitives qui doivent d'abord être appariées, puis reconstruites, consiste à comparer les images réelles aux images conceptuelles. Ces dernières sont obtenues en tenant compte des points de vue des deux caméras par l'intermédiaire des paramètres d'étalonnage. Les images conceptuelles gauche et droite résultantes sont alors utilisées comme référence. En tenant compte du point de vue de la caméra, ces représentations accentuent les segments visibles dans les deux images. Par conséquent, seuls ces contours vont subir la procédure d'appariement décrite ci-dessus et les caractéristiques non visibles des points de vue courants des caméras déterminés par étalonnage sont éliminées. Ceci conduit à l'obtention d'un

modèle réduit de la pièce qui ne contient que les caractéristiques visibles dans les images réelles. Les deux images synthétiques correspondantes peuvent alors être superposées aux images réelles pour les deux points de vue des caméras composant la tête stéréoscopique. De cette façon, l'utilisation des modèles réduits CAO permet de guider la mise en correspondance correcte des primitives des deux images utiles à la reconstruction de l'objet, même si les images ne contiennent pas le même nombre de contours. Enfin, ce modèle CAO joue un rôle très important dans notre application, étant donné que la comparaison d'images réelles et conceptuelles autorise également une métrologie de grande précision, par exemple en vue d'un contrôle qualité.

Une deuxième contribution à cette thèse consiste à améliorer la maîtrise des paramètres d'éclairage par une modélisation adéquate des phénomènes d'illumination. Cela permet d'ajuster le traitement d'images aux conditions d'illumination observées localement dans l'image et d'anticiper, à partir des images conceptuelles, la pertinence des informations contenues dans les images réelles.

Les applications de la vision par ordinateur nécessitent de plus en plus souvent une reconstruction 3D précise des objets à manipuler, à évaluer, etc. Plus précisément, à partir d'une analyse multi-images, on détermine des descriptions scéniques en faisant appel à une connaissance *a priori* (modèles, capteurs, illumination, comportement des opérateurs, stratégie, etc.) et à des techniques d'acquisition variées (p. ex. stéréovision, lumière structurée, etc.). Or, la construction de ces descriptions est grandement simplifiée si l'on fait appel à un outil de planification utilisant, dans notre étude, des graphes de situations. L'utilisation de ces graphes autorise une automatisation des différentes procédures de la reconstruction 3D et permet aussi une replanification des traitements en cours d'exécution en fonction des résultats obtenus (auto-apprentissage).

Le système de planification développé dans notre laboratoire est constitué de deux modules. Le premier est réservé à l'acquisition des données et à la reconstruction 3D partielle de l'objet et sa gestion se fait par graphes de situations (SGT ou Situation Graph Trees). Le second module prend en charge le contrôle de la séquence d'acquisition ainsi que la reconstruction 3D complète et l'évaluation dimensionnelle de l'objet. Sa gestion se fait à l'aide d'un plan de tâches hiérarchiques (HTP ou Hierarchical Task Plan). L'automatisation de la reconstruction 3D par vision stéréoscopique a conduit au développement d'un SGT spécifique. Ce système se charge de la gestion des traitements permettant l'analyse dimensionnelle complète d'une pièce manufacturée. De ce fait, notre troisième contribution, dans le cadre de cette thèse, consiste en l'utilisation des graphes de situations pour réaliser des opérations d'acquisition et de reconstruction partielle planifiées. Une caractéristique intéressante à exploiter de ces graphes, est qu'ils sont également capables de s'adapter dynamiquement aux conditions réelles de prise de vue (effets d'illumination, position de la tête stéréoscopique) et donc de modifier automatiquement les traitements des données et éventuellement les conditions effectives d'acquisition.

Nous nous sommes ainsi focalisés sur ces aspects de la planification et de l'automatisation de tâches dédiées à la reconstruction de l'objet. Toute la chaîne de traitement est planifiée hors ligne par un opérateur expérimenté; cette planification prend en compte des informations *a priori* liées au modèle de l'objet ainsi qu'au système d'acquisition (en

particulier le capteur stéréoscopique et le système d'illumination).

Nous avons ainsi développé une approche permettant une replanification dynamique adaptée au type de problèmes rencontrés en utilisant le système. L'intérêt réside dans la possibilité de réagir à des actions qui ne sont pas exécutées comme prévu par le planificateur (opérateur). Notre graphe de situations doit alors replanifier certaines actions, qui vont prendre en compte des valeurs modifiées des paramètres. Sur cette base, nous avons développé des règles de décision permettant au système de planification d'adapter les traitements en cours d'exécution. Ces règles se basent en premier lieu sur l'ajustement de l'éclairage de la scène et, dans un second temps, sur le déplacement de la tête de mesure.

Intégrés dans la chaîne de traitement, les modules développés sont capables de fournir une reconstruction partielle élaborée des objets à évaluer. Cette reconstruction doit ensuite être comparée avec le modèle CAO correspondant, pour décider de la qualité de l'objet. En fonction du résultat, le système permet d'estimer quantitativement les caractéristiques dimensionnelles des pièces ainsi que les erreurs d'usinage. De ce fait, nous mettons à disposition une information bien plus significative sur la qualité de la production.

Notre mémoire de thèse est constitué de quatre chapitres. Le premier fait l'objet d'un état de l'art de la vision stéréoscopique. Le problème de la mise en correspondance est largement étudié dans la littérature et est présenté ici de manière à aborder les différentes méthodes d'appariement existantes. Nous présentons également les principes de base de la stéréovision, à savoir la modélisation et l'étalonnage du capteur, la géométrie épipolaire, l'extraction des indices visuels ainsi que les méthodes de reconstruction 3D.

Le deuxième chapitre présente l'approche que nous proposons pour apporter une solution aux problèmes liés à l'appariement de contours dans des images stéréoscopiques d'un objet manufacturé. Dans ce chapitre, nous décrivons les différentes étapes de notre algorithme de mise en correspondance, de l'extraction des indices visuels jusqu'à l'obtention des points 3D. La modélisation CAO et la modélisation de l'illumination font l'objet du chapitre 3. Nous abordons ainsi, dans un premier temps, la définition et le développement d'une procédure de modélisation CAO adaptée au type de pièces à évaluer, ainsi que la génération des images conceptuelles. Dans un second temps, nous focalisons l'attention sur les phénomènes d'illumination dans un système de vision, en précisant comment les gérer convenablement. Ainsi, quelques premiers essais de modélisation de l'illumination sont présentés. Ces résultats montrent qu'il est possible de maîtriser des paramètres d'éclairage afin de les adapter pour optimiser un traitement donné. Ceci peut être exploité en cours d'exécution de l'application, en ligne, en repositionnant par exemple la source lumineuse lorsque les indices visuels de l'objet à inspecter ne sont pas entièrement ou correctement perçus.

Enfin, dans le chapitre 4, nous abordons la mise en place et l'automatisation de notre système d'inspection de pièces manufacturées. Nous commençons tout d'abord par introduire quelques généralités sur l'inspection automatique. Puis nous étudions les graphes de situations, tout en présentant leur contexte habituel d'utilisation, ainsi que leur adaptation à notre application de reconstruction 3D en vue, par exemple, d'un contrôle qualité. Nous mettons ensuite l'accent sur l'analyse d'une reconstruction partielle par stéréovision et sur sa replanification dynamique en utilisant les graphes de situations.

Finalemant, une conclusion générale dresse un bilan des apports de notre approche et fait ressortir les perspectives liées à ce travail.

Chapitre 1

Vision stéréoscopique

1.1 Introduction

Depuis que David Marr [Mar82] a proposé vers la fin des années 70 une méthodologie pour simuler la vision humaine, les systèmes binoculaires ont présenté un intérêt croissant, et sont devenus très répandus. Marr a défini la vision comme étant un processus qui crée, à partir d'un ensemble d'images, une représentation complète et précise de la scène observée. Ce processus théorique est traduit en trois étapes de traitements algorithmiques :

- la segmentation. L'extraction de primitives géométriques de bas niveau (segments, contours, etc.), dans une image ou dans une séquence d'images, permet la construction d'un schéma primaire (*primal sketch*) ;
- la reconstruction. La construction d'une représentation de l'espace centré sur l'observateur à partir de ces primitives de base vise à estimer les caractéristiques tridimensionnelles de la scène par rapport à l'observateur ;
- la reconnaissance. La représentation de l'espace précédente est ensuite fusionnée avec une connaissance 3D complémentaire pour obtenir une représentation de l'espace centrée sur la scène. Cette connaissance 3D complémentaire peut être la position de l'objet par rapport à la caméra.

Ces étapes permettent la transformation du signal image en une représentation symbolique de la scène observée. Malgré son ancienneté (1970), cette méthodologie est toujours d'actualité car elle a guidé les recherches en vision artificielle pendant des années et a conduit à de nombreuses avancées, notamment pour l'analyse des contours, des textures, des ombrages ou encore l'analyse du mouvement ou la stéréovision passive [Mar96, Dev97].

Aloimonos [AWB87], Bajcsy [Baj86] et Ballard [Bal91] ont modifié l'approche de Marr en proposant le concept de vision active. Cette dernière, telle qu'elle est décrite par Aloimonos, est une analyse théorique du processus de vision dont la motivation principale est l'optimisation, en réduisant la complexité, de la tâche visuelle à l'aide d'un capteur actif. Ces auteurs ont montré que les problèmes fondamentaux de la perception sont moins complexes à résoudre si l'on utilise un capteur actif. D'autres travaux plus récents font également appel aux idées de ces auteurs, en soulignant l'apport incontestable de la vision active. On peut citer [Bou93], [CBBJ96], [Mar96] qui ont traité ainsi le délicat problème de l'estimation de la structure 3D à partir du mouvement (« structure from controlled motion »). Dans ce cas, le contrôle du mouvement de la caméra est explicitement pris en compte afin d'optimiser la reconstruction 3D .

La mise en œuvre des systèmes de vision artificielle repose sur l'analyse de l'information contenue dans des images. L'entrée d'un tel système est constituée de plusieurs images et la sortie représente une description géométrique de la scène observée.

Les données tridimensionnelles d'une scène sont acquises, soit à partir des informations 2D issues d'une ou plusieurs caméras, soit directement par des capteurs 3D classiques (p. ex, un télémètre laser). Dans le domaine de la vision par ordinateur, plusieurs méthodes sont couramment utilisées pour obtenir une représentation tridimensionnelle de la scène observée. Dans la littérature, on peut ainsi distinguer, selon le nombre d'images nécessaires, deux classes d'approches. Les approches mono-image n'utilisent qu'une seule image pour reconstruire la scène perçue [Bou93, Alo88]. Ces approches se basent soit

sur l'intensité lumineuse [HB89], soit sur la texture [Alo88] ou encore sur les contours [DRLR89] pour obtenir la reconstruction 3D. Toutes ces techniques de vision monoculaire nécessitent des contraintes strictes d'éclairage ou de géométrie et sont souvent instables [Bou93]. Le deuxième type d'approches, appelé approches multi-images, fait appel à deux ou plusieurs images pour la détermination de la structure 3D de la scène. Ces dernières approches englobent deux catégories qui diffèrent selon le système de vision utilisé. La première réalise la reconstruction 3D à partir du mouvement en utilisant une seule caméra mobile [Via92] et est appelée « vision dynamique ». Pour une étude comparative complète des méthodes existantes, le lecteur pourra se référer à la thèse de Boukir [Bou93]. La seconde approche, plus classique, est la stéréovision, qui utilise un système de plusieurs caméras [Aya89, WCR92].

Analyser des images en stéréovision est une méthode très répandue pour l'extraction des structures 3D d'une scène. De ce fait, de nombreux chercheurs s'y sont intéressés. Simuler le mécanisme de perception utilisé par la plupart des être vivants était une des motivations initiales de l'étude de la stéréovision en vision artificielle. La performance du système visuel humain nous fait parfois oublier la complexité des tâches réalisées en continu pour reconstruire notre environnement 3D. En effet, le processus peut être décomposé en trois étapes :

- sélectionner l'élément à traiter dans la première vue ;
- identifier son correspondant dans la deuxième vue ;
- calculer la position 3D de l'élément en connaissant les points des vue par rapport à un référentiel absolu.

Un système de vision stéréoscopique comporte donc une chaîne de traitement qui doit permettre d'aboutir à l'information tridimensionnelle recherchée à partir des images acquises. Nous pouvons présenter les étapes de cette chaîne de traitement comme suit :

- *La modélisation et l'étalonnage du capteur.* Il s'agit de déterminer la transformation qui associe à un point de l'espace tridimensionnel sa projection dans l'image donnée par la caméra. Il s'agit d'un problème essentiel de la vision artificielle dès que l'on s'intéresse à la géométrie dans l'espace de la scène perçue à travers une ou plusieurs images.
- *L'extraction des indices visuels.* La segmentation d'images est l'opération de base de tout système de vision. Il est très difficile d'analyser une image brute contenant une trop grande quantité d'information. Le but de la segmentation est donc l'extraction des indices visuels pertinents de l'image acquise. Le choix de la représentation de l'objet est une étape qui reste encore un sujet d'importance pour les chercheurs du domaine. Certains utilisent les points élémentaires (pixels de l'image), d'autres préfèrent utiliser des primitives symboliques (régions, contours, etc.). L'extraction des indices visuels représentant ces attributs qui caractérisent des entités plus complexes est une étape incontournable du processus de vision stéréoscopique.
- *La mise en correspondance.* Cette étape d'identification et d'association des primitives extraites et provenant de la projection d'une même entité dans deux images distinctes s'avère être très complexe et nécessite une réflexion approfondie. La résolution de ce problème, dit d'appariement, fait l'objet de nombreuses recherches dans le monde entier, et les différentes approches se distinguent essentiellement par les contraintes imposées aux attributs présents dans les images d'une scène (segments,

1.2 Modélisation géométrique d'une caméra

courbes, objet lisse, contours d'occultation, etc.), et par les applications (robotique mobile, réalité virtuelle, photogrammétrie, etc.). Les diverses méthodes d'appariement dépendent du choix de l'application.

- *La reconstruction 3D.* L'appariement fournit des couples de points image appariés. La position du point 3D correspondant à cet appariement est déterminée classiquement par triangulation.

Après avoir introduit la vision stéréoscopique, nous allons mettre en lumière les points marquants de chaque étape de la chaîne de traitements permettant d'aboutir à une reconstruction tridimensionnelle.

1.2 Modélisation géométrique d'une caméra

La stéréovision binoculaire est une approche très répandue en vision artificielle. Le système possède deux capteurs passifs permettant d'acquérir des images de la scène à partir de deux points de vue différents.

Afin de pouvoir traiter les données acquises, il est nécessaire de définir un modèle géométrique du processus de formation de la vue obtenue. Ce modèle nous permettra de définir les paramètres d'étalonnage de la caméra. Ces paramètres peuvent être intrinsèques ou extrinsèques et permettent de définir la transformation homographique qui fait correspondre les points de l'espace aux points du plan de projection de la caméra.

Modèle projectif (*sténopé*)

La plupart des méthodes d'étalonnage reposent sur un modèle projectif. En effet, ce modèle permet de modéliser la plupart des capteurs projectifs. Cette modélisation géométrique revient à effectuer une projection centrale sur un plan 3D, appelé plan rétinien, puis un changement de coordonnées linéaires dans le plan 2D associé. C'est le modèle projectif linéaire appelé *sténopé* (petit trou ou *pinhole*)¹, modèle le plus fréquemment employé [Tsa86, Tos87, CR89], car le plus simple à manipuler. Dans ce modèle, l'optique est modélisée par une simple projection centrale, de centre O , sur le plan rétinien (voir figure 1.1). On suppose par ailleurs que tous les rayons de projection sont rectilignes et passent par le centre de la lentille. Les paramètres extrinsèques de la caméra sont la rotation et la translation (R, t) qui correspondent au déplacement entre le repère du monde R_{monde} et le repère de la caméra R_{camera} . Les paramètres intrinsèques sont associés à la nature du capteur utilisé, c'est-à-dire sa focale f , les coordonnées du centre optique O (voir figure 1.1) et les paramètres d'échantillonnage spatial (la géométrie du capteur est discrète et forme une rétine rectangulaire de paramètres k_u, k_v ; voir annexe A). Ceux-ci font correspondre le repère de la caméra x_c, y_c, z_c et le repère lié au plan image u, v . Un point P dans l'espace est repéré par rapport au référentiel du monde R_{monde} . Les paramètres extrinsèques de la caméra sont donnés par la matrice de rotation R et le vecteur de translation t qui correspondent au déplacement entre le repère de la caméra R_{camera} et le repère du monde R_{monde} . Le modèle sténopé permet de modéliser fidèlement la plupart

¹ Du grec « steno » qui signifie étroit et « ops » qui signifie œil : le centre optique de projection peut être défini dans un plan opaque percé d'un petit trou. Le terme anglais « pinhole camera », apparu en 1856, est, à cet égard, plus explicite : appareil photo avec un trou.

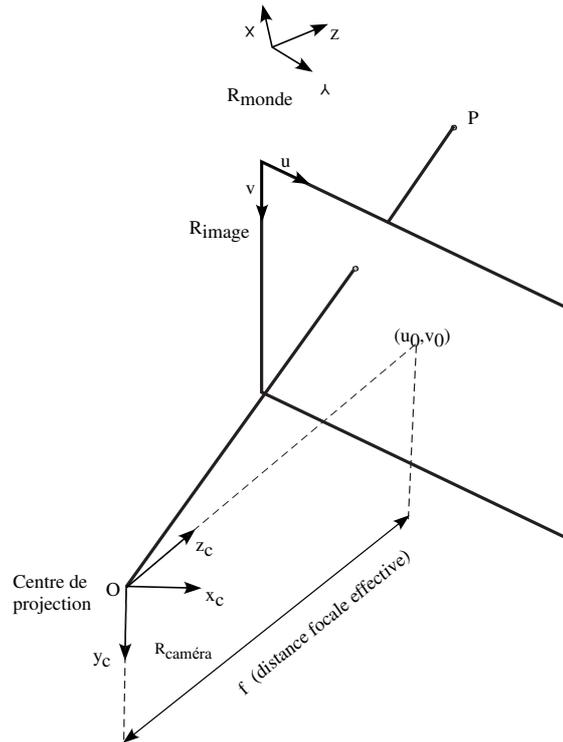


FIG. 1.1 – Modèle sténopé pour une caméra et ses paramètres

des capteurs projectifs. De plus, il permet de simplifier les mathématiques mises en jeu pour l'estimation des paramètres du modèle [Dev97].

D'autres modèles de projection ont été proposés dans la littérature. Citons entre autres le modèle de projection orthographique et le modèle de projection perspective faible [Bou94]. Le premier est un modèle très simple et approxime la géométrie d'une caméra. Les rayons optiques sont orthogonaux au plan image. Le modèle de projection perspective faible, quant à lui, suppose que les points subissent d'abord une projection orthographique sur le plan perpendiculaire à l'axe optique et passant par le centre de gravité des points de l'objet observé. Les points deviennent coplanaires et subissent ensuite une projection perspective sur le plan image. En comparant avec la projection orthographique, l'avantage de ce modèle est qu'il tient compte de la distance entre la caméra et l'objet observé. Dans ce cas, on peut dire qu'il s'agit d'une projection orthographique améliorée. Il s'agit d'une approximation très acceptable dans le cas où « l'épaisseur » de l'objet observé est relativement petite par rapport à la distance entre la scène et la caméra [Lin99, Bou94].

1.3 Étalonnage du capteur

La reconstruction 3D des primitives de l'environnement impose la détermination préalable des transformations géométriques liant le plan image à l'espace tridimensionnel. En effet, les mesures de la structure 3D des objets de l'environnement perçue par la caméra sont exprimées dans le repère de référence (R_{monde}) qui est différent du repère de projection (R_{camera}). Il est ainsi nécessaire de déterminer la transformation rigide liant le repère

1.3 Étalonage du capteur

de la scène au repère de la caméra.

Plus particulièrement, dans le cadre de cette étude, on s'intéresse à l'étalonnage d'un système stéréoscopique qui est formé par l'ensemble de deux caméras rigidement liées observant la même scène. Le système stéréoscopique est constitué de deux caméras indépendantes et on peut appliquer une procédure classique d'étalonnage, utilisant une ou plusieurs images d'une mire ou d'un objet étalon de géométrie connue [Dev97].

Le principe de l'étalonnage est d'utiliser une grille (mire) ou tout autre objet, pour lesquels les positions de points marqués sur l'objet, appelés points d'étalonnage (ou de contrôle), sont connues et dont on observe la position dans les images. Ces points d'étalonnage sont sélectionnés dans des zones particulières de l'image (coins de la mire [Tos87], intersections de droites [BMB94], points de la mire [Wil94]), ou peuvent être toute autre primitive extraite des images. À partir de cette observation et connaissant les points de l'espace à l'origine des points projetés dans l'image, on peut en déduire les paramètres de la transformation qui amène un point appartenant à la mire vers le point correspondant dans l'image.

L'étalonnage est donc un processus qui permet d'estimer les paramètres d'un modèle de caméra. Plus précisément, le problème de l'étalonnage peut être présenté comme suit : étant donné un ensemble de points de coordonnées 3D (X, Y, Z) connues, il s'agit de déterminer les paramètres de la projection associée au capteur pour que leurs projections correspondent au mieux aux points correspondants extraits des images. La fonction de projection peut être décomposée en deux transformations distinctes : la première est la transformation entre le repère de l'espace 3D et le système de coordonnées de la caméra ; ce déplacement 3D qui englobe une rotation et une translation forme ce qu'on appelle les paramètres extrinsèques du capteur. La seconde est la transformation entre les coordonnées des points 3D, exprimées dans le repère attaché à la caméra et les points 2D dans le plan image de la caméra.

1.3.1 Méthodes d'étalonnage

Ce paragraphe présente une synthèse non exhaustive des principales méthodes d'étalonnage mises au point dans le cadre de travaux portant sur la vision par ordinateur.

1.3.1.1 Étalonage selon Tsai

Cette méthode est devenue une référence depuis de nombreuses années. Effectivement, un grand nombre de méthodes en reprennent les fondements et son modèle de caméra. Roger Tsai [Tsa86, Tsa91] utilise un déplacement du plan d'étalonnage et modélise la formation de l'image en déterminant des facteurs d'échelles qui permettent le passage entre le plan image et l'image échantillonnée. Salvi [SAB02] a renommé cette technique « étalonnage à deux étapes ». La première étape utilise une modélisation simplifiée de la caméra, tandis que la deuxième raffine l'étalonnage en estimant un coefficient qui exprime la distorsion radiale du système optique [Tsa87]. Dans cette méthode, la mire utilisée est le damier repris dans un grand nombre d'autres approches. Le plan d'étalonnage doit se déplacer de façon parallèle à sa position initiale dans le but d'utiliser une contrainte de parallélisme et d'éliminer les distorsions axiales [Bén02].

1.3.1.2 Étalonnage selon Faugeras et Toscani

Cette méthode d'étalonnage est parmi les méthodes les plus utilisées [Tos87]. Toscani a développé sa procédure d'étalonnage d'une caméra en utilisant un formalisme mathématique couramment employé et d'implémentation immédiate. Par l'utilisation d'une mire d'étalonnage et par l'exploitation des équations de la projection perspective linéaire du modèle sténopé, Toscani extrait les paramètres intrinsèques et extrinsèques par la résolution d'un système à $2n$ équations, n étant le nombre de points non coplanaires appartenant à la mire d'étalonnage (n doit être supérieur à 6). Une contrainte physique est adoptée pour résoudre ce système d'équations. Elle repose sur des contraintes d'orthogonalité de la matrice de rotation qui ne subit pas l'influence des paramètres intrinsèques. Cette contrainte est appelée contrainte de Faugeras-Toscani [FT86].

1.3.1.3 Étalonnage à base d'ellipses

Une méthode d'étalonnage de caméra à base d'ellipses a été présentée par Tarel [Tar96]. Tarel a relié la description d'un disque elliptique de l'espace aux moments géométriques de son image. Sa méthode n'est pas basée sur une correspondance entre points mais repose sur l'utilisation des moments d'une ellipse.

Cette technique a l'avantage d'être robuste au bruit de l'image et fournit une précision d'étalonnage inférieure à 1/10 de pixel sur la position du barycentre. Les moments géométriques d'une région ont été choisis comme données à fournir à l'algorithme d'estimation, car ils peuvent être calculés dans une image avec précision et robustesse en se basant sur la photométrie des surfaces. La relation liant exactement la description d'une ellipse de l'espace 3D et les moments d'ordre inférieurs ou égaux à deux de sa projection est explicitée en fonction des paramètres du capteur. L'estimation est un problème non-linéaire que l'auteur résout avec un algorithme itératif de type Gauss-Newton.

1.3.1.4 Étalonnage selon Zhang

Cette méthode est basée sur la détection d'une mire d'étalonnage représentant un damier dont les caractéristiques sont connues et mesurées. L'utilisation d'une mire plane simplifie sa fabrication et permet une simplification du modèle à adopter car la coordonnée Z du point dans l'espace est annulée en liant le repère du monde au plan. Zhang [Zha99, Zha00] a proposé une nouvelle technique flexible pour étalonner facilement une caméra. La technique utilise une mire d'étalonnage vue à partir d'au moins trois points de vue d'orientations différentes. La caméra et la mire peuvent être déplacées librement sans avoir besoin de connaître le mouvement. La distorsion radiale de l'objectif est modélisée. Le processus d'étalonnage proposé se compose d'une recherche d'une solution initiale, suivie d'une optimisation non-linéaire basée sur le critère du maximum de vraisemblance. Cette méthode reprend l'approche proposée par Tsai, mais est d'une utilisation plus souple. En effet, le fait de pouvoir utiliser simplement un plan déplacé d'une image à l'autre sans que le mouvement ait besoin d'être connu lui donne une grande flexibilité.

1.3 Étalonage du capteur

1.3.2 Auto-étalonage

Par auto-étalonage, on entend le fait qu'aucune connaissance *a priori* sur la scène observée n'est nécessaire. C'est la technique d'étalonage la plus récente qui ne nécessite pas la connaissance de points de coordonnées 3D. Ces techniques sont également connues comme étant des méthodes dites *d'auto-calibration* [JNS04] et permettent de retrouver la structure euclidienne de l'espace à un facteur d'échelle global près. Ce type d'approches ne demande pas d'intervention manuelle, ni de mire d'étalonage. L'auto-calibration est fortement adaptée aux systèmes à vision active.

Les méthodes d'auto-étalonage sont progressivement apparues [FLM92] pour apporter la souplesse d'utilisation faisant défaut aux méthodes d'étalonage classiques. Le principe de cette technique est de remplacer la mire par la scène 3D elle-même. La seule connaissance *a priori* utilisée sur la scène est sa rigidité. L'inconvénient de ces algorithmes, non-linéaires par essence, est l'instabilité rendant leur mise en application difficile. Plus particulièrement, certains mouvements de caméra conduisent à une dégénérescence des contraintes [GS03, MBFM05].

Dans toutes ces techniques d'étalonage, la résolution des équations mises en œuvre passe par des algorithmes d'optimisation numérique minimisant une fonction de coût appropriée. Mais, les résultats de ces algorithmes dépendent du choix de cette fonction de coût et du choix de sa paramétrisation.

1.3.3 Étalonage faible

L'étalonage faible consiste à étalonner deux caméras en l'absence d'une connaissance *a priori* de la scène observée et des paramètres des caméras. À partir d'une liste de correspondances de points entre les images des deux caméras observant la même scène, on peut étalonner le capteur. La définition de la relation entre les deux caméras va permettre d'établir ce que l'on appelle la « géométrie épipolaire », qui permet de définir la relation qui existe entre les pixels des deux images, de manière à faciliter la mise en correspondance [HZ00]. Dans le cas où les paramètres intrinsèques des caméras sont connus, la géométrie épipolaire est donnée par la matrice essentielle [LH87]. Le calcul de cette matrice revient à déterminer sa structure à partir de correspondances [DF96]. Le deuxième cas est celui où l'on ne dispose pas des paramètres intrinsèques. Le calcul de la géométrie épipolaire se fait alors par l'intermédiaire de la matrice fondamentale [DF96] qui lie toujours les coordonnées pixels des images et qui fournit les informations permettant de simplifier le problème de mise en correspondance de points entre les images, pour aboutir ensuite à la reconstruction 3D.

Utiliser des points en correspondance dans une paire d'images permet d'estimer la matrice fondamentale, donc de déterminer la géométrie épipolaire dans les repères liés aux images. Plusieurs auteurs ont élaboré une reconstruction tridimensionnelle à partir d'un étalonage de ce type, dit faible. Zhang et *al* [ZLF93] ont exprimé la matrice fondamentale à partir de l'appariement d'ensembles de points dans les deux images. On obtient ainsi les paramètres nécessaires pour mener à bien une reconstruction projective de la scène. Dans le même ordre d'idée, Deverney et Faugeras [DF96] partent eux aussi d'une paire de caméras faiblement étalonnées, *i.e.* dont on ne connaît que la géométrie épipolaire. L'idée

est sensiblement la même : à partir d'un tel montage, la reconstruction projective fournit des points aux coordonnées homogènes ². En effectuant un déplacement du capteur, on obtient un second ensemble de points reconstruits de manière projective. Il s'agit ensuite de déterminer la transformation nécessaire pour amener ces deux ensembles de points, exprimés dans leur espace projectif, dans un repère euclidien commun.

1.4 Géométrie épipolaire et matrice fondamentale

La matrice fondamentale joue ainsi un rôle clef en vision 3D par ordinateur. La matrice fondamentale est de taille (3×3) et a 7 degrés de liberté. Son estimation est primordiale pour beaucoup d'algorithmes de mise en correspondance et de reconstruction 3D. La plupart des algorithmes d'estimation de cette matrice reposent sur des correspondances de points. La géométrie dite des deux vues est décrite par la géométrie épipolaire qui encapsule le positionnement relatif des deux caméras. Nous commençons par définir explicitement la géométrie épipolaire (voir figure(1.2)).

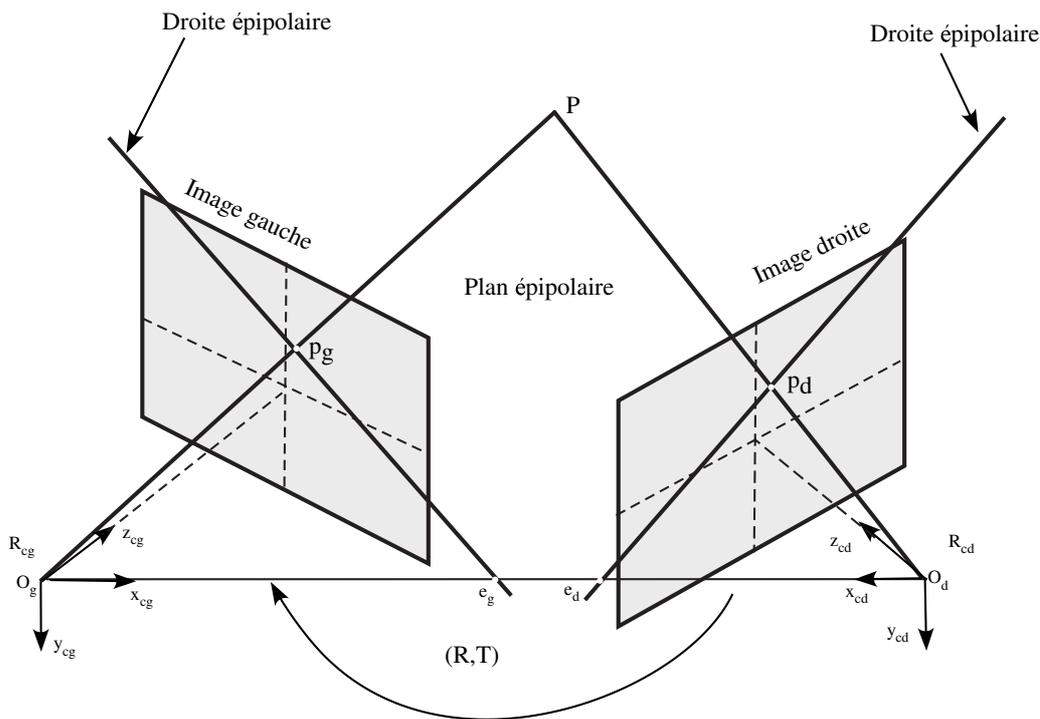


FIG. 1.2 – La géométrie épipolaire

²Soit M un point du plan, de coordonnées cartésiennes x, y . On appelle coordonnées homogènes de M tout triplet de réels X, Y , et T tels que : $x = X/T$, et $y = Y/T$. Les coordonnées homogènes ne sont pas uniques, et sont définies à une constante multiplicative près. Un point M pour lequel $T = 0$ est un point à l'infini, et l'ensemble de tous ces points s'appelle la droite à l'infini.

1.4.1 La géométrie épipolaire

La géométrie épipolaire est intrinsèque à tout système de deux caméras quel que soit le modèle utilisé pour ces caméras. Elle a été introduite par Longuet-Higgins [LH87]. Cette géométrie décrit une relation géométrique entre deux images prises sous deux angles de vue différents. La précision de l'estimation de la géométrie épipolaire est très importante parce qu'elle conditionne la précision des algorithmes de mise en correspondance entre points d'un couple d'images stéréoscopiques, ces algorithmes reposant souvent sur la connaissance *a priori* de cette géométrie.

Soient p_g, p_d , les deux projections gauche et droite correspondant au même point 3D P de la scène. O_g et O_d sont les centres optiques correspondant aux caméras gauche et droite respectivement. La matrice fondamentale conduit à l'expression algébrique de la contrainte de coplanarité des points (p_d, p_g, O_d, O_g) , exprimée à partir des images gauche et droite. Cette contrainte est aussi appelée contrainte épipolaire.

On peut facilement observer que le point p_g de l'image gauche représente la projection de tous les points de l'espace appartenant à la droite $(O_g P)$. Les correspondants possibles de p_g dans l'image droite sont contraints de se trouver sur une droite, appelée droite épipolaire. Cette dernière n'est autre que la projection de la droite $(O_g P)$ dans l'image droite. Il en découle que l'ensemble des droites épipolaires de l'image droite se coupent en un même point qui est la projection de O_g dans le plan image droit. Ce point particulier s'appelle l'épipole e_d . De la même manière, la projection de O_d dans le plan image gauche sera appelée épipole e_g . En d'autres termes, un point P de l'espace définit avec les deux centres de projection O_g et O_d un plan appelé plan épipolaire. Ce dernier coupe les deux plans images en deux droites qui sont les droites épipolaires. Tous les plans épipolaires passent par la droite $(O_g O_d)$ et donc par les épipoles e_g et e_d . Cette relation entre les deux images gauche et droite est appelée géométrie épipolaire et donne une contrainte pour la mise en correspondance : étant donné un point p_g de l'image gauche, son correspondant p_d dans l'image droite se trouve obligatoirement sur la droite épipolaire correspondante.

1.4.2 Expression matricielle de la matrice fondamentale

Supposons que le repère de la caméra gauche R_{cg} devienne le repère absolu de la scène. La matrice de rotation R et le vecteur de translation t caractérisent le passage du repère lié à la caméra gauche vers celui de la caméra droite. Les équations de projection dans les deux images deviennent :

$$\tilde{p}_d = K_d \cdot [I \ 0] \cdot \tilde{P} \quad (1.1)$$

$$\tilde{p}_g = K_g \cdot [R \ t] \cdot \tilde{P} \quad (1.2)$$

Où I est la matrice identité de taille (3×3) . \tilde{p}_d, \tilde{p}_g sont respectivement les pixels des images I_d, I_g exprimés en coordonnées homogènes, \tilde{P} le point correspondant de l'espace et K_g, K_d respectivement les matrices de paramètres intrinsèques des caméras gauche et droite. En éliminant \tilde{P} dans les deux équations, on obtient l'équation suivante liant \tilde{p}_d et \tilde{p}_g :

$$\tilde{p}_g^t \cdot K_g^{-t} \cdot \hat{t} \cdot R \cdot K_d^{-1} \cdot \tilde{p}_d = 0 \quad (1.3)$$

Où \hat{t} est la matrice « produit vectoriel » associée à la translation t . La matrice $K_g^{-t} \cdot \hat{t} \cdot R \cdot K_d^{-1}$ ainsi obtenue est appelée matrice fondamentale et est généralement

notée F . C'est une matrice définie à un coefficient multiplicatif près. On voit que cette matrice est définie uniquement par la connaissance des caractéristiques des deux caméras (paramètres intrinsèques) ainsi que par leur position relative l'une par rapport à l'autre. Les deux équations simples reliant les projetés dans les deux images sont donc les suivantes :

$$\tilde{p}_g^t \cdot F \cdot \tilde{p}_d = 0 \quad (1.4)$$

$$\tilde{p}_d \cdot F^t \cdot \tilde{p}_g = 0 \quad (1.5)$$

où F^t est la matrice transposée de F .

Les épipoles dans chaque image correspondent alors aux vecteurs propres associés aux valeurs propres nulles de F et F^t . Un lien direct peut être établi entre cette matrice et la matrice essentielle E [Ois98], qui ne prend pas en compte les caractéristiques des caméras [LH87]. La matrice essentielle s'exprime simplement à l'aide de la rotation et de la translation liant les repères des caméras et traduit la relation unissant deux points p_g et p_d des plans images exprimés en coordonnées normalisées ³ :

$$\tilde{m}_g^t \cdot E \cdot \tilde{m}_d = 0 \quad (1.6)$$

avec $E = [R \ t]$ et \tilde{m}_g, \tilde{m}_d les deux points image p_g, p_d exprimés dans les repères caméras gauche et droite respectivement.

L'expression de F en fonction de E est alors la suivante :

$$F = K_g^{-t} \cdot E \cdot K_d^{-1} \quad (1.7)$$

1.4.3 Méthodes d'estimation de la matrice fondamentale

De nombreuses méthodes d'estimation de la matrice fondamentale existent et presque toutes n'utilisent que des correspondances de points image comme données. Ces méthodes se distinguent par la paramétrisation de F ainsi que par la méthode d'optimisation. L'estimation de la matrice fondamentale repose sur la connaissance d'un certain nombre de paires de points en correspondance. Ces dernières années, plusieurs méthodes d'estimation de la matrice fondamentale ont été proposées. Une synthèse des différentes méthodes d'estimation de la géométrie épipolaire est donnée par Zhang dans [Zha96] et par Xavier et Joaquim dans [XJ03]. Après Hartley [HZ00], Xavier et Joaquim ont étudié 19 méthodes pour calculer cette matrice. Ces méthodes peuvent être classées en méthodes linéaires, itératives ou robustes. Les études ont montré que les méthodes linéaires, et notamment celle introduite par Longuet-Higgins [LH87], sont très sensibles au bruit dû aux faux appariements [ZDFL95]. Les méthodes linéaires souffrent également de deux défauts, liés à l'absence de contrainte sur le rang de la matrice fondamentale et à l'absence de normalisation du critère de minimisation, ce qui entraîne des erreurs dans l'estimation de la matrice fondamentale. Les méthodes itératives, comme par exemple celles utilisant la technique d'optimisation de Levenberg-Marquardt [HZ00], peuvent, quant à elles, générer une mauvaise localisation des points dans l'image, à cause du bruit introduit par

³Soit m un point exprimé dans le repère caméra par les coordonnées (x_c, y_c, z_c) . x_c/z_c et y_c/z_c sont appelées les coordonnées normalisées de m .

1.4 Géométrie épipolaire et matrice fondamentale

la segmentation des images. Par contre, les méthodes robustes, comme par exemple la technique des *M-Estimators* [HZ00], sont capables de donner un résultat précis avec des images bruitées, et peuvent aussi prendre en compte les points aberrants (p. ex, les faux appariements entre points d'un couple d'images stéréoscopiques). Xavier et Joaquim ont montré que les meilleurs résultats sont obtenus lorsqu'on l'utilise des critères non-linéaires.

L'équation (1.4) est la relation liant un point $p_{di} (u_{di}, v_{di})^t$ de l'image droite à un point $p_{gi} (u_{gi}, v_{gi})^t$ de l'image de gauche. Cette équation peut se réécrire en coordonnées homogènes sous une forme linéaire :

$$u_i^t \cdot F = 0 \quad (1.8)$$

Où

$$u_i = (u_{ig} \cdot u_{id}, u_{ig} \cdot v_{id}, u_{ig}, v_{ig} \cdot u_{id}, v_{ig} \cdot v_{id}, v_{ig}, u_{id}, v_{id}, 1)^t \quad (1.9)$$

et

$$F = [F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}]^t \quad (1.10)$$

Avec F_{ij} l'élément de la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne de la matrice F . Cette équation est donc valable pour chaque paire de points en correspondance. Elle se généralise alors pour un ensemble de n couples de points :

$$U_n \cdot F = 0 \quad (1.11)$$

Où

$$U_n = [u_1, \dots, u_n]^t \quad (1.12)$$

Cette équation primordiale est un point de départ pour la plupart des méthodes de détermination de la matrice fondamentale.

1.4.3.1 Méthodes linéaires

La méthode linéaire dite des sept points est basée sur le calcul de la matrice fondamentale en utilisant seulement sept points en correspondance entre les deux images formant le couple stéréoscopique [Zha96]. L'avantage majeur de cette méthode est sa simplicité : seuls sept points sont nécessaires pour l'estimation de F . Cependant, cela devient un inconvénient lorsque quelques points sont mal localisés [XJ03]. En pratique, on dispose de plus de 7 points se correspondant. Si l'on ignore la contrainte sur le rang de la matrice F , qui est égal à 2, on peut utiliser la méthode des moindres carrés pour résoudre l'équation suivante :

$$\min_F \sum_i (p_{gi}^t \cdot F \cdot p_{di})^2 \quad (1.13)$$

En imposant une contrainte rendant la norme de F égale à 1, le problème devient un problème de minimisation classique sous la forme suivante [Zha96] :

$$\left\{ \begin{array}{l} \min_F \|U_n \cdot F\|^2 \\ \|F\| = 1 \end{array} \right. \quad (1.14)$$

La résolution est alors effectuée en employant la technique des multiplicateurs de Lagrange.

$$\min_F \psi(F, \lambda) \quad (1.15)$$

avec

$$\psi(F, \lambda) = \|U_n \cdot F\|^2 + \lambda \cdot (1 - \|F\|^2) \quad (1.16)$$

$$U_n^t \cdot U_n \cdot F = \lambda \cdot F \quad (1.17)$$

La solution pour F est le vecteur propre qui correspond à la plus petite valeur propre λ . L'estimation de la matrice fondamentale peut se faire plus simplement en utilisant l'*algorithme des huit points*. Mais la solution obtenue n'est pas forcément optimale. Par contre, la matrice fondamentale possède par construction deux propriétés intéressantes : son rang est égal à 2 et elle est de dimensions 3×3 . En utilisant ces caractéristiques liées à la nature de la matrice, il est alors possible d'améliorer les méthodes d'estimation de F . Il existe une solution *a posteriori* pour trouver une matrice de déterminant nul \hat{F} proche de F . On estime la proximité des deux matrices par la norme dite de Frobenius. Pour obtenir \hat{F} , la matrice F est décomposée sous la forme suivante par une technique de type SVD (*Singular Value Decomposition*) :

$$F = U \cdot S \cdot V^t \quad (1.18)$$

Où $S = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ est une matrice diagonale avec $\sigma_1 > \sigma_2 > \sigma_3$ et U et V des matrices orthogonales. On peut alors démontrer que la matrice $\hat{F} = U \cdot \hat{S} \cdot V^t$ avec $\hat{S} = \text{diag}(\sigma_1, \sigma_2, 0)$ est la matrice de rang 2 qui minimise la norme de Frobenius de $F - \hat{F}$. Cet algorithme a été perfectionné par Hartley [Har97] pour le rendre encore plus robuste. Ainsi, il a proposé un algorithme dit des huit points normalisés. Il a en effet montré que l'application de l'algorithme des 8 points est souvent instable. La solution qu'il propose est de remplacer l'origine dans chacune des images par le centroïde des points appariés. Ensuite, on applique un facteur d'échelle de telle façon à ce que la norme moyenne des vecteurs associés aux points soit égale à $\sqrt{2}$. Ces deux opérations reviennent à multiplier les points de l'image gauche (droite) par une matrice (3×3) . Cette approche a grandement amélioré le résultat de la méthode des huit points.

1.4.3.2 Méthodes itératives

Une méthode élégante consiste à minimiser la distance entre les points et leurs droites épipolaires correspondantes. Soit $(p_{gi}, p_{di}), (i = 1, \dots, n)$ un ensemble d'appariements de points images. L'équation $p_{di}^t \cdot F \cdot p_{gi} = 0$ est directement liée à la distance d du point p_{di} à sa droite épipolaire $L_d = F \cdot p_{gi}$. Une première idée est alors d'utiliser un critère non-linéaire minimisant la somme :

$$\sum_{i=1}^n d^2(p_{di}, F \cdot p_{gi}) \quad (1.19)$$

avec :

$$d(p_{di}, F \cdot p_{gi}) = \frac{|p_{di}^t \cdot F \cdot p_{gi}|}{\sqrt{(F \cdot p_{gi})_1^2 + (F \cdot p_{di})_2^2}} \quad (1.20)$$

où le terme $(F \cdot p_{gi})_i$ est $i^{\text{ème}}$ élément du vecteur $F \cdot p_{gi}$. Pour donner un rôle symétrique aux deux images et obtenir une géométrie épipolaire cohérente, il convient de minimiser le critère :

$$\sum_{i=1}^n \{d^2(p_{di}, F \cdot p_{gi}) + d^2(p_{gi}, F^t \cdot p_{di})\} \quad (1.21)$$

où :

$$d(p_{gi}, F^t \cdot p_{di}) = \frac{|p_{gi}^t \cdot F^t \cdot p_{di}|}{\sqrt{(F \cdot p_{di})_1^2 + (F \cdot p_{di})_2^2}} \quad (1.22)$$

est la distance du point p_{gi} à sa ligne épipolaire $F^t \cdot p_{di}$. Compte tenu du fait que :

$$p_{gi}^t \cdot F^t \cdot p_{di} = p_{di}^t \cdot F \cdot p_{gi} \quad (1.23)$$

on aboutit finalement à la minimisation du critère suivant, qui est un critère normalisé :

$$\sum_{i=1}^n \left\{ \frac{(p_{di}^t \cdot F \cdot p_{gi})^2}{(F \cdot p_{gi})_1^2 + (F \cdot p_{gi})_2^2} + \frac{(p_{di}^t \cdot F \cdot p_{gi})^2}{(F^t \cdot p_{di})_1^2 + (F^t \cdot p_{di})_2^2} \right\} \quad (1.24)$$

La minimisation de ce critère est un problème de minimisation non-linéaire que l'on peut résoudre par une méthode itérative du type Levenberg-Marquardt [APS01]. La valeur initiale peut être fournie par une méthode linéaire. Une autre façon d'améliorer le critère est la suivante. Les coordonnées des appariements images (p_{di}, p_{gi}) sont généralement entachées d'un certain bruit. On peut alors multiplier chaque terme $(p_{di}^t \cdot F \cdot p_{gi})^2$ dans l'équation 1.24 par un coefficient, de telle sorte que tous les termes aient la même covariance. Si le bruit est de moyenne nulle, uniformément réparti et indépendant selon chaque axe de l'image, alors on montre que le nouveau critère à minimiser est le critère normalisé suivant [Luo92] :

$$\sum_{i=1}^n \left\{ \frac{(p_{di}^t \cdot F \cdot p_{gi})^2}{(F \cdot p_{gi})_1^2 + (F \cdot p_{gi})_2^2 + (F^t \cdot p_{di})_1^2 + (F^t \cdot p_{di})_2^2} \right\} \quad (1.25)$$

Ce critère à minimiser est semblable au précédent. Il peut également être minimisé par une méthode numérique itérative du type Levenberg-Marquardt. D'autre part, Chojnacki *et al.* [CBvdHG04] ont proposé récemment deux nouvelles méthodes itératives basées sur une estimation faisant appel au maximum de vraisemblance (*maximum likelihood*). Ces méthodes sont appelées *Fundamental Numerical Scheme* (FNS) et *Constrained Fundamental Numerical Scheme* (CFNS). Elles sont basées sur la technique de minimisation de Newton-Raphson, la matrice fondamentale étant estimée par la minimisation du critère suivant :

$$\sum \frac{(p_{di}^t \cdot F \cdot p_{gi})^2}{p_{di}^t \cdot F \cdot F^t \cdot p_{di} + p_{gi}^t \cdot F \cdot F^t \cdot p_{gi}} \quad (1.26)$$

Cette technique donne de très bons résultats, comparés à ceux obtenus avec des méthodes linéaires et des méthodes itératives qui minimisent la distance entre les points et les droites épipolaires. Bien que les méthodes itératives soient plus précises que les méthodes linéaires, elles ne peuvent s'affranchir des points aberrants.

1.4.3.3 Méthodes robustes

La première méthode robuste abordée ici repose sur la technique des M-estimateurs (*M-Estimators*) [HZ00]. Cette méthode conduit à un bon résultat en présence d'un bruit gaussien au niveau des points sélectionnés de l'image, mais elle est limitée dans sa capacité à prendre en compte des points aberrants. Deux autres techniques sont classées parmi les méthodes robustes et sont similaires : la méthode dite *Random Sampling Consensus* (RANSAC) et la méthode dite *Least Median Squares* (LMeds). Ces deux dernières consistent à sélectionner aléatoirement l'ensemble de points utilisés pour l'approximation de la matrice fondamentale. La méthode du LMeds calcule pour chaque estimation de F la distance euclidienne entre les points et les droites épipolaires, et le choix de F correspond à la minimisation de cette distance. La méthode RANSAC, quant à elle, calcule pour chaque valeur de F le nombre de points pouvant convenir (*inliers*). La matrice F choisie est celle qui maximise ce nombre. Une fois les points aberrants éliminés, la matrice F est recalculée pour obtenir une meilleure estimation. Les expériences ont montré que la technique LMeds donne un meilleur résultat que la méthode RANSAC en terme de précision [XJ03].

Récemment, d'autres méthodes d'estimation de F ont été développées. Parmi celles-ci, on peut citer la méthode MLESAC (*Maximum Likelihood Sample Consensus*) [LF98], ou encore la technique MAPSAC (*Maximum A Posteriori Sample Consensus*) [Tor02, ATLB04]. Ces méthodes robustes sont basées sur la technique RANSAC et elles sont robustes face aux mauvais appariements.

Nous avons ainsi donné un état de l'art des méthodes d'estimation de la matrice fondamentale les plus courantes. Nous avons vu que cette matrice modélise la géométrie de deux caméras non étalonnées. Son estimation est souvent faite à partir de listes de correspondances de points entre deux images. En raison de la simplicité des calculs, la méthode que nous avons retenue pour l'estimation de la matrice fondamentale est celle dite des huit points normalisés.

Cette matrice peut ensuite être utilisée pour reconstruire la structure 3D par triangulation des correspondances d'indices visuels.

1.5 Extraction des indices visuels

En stéréovision, l'étape de la mise en correspondance d'indices image suppose le traitement d'une grande quantité de données que l'on est tenté de réduire. De ce fait, on ne garde que les données nécessaires à l'élaboration d'une méthode d'appariement des données extraites de deux images stéréoscopiques. Une telle simplification est basée sur les indices visuels [Aya89], qui décrivent de manière succincte l'information pertinente contenue dans les images.

Analyser une image brute s'avère être une tâche très difficile car l'image contient beaucoup d'informations qui ne sont pas nécessairement pertinentes. Deux étapes sont primordiales pour résoudre la plupart des problèmes posés par l'analyse d'image. Extraire des images une représentation symbolique est considéré comme étant la première étape ; la seconde étape revient à étudier cette représentation symbolique afin de résoudre une tâche particulière.

On appelle indice visuel tout objet extrait de l'image et contenant de manière compacte

1.5 Extraction des indices visuels

l'information pertinente pour son analyse [Lux85]. Ces indices peuvent être des points d'intérêts, des régions, des contours ou bien des segments. Dans les paragraphes suivants, nous présentons différents types d'indices visuels, en montrant leur intérêt pour la mise en correspondance en stéréovision.

1.5.1 Points d'intérêts

Les points d'intérêts sont des points pertinents qui sont nécessairement représentatifs. Par exemple, on peut utiliser les points faisant la jonction entre plusieurs régions, ou bien des points particuliers le long d'une courbe comme les points anguleux. Ces points doivent pouvoir être appariés facilement dans l'autre image. Ils ont un voisinage particulier, facile à identifier sans ambiguïté par une mesure de corrélation. En fait, les points d'intérêts seuls ne sont pas d'un grand apport pour une reconstruction 3D. Ils sont généralement utilisés dans le processus d'étalonnage du capteur pour trouver le modèle de la caméra : dans une paire d'images stéréoscopiques, les points d'intérêt peuvent servir à créer la liste d'appariements initiale ; on est alors en mesure d'estimer la géométrie épipolaire, puis d'obtenir les paramètres des caméras.

1.5.2 Régions

L'objectif de la segmentation d'une image en régions est de partitionner l'image en zones correspondant à des entités de la scène observée [CMB⁺95]. Une région est définie comme étant un ensemble de points connexes qui vérifient un critère d'homogénéité. De nombreuses propriétés sont attachées à cet ensemble de points, comme la longueur des frontières, la surface, le centre de gravité, l'intensité moyenne, etc. Du fait qu'elles sont extraites à partir de propriétés photométriques, les régions sont sensibles aux variations d'intensité lumineuse. Ainsi, dans un couple d'images stéréoscopiques, les régions peuvent avoir des formes et des tailles différentes. La mise en correspondance de régions est donc une mise en correspondance d'un ensemble de pixels. Dans la littérature, la mise en correspondance de régions est bien moins utilisée que l'appariement exploitant d'autres primitives géométriques (indices visuels). La *segmentation en régions* est consommatrice en temps de calcul [Ghe92]. Par ailleurs, les primitives de type région couvrent tous les points de l'image et, en appariant les régions, on apparie tous les points de l'image, ce qui permet d'avoir une reconstruction 3D dense (tous les points des images seront reconstruits dans l'espace 3D).

Les algorithmes de segmentation en régions peuvent être classés selon deux grandes classes [HM93, Rou99]. Ceux qui réalisent une partition de l'image en utilisant par exemple une décomposition en quadarbres (quadtrees, en anglais) sont appelés *algorithmes de segmentation en régions par classification*. Dans ces méthodes, on associe à chaque pixel la classe de niveaux de gris à laquelle il appartient. De ce fait, les régions sont constituées d'ensembles de pixels connexes appartenant à une même classe. Cette méthode donne de bons résultats pour les images comprenant un nombre peu important d'objets ayant des niveaux d'intensités distincts [HM93].

La deuxième classe des algorithmes de segmentation en régions procède par agrégation itérative de pixels connexes possédant des caractéristiques similaires. Ces algorithmes associent à chaque pixel un vecteur de propriétés, deux pixels étant regroupés si leurs vecteurs de propriétés sont suffisamment similaires. Ils sont appelés *algorithmes de croissance de régions*. L'algorithme de division-fusion (*Split and Merge*) est une variante de ces algorithmes de croissance de régions. Son principe repose sur la subdivision successive de l'image en sous-régions jusqu'à obtention d'une zone homogène. Puis les régions adjacentes sont fusionnées deux par deux tant que l'ensemble regroupé est homogène [GWES02, DR03].

Les algorithmes de segmentation en régions par classification et de croissance de région permettent de donner un ensemble de pixels étiquetés en fonction de la région à laquelle ils appartiennent.

Vezien [JT95] a proposé une approche générique pour réaliser la reconstruction 3D de facettes planes à partir d'une paire d'images stéréoscopiques segmentées en régions. L'avantage de cette méthode réside dans la possibilité d'utiliser différentes caractéristiques pour les régions (moments d'inertie, représentation paramétrée ou non du contenu photométrique).

1.5.3 Contours

Les avantages, qui rendent la mise en correspondance de primitives de type contour très intéressante, sont très nombreux. Tout d'abord, ils conservent une grande partie du contenu sémantique de l'image. De plus, le temps de calcul des algorithmes de mise en correspondance exploitant des points de contour est réduit, à cause du nombre réduit de points de contour contenus dans l'image. Ce dernier avantage est très appréciable pour des applications qui exigent des traitements en temps réel. Un autre avantage de l'utilisation des primitives de type contour réside dans la précision des mesures que l'on peut effectuer sur les contours ; ce qui apporte de la fiabilité aux applications, par exemple dans le domaine du contrôle dimensionnel. Enfin, les contours sont robustes aux variations géométriques et photométriques.

On définit les contours comme étant les discontinuités photométriques présentes dans les images. Pour la détection des points de contour, on applique par exemple un opérateur dérivateur sur l'image. On recherchera ensuite les extréma locaux du gradient, ou bien les passages par zéro d'un Laplacien des niveaux de gris. Une image est forcément entachée d'une certaine quantité de bruit. De ce fait, il est nécessaire d'appliquer un opérateur de filtrage avant de pouvoir isoler les maxima de gradient ou les passages par zéro du Laplacien. Canny [Can86] a construit une approche quantitative de la détection de points de contour en déterminant les critères que doit respecter un opérateur de détection de points de contour [HM93]. Ces critères sont :

- *une bonne détection*. L'opérateur doit donner une réponse au voisinage d'un contour,
- *une bonne localisation*. La localisation de la réponse de l'opérateur doit idéalement se situer sur la transition, même si celle-ci est noyée dans un bruit. Le contour doit être localisé avec précision,

1.5 Extraction des indices visuels

- *une unicité de la réponse*. Une transition aussi bruitée soit-elle ne doit engendrer qu'une seule réponse de l'opérateur (un seul maximum du gradient ou un seul passage par zéro du Laplacien). Un contour doit provoquer une seule réponse de l'opérateur d'extraction. C'est ce critère qui justifie la nécessité d'un filtrage limitant le nombre de réponses de l'opérateur dans le voisinage d'un contour.

Les travaux sur la détection de points de contour ont été menés dès les premiers développements du traitement des images. Au cours des années, les chercheurs ont développé des algorithmes de plus en plus robustes. À la naissance du traitement d'images, la détection de points de contour ne servait qu'à quelques applications de bas niveau. Actuellement, des applications de plus haut niveau sont conçues et nécessitent une grande précision. Des détecteurs de points de contour très performants ont ainsi été développés. Les méthodes les plus utilisées pour la détection des points de contour sont les méthodes dérivatives [Rab05]. Elles sont divisées en trois familles :

- les méthodes utilisant le calcul du gradient déterminent les gradients directionnels de l'image en chaque point. Ensuite, par exemple, on calcule la norme du gradient pour déterminer finalement le maximum local de cette norme qui définit le point de contour,
- les méthodes basées sur le calcul du Laplacien de l'image lissée et sur la recherche de ses passages par zéro,
- les méthodes basées sur le calcul des dérivées secondes. Elles utilisent la norme du gradient pour calculer la dérivée seconde dans la direction du gradient. Le passage par zéro de la dérivée seconde permet de détecter les points de contour.

1.5.3.1 Types de contours

D'après Ayache [Aya89], on peut modéliser de manière approximative l'intensité lumineuse émise par un élément de surface dans la direction d'une caméra par une fonction continue des paramètres suivants : l'intensité lumineuse incidente, le vecteur normal à l'élément de surface et la réflectance de l'élément de la scène. La plupart des contours correspondent à une discontinuité de l'un de ces paramètres lorsque la direction d'observation varie. On distingue quatre types principaux de contour comme cela est montré dans la figure 1.3 :

- *arêtes*. Orientation des surfaces,
- *occultations*. Orientation de profondeur,
- *ombres*. Discontinuité de l'intensité de l'éclairage incident,
- *marques*. Discontinuité de la réflectance de la surface.

1.5.3.2 Caractéristiques des contours

Les contours sont les parties de l'image séparant des régions homogènes, *i.e.* des ensembles de pixels voisins d'intensité assez proche qui représentent le fond de l'image et les facettes apparentes des objets. Leur détection est le plus souvent basée sur la détermination des maxima des variations spatiales des intensités. Un contour est caractérisé [Cou94] par :

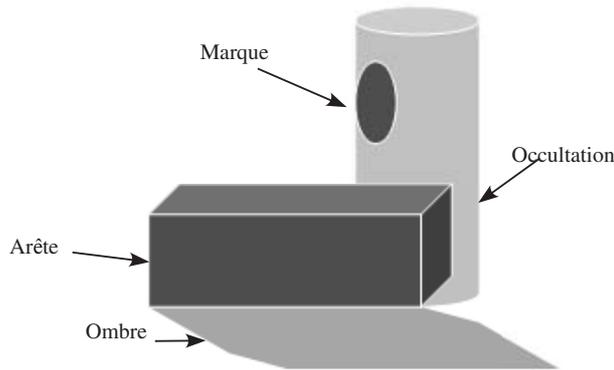


FIG. 1.3 – Différents types de contour

- *Sa direction.* Un contour est orthogonal à la direction du gradient d'intensité induite par la transition d'une zone homogène à une autre,
- *Sa continuité.* Un contour représente une chaîne connexe de pixels de l'image, qui sépare deux zones homogènes,
- *Son épaisseur.* Un contour est d'autant plus proche de la frontière exacte entre deux régions homogènes que son épaisseur est faible. C'est la raison pour laquelle la détection par maxima du gradient est souvent suivie, par exemple, par une opération morphologique d'amincissement.

1.5.4 Segments

Les segments sont plus discriminants que les points de contour pris séparément. Ils permettent de réduire la complexité de l'algorithme de mise en correspondance en réduisant le nombre d'appariements à réaliser, car le nombre de segments est toujours inférieur au nombre de points. D'autre part, on peut prendre en compte de manière explicite la continuité des contours ; en effet, lorsque deux points se correspondent, bien souvent leurs voisins se correspondent également. De plus, les attributs géométriques mesurés sur les segments de contour sont beaucoup plus riches en information que ceux mesurables sur des points isolés. Enfin, la précision sur la mesure de la position et de l'orientation d'un segment est souvent beaucoup plus grande que celle liée à des mesures sur des points isolés.

Les types d'indices visuels que nous avons étudiés permettent de représenter les informations contenues dans une image d'une façon très compacte et très structurée. Cette réduction des données est très utile lorsqu'on applique les algorithmes de mise en correspondance.

1.6 Mise en correspondance

L'être humain n'éprouve aucune difficulté à appairer des images quelles que soient les ambiguïtés dues aux oclusions, à la luminosité, etc. Par contre, cette correspondance

1.6 Mise en correspondance

peut s'avérer relativement difficile à réaliser d'un point de vue purement algorithmique. L'appariement des images est un des plus difficiles problèmes de la vision. Il apparaît dès qu'on utilise plusieurs images représentant la même scène. Étant donné un point de la scène observée, il faut retrouver sa projection dans chacune des deux images. Le problème de la mise en correspondance entre deux images consiste à trouver pour chaque primitive de la première image, une primitive correspondante dans l'autre image. Sans aucune information *a priori*, pour trouver la correspondante d'une primitive, il faut la comparer à toutes les primitives de l'autre image. Cette comparaison s'avère être très délicate, d'une part parce que la seule information disponible dans les images est l'information « intensité » qui est très sensible aux changements de luminosité qui peuvent se produire d'une vue à l'autre. D'autre part, cette variation de luminosité peut influencer l'extraction des caractéristiques images de plus haut niveau (contours, régions, points d'intérêts, segments, etc.), ce qui influence forcément leur appariement. Un second problème réside dans le fait que certains objets ou des parties d'objets de la scène peuvent être non perçus dans une des deux images, ou bien cachés par d'autres objets. Ce phénomène est connu sous le nom d'occlusion et les primitives présentes mais occultées dans la première image ne peuvent pas être mises en correspondance avec celles de la deuxième image. Un autre problème peut provenir de la nature des composants constituant la scène qui peut rendre la tâche de mise en correspondance encore plus difficile. En effet, certains type d'objets peuvent, par exemple, contenir une structure périodique qui provoque de faux appariements très difficiles à détecter.

1.6.1 Contraintes de mise en correspondance

La difficulté de la tâche de mise en correspondance entre les primitives de deux images a poussé plusieurs chercheurs à trouver des contraintes pour mieux rechercher les appariements potentiels. Ces contraintes vont permettre de contrôler fortement le phénomène des correspondances multiples (ou faux appariements). Certaines de ces contraintes portent sur les relations qui existent d'une image à l'autre, comme la contrainte épipolaire. D'autres concernent plutôt les éléments à mettre en correspondance, comme des points, des contours ou des régions. Seuls les critères géométriques possèdent une véritable base théorique, confortée par l'étalonnage. Les contraintes dites figurales ont une base heuristique et reposent principalement sur la recherche de ressemblances entre les éléments que l'on peut trouver dans la paire d'images stéréoscopiques. Cependant, lorsque les prises de vue sont relativement éloignées l'une de l'autre, ces critères figuraux ne sont plus utilisables car la ressemblance apparente entre deux représentations devient trop faible, essentiellement à cause des phénomènes de perspective. D'autres contraintes, comme par exemple les contraintes à critère morphologique, *i.e.* basées sur la recherche de formes prédéfinies, s'appliquent particulièrement bien aux primitives comme les contours ou les régions. Les contraintes d'appariement présentées ci-dessous sont très utilisées, dans la majorité des travaux, pour limiter la combinatoire de la mise en correspondance. Nous allons nous focaliser sur les plus couramment utilisées dans le domaine de la vision binoculaire.

Contrainte épipolaire Cette contrainte découle de la géométrie du système de stéréovision. Elle réduit la complexité du processus de recherche du bon candidat à une recherche sur une ligne épipolaire. Soient les deux points p_g et p_d , projections sur les plans images

gauche et droite d'un point P appartenant à la scène. Les points (P, p_g, p_d) forment le plan épipolaire, qui donne les deux lignes épipolaires l_g et l_d . Toutes les lignes épipolaires passent par un seul point appelé épipole. Les candidats à l'appariement de chaque points de l_g se trouvent sur l_d , et inversement. Cette contrainte épipolaire est une contrainte géométrique qui réduit l'ensemble des correspondants potentiels d'un point d'une image à une droite dans l'autre image.

Contrainte de continuité de la disparité La disparité est la mesure de la différence de position, le long des droites épipolaires rectifiées, des deux projections d'un point 3D P . Si l'on considère respectivement les deux projetés gauche et droite $p_g(u_g, v_g), p_d(u_d, v_d)$ du point P (voir figure 1.4) dans les repères image $(O_g, u, v), (O_d, u, v)$, si l'on se ramène au modèle simplifié du capteur stéréoscopique (voir figure 1.7), alors la disparité est donnée par la mesure $(u_g - u_d)$. La contrainte de continuité de la disparité peut être énoncée comme suit. Soit (p_{g1}, p_{d1}) un appariement de disparité d . Alors, un point p_{g2} , voisin de p_{g1} , doit trouver un correspondant p_{d2} dans la deuxième image avec une disparité proche de d [Som97].

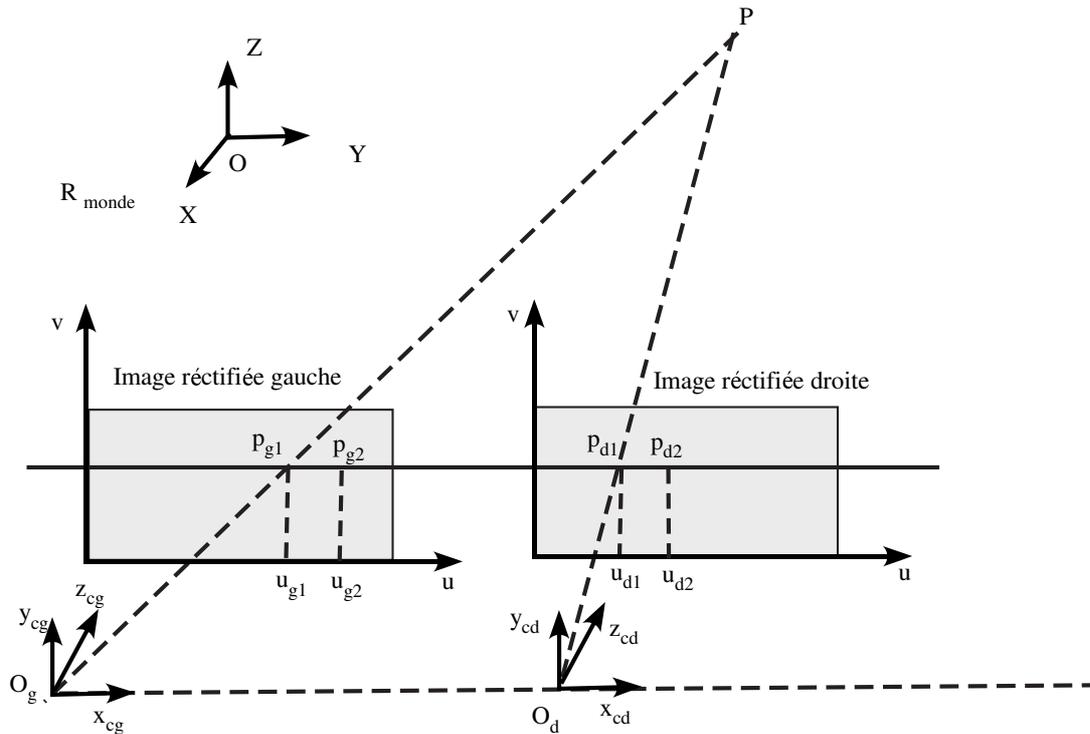


FIG. 1.4 – Contrainte de continuité de la disparité

Contrainte d'ordre La contrainte d'ordre implique que la projection des points d'une scène conserve le même ordre dans les deux projections images. Si un point p_g se trouve à gauche d'un autre point q_g dans l'image de gauche, alors la contrainte d'ordre exige que les points correspondants p_d et q_d dans l'image droite soient dans le même ordre (voir figure 1.5). Cette contrainte a été utilisée pour la première fois par Baker et Binford

1.6 Mise en correspondance

[BB81] pour simplifier la combinatoire de la mise en correspondance. Dans ses travaux, Ayache [AF87] a construit une relation d'ordre Gauche-Droite dans la carte de disparités. Toutefois, cette contrainte n'est pas applicable lorsque la scène est composée, par exemple, d'objets transparents, parce que les projections sont vues dans ce cas dans les deux images dans un ordre inversé (voir figure 1.6).

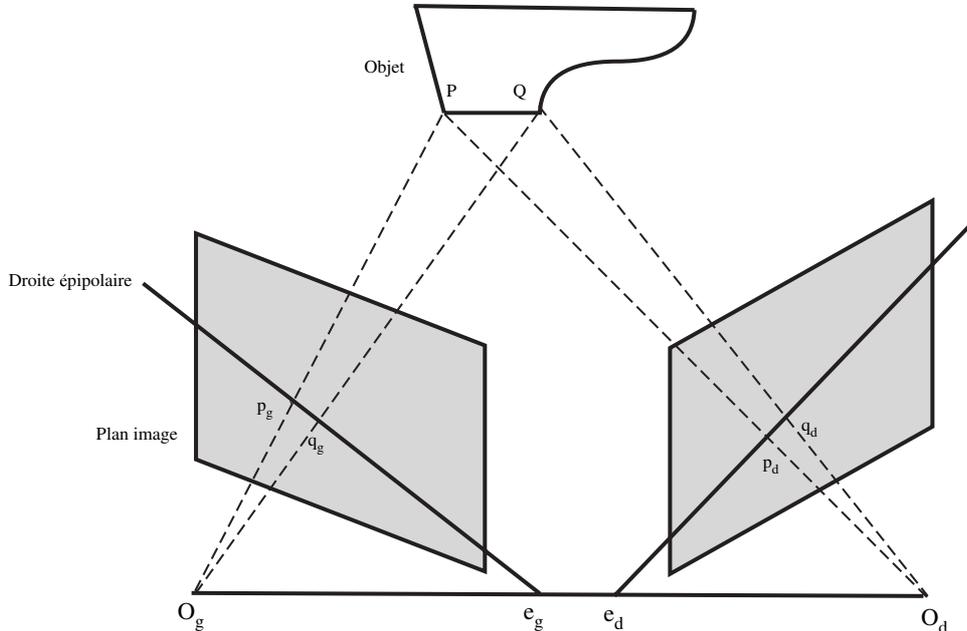


FIG. 1.5 – Contrainte d'ordre

Contrainte d'unicité Par définition, un point de l'image gauche devrait avoir un et un seul correspondant dans l'image droite. La contrainte d'unicité impose qu'une primitive soit mise en correspondance avec une seule primitive de l'autre image. Cela suppose qu'une primitive est la projection d'une unique entité 3D. Cette contrainte n'est pas respectée dans le cas où plusieurs éléments 3D sont alignés et se projettent en une même primitive dans une image.

Contrainte de continuité figurale C'est une contrainte simple et efficace dans le cas où les points à mettre en correspondance se trouvent le long de contours. En se basant sur la conservation de la continuité par la projection perspective, on suppose qu'une courbe continue dans l'espace physique se projette en une courbe continue sur les images. On évite ainsi que les points d'un contour dans l'image gauche soient appariés avec des points de plusieurs contours dans l'image droite. La continuité figurale, utilisée dans quelques travaux [BT80, MF81, BB81, Gri85, KA87], semble être un critère d'appariement pertinent.

Stéréoscopie trinoculaire et critère multiple Au delà de la vision binoculaire, Faugeras [Fau93] a proposé une approche trinoculaire pour réduire les ambiguïtés de l'appa-

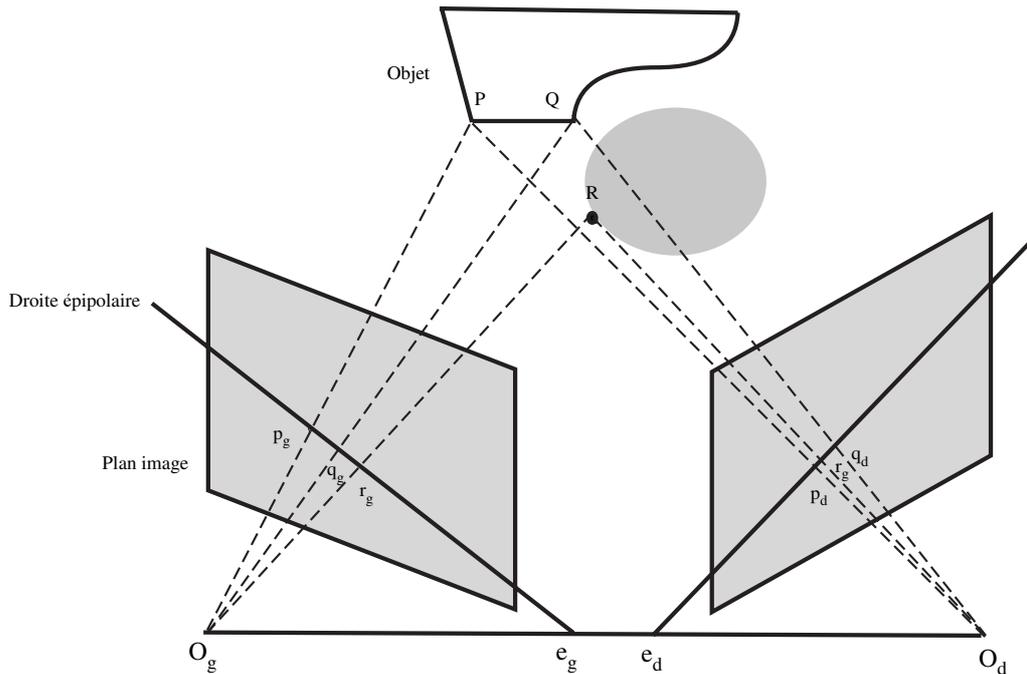


FIG. 1.6 – Mise en défaut de la contrainte d'ordre

riement. La troisième caméra dans la stéréoscopie trinoculaire permet de contraindre la mise en correspondance des éléments des images des deux autres caméras. On teste ainsi la validité de l'appariement en vérifiant la présence cohérente du même élément dans la troisième image. L'utilisation d'une troisième caméra renforce les contraintes géométriques et permet de simplifier les algorithmes de mise en correspondance.

Rectification d'images La rectification d'une paire d'images stéréoscopiques permet de se ramener à une géométrie épipolaire simple où les droites épipolaires sont parallèles à l'axe horizontal passant par les deux centres optiques. Ceci permet de réduire la recherche du point correspondant de l'image droite à une droite horizontale de l'image gauche située à la même ordonnée. Dans cette configuration, les deux épipoles e_g et e_d sont projetés à l'infini. Lorsqu'on procède à un étalonnage des deux caméras, la rectification des images est simple [FTV97]. Par contre, cette rectification devient ardue lorsque les positions des deux caméras ne sont pas connues [Har97]. La figure (1.7) indique le principe d'une rectification des images.

1.6.2 Méthodes de mise en correspondance

De très nombreuses méthodes de mise en correspondance ont été conçues en utilisant différentes stratégies exploitant des primitives et des contraintes. Ces méthodes ont été classées, selon l'ouvrage de Faugeras [Fau93], en trois grandes familles, à savoir les appariements par corrélation, par programmation dynamique et par relaxation. Deux autres méthodes sont également intéressantes. La première est dite appariement par les

1.6 Mise en correspondance

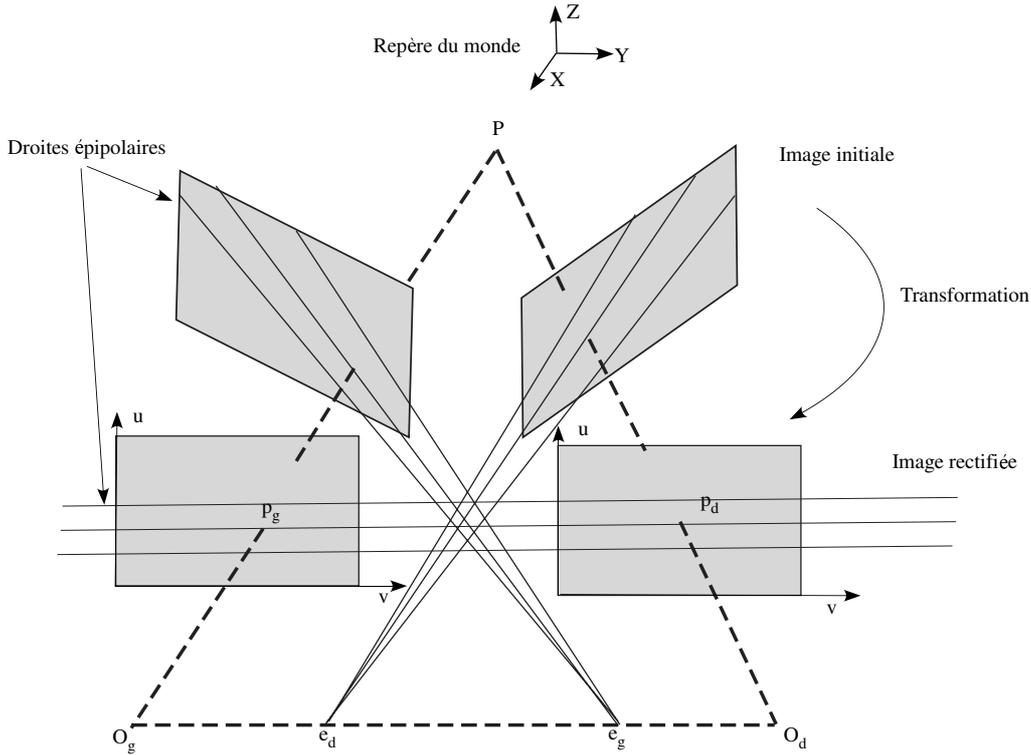


FIG. 1.7 – Rectification d'images

invariants et triangulation de Delaunay. La deuxième, dite mise en correspondance hiérarchique, est développée pour la reconstruction de terrain à partir de deux vues aériennes [HM93]. Nous proposons de donner les principes de ces méthodes dans les paragraphes suivants.

1.6.2.1 Mise en correspondance par corrélation

Une méthode courante utilisée pour l'appariement entre images repose sur l'utilisation de la corrélation [Zha89, Fau93, ZW94]. La corrélation est une caractérisation élémentaire du signal autour d'un point, celui-ci étant caractérisé par l'ensemble des valeurs de niveaux de gris de son voisinage. La corrélation entre images stéréoscopiques consiste, à partir d'une fenêtre placée dans l'une des images, à calculer le degré de corrélation avec une autre fenêtre se déplaçant le long de la ligne épipolaire correspondante dans l'autre image. On obtient ainsi une suite de valeurs de corrélation le long de la ligne épipolaire. Le pixel correspondant au meilleur score sera choisi comme étant à appairer au pixel du centre de la fenêtre fixe dans l'autre image. Une méthode, proposée par Hu [HA94], permet de rendre l'algorithme d'appariement par corrélation robuste aux rotations en utilisant des mesures de corrélation dans plusieurs directions. Mais le calcul dans plusieurs directions s'avère coûteux en temps. En outre, la mesure de corrélation conduit à un certain nombre de faux appariements. Deriche *et al* [DZLF94] ont amélioré les résultats obtenus par corrélation en proposant d'exploiter des caractéristiques du signal comme, par exemple, la direction du gradient, la courbure et la disparité qui est supposée être en-dessous d'un certain seuil.

Certaines méthodes utilisent des points particuliers de l'image, dans le but de diminuer le coût de calcul de la corrélation. À titre d'exemple, Zhang *et al* [ZDFL95] ont appliqué la corrélation à des points d'intérêt. Et, afin d'améliorer le résultat obtenu pour un couple de point mis en correspondance, les points dans des voisinages respectifs de même taille doivent également se correspondre. Spécifiquement, ces points du voisinage doivent avoir une même position relative dans les deux images. Pour les appariements effectués par calcul d'une mesure de ressemblance, de nombreuses mesures ont été proposées (voir, p. ex., [Chr98]).

Mesures de corrélation standard Ces mesures sont basées sur les différences d'intensité lumineuse entre deux sous-fenêtres de taille $(2n + 1) \cdot (2n + 1)$ pixels dans chacune des deux images. Les mesures de corrélation sont considérées comme étant les mesures les plus utiles pour effectuer un appariement de points entre images. Aschwanden [AG92] a montré dans ses travaux que les mesures de corrélation donnent un bon résultat sauf dans le cas d'un changement de luminosité. Quant à Zhang [ZDFL95], il a proposé une méthode d'appariement robuste basée sur des contraintes géométriques locales, au voisinage des points en correspondance, et sur l'estimation de la géométrie épipolaire. Dans [Fau93], une approche a été développée pour améliorer le résultat d'Aschwanden. Cette méthode consiste à faire une correction globale sur une des deux images en cas de changement de luminosité important entre les deux vues.

Mesures de corrélation robustes Ce type de mesures permet de prendre en compte le problème d'occlusion. Zabih [ZW94] a défini une transformation, appelée *rank*, qui permet de transformer chaque pixel en une valeur indiquant le nombre de pixels voisins (parmi les 8) qui ont un niveau de gris inférieur. Ensuite, la mesure de la somme des écarts absolus (*Sum of Absolute Distances* ou SAD) est utilisée pour calculer la ressemblance entre les points des deux images. Il a également décrit une variante, la transformation appelée *census*, qui transforme chaque pixel en une chaîne de 8 bits indiquant quels sont les pixels voisins qui ont un niveau de gris inférieur. Il a utilisé la *distance de Hamming* pour les mesures de distance entre images transformées. Bhat [BN96] a proposé de transformer chaque pixel en une suite de 9 chiffres correspondant à la numérotation des 9 pixels du voisinage (3×3) dans l'ordre croissant de leur intensité. De son côté, Lan [Lan97] a proposé une méthode statistique robuste pour détecter les occultations. Le principe de cette approche est de calculer une mesure à partir de la somme des écarts quadratiques centrés (*Zero-Mean Sum of Squared differences* ou ZSSD), en ne tenant compte que des pixels dans la fenêtre qui ne sont pas occultés. Les pixels occultés sont déterminés en supposant que les pixels formant les fenêtres et en correspondance subissent une translation d'intensité. Lan estime ensuite la distribution de l'erreur. Il peut ainsi déterminer les pixels occultés, puis effectuer une corrélation sur les parties de l'image non occultées. Le critère SAD revient à minimiser la somme des différences des intensités sur l'ensemble de la fenêtre de corrélation. Il est donné par :

$$SAD = \sum_{i=-n}^n \sum_{j=-n}^n | I_2.(x_2 + i, y_2 + j) - I_1.(x_1 + i, y_1 + j) | \quad (1.27)$$

1.6 Mise en correspondance

Le critère ZSSD revient à minimiser la somme des différences de l'écart des intensités à leurs moyennes calculées sur l'ensemble de la fenêtre de corrélation.

$$ZSSD = \sum_{i=-n}^n \sum_{j=-n}^n ((I_2.(x_2 + i, y_2 + j) - \bar{I}_2) - (I_1.(x_1 + i, y_1 + j) - \bar{I}_1))^2 \quad (1.28)$$

où I_1 et I_2 sont les fonctions d'intensité dans chacune des deux images.

Il est utile de souligner que la mise en correspondance par corrélation est une méthode très efficace dans la mesure où la paire d'images présente des régions texturées permettant de faciliter l'appariement ; en revanche, elle échoue souvent lorsque les régions ne sont pas suffisamment texturées.

1.6.2.2 Mise en correspondance par programmation dynamique

Le problème de l'appariement est posé ici sous la forme d'une minimisation d'une fonction à plusieurs variables discrètes, et on trouve des exemples de ces algorithmes dans [BB81, OK85]. C'est une méthode mathématique qui permet de réduire la complexité de l'algorithme d'appariement. On peut également trouver une approche plus simple de la programmation dynamique, en mettant en correspondance les contours extraits d'une paire d'images rectifiées. La recherche de points correspondants s'effectuera donc le long de lignes horizontales dans la paire d'images. Les données traitées dans ce cas à l'aide de la programmation dynamique sont les intervalles entre deux contours.

1.6.2.3 Mise en correspondance par relaxation

La mise en correspondance peut aussi être considérée comme un problème d'étiquetage. Les primitives de l'image droite sont considérées comme des étiquettes, celles de gauche comme les objets à affecter aux étiquettes. L'idée des méthodes de relaxation est d'utiliser les relations de voisinage pour trouver le meilleur étiquetage possible. Pendant la phase d'initialisation, on définit une mesure de ressemblance b_{gd} entre deux primitives homologues g et d des images gauche et droite. Ensuite, dans un processus itératif, la mesure de ressemblance est augmentée si les primitives voisines de g et d peuvent aussi être appariées, sinon elle est abaissée. Après un certain nombre d'itérations, b_{gd} n'évolue plus. En regardant la liste des mesures de ressemblance, on forme la liste des primitives homologues. Davis et Rosenfeld [DR81] décrivent ainsi un processus de relaxation pour effectuer l'appariement entre deux images comme suit :

- un ensemble d'étiquettes (appariements) est sélectionné pour chaque point et on associe une mesure de confiance à chaque étiquette,
- les différentes étiquettes possibles pour chaque point sont comparées avec celles des points voisins. Ensuite, on modifie itérativement les étiquettes, ainsi que leurs mesures de confiance, pour réduire les inconsistances d'étiquetage entre points voisins.

L'avantage majeur de cette méthode réside dans l'évolution itérative de la mesure de ressemblance, puisque l'on corrige et améliore les appariements initiaux grâce aux relations de voisinage.

L'idée générale de ces méthodes est donc de proposer une première liste d'appariements et de déduire des contraintes à partir de ces appariements pour faire d'autres associations. Ensuite, on réitère le processus.

1.6.2.4 Mise en correspondance par invariants et triangulation de Delaunay

Le principe de cette méthode repose sur le fait que chaque point réel projeté dans les plans images possède des caractéristiques invariantes (p. ex., son intensité lumineuse). On peut ainsi caractériser chaque point de chaque image par une liste d'invariants. Par l'intermédiaire de la contrainte épipolaire, on cherche à mettre en correspondance les points qui ont des listes d'invariants similaires. Pour cela, on choisit les points les plus remarquables à mettre en correspondance. Ensuite, on applique une triangulation de Delaunay, en créant des triangles dont les sommets sont les points remarquables mis en correspondance et sélectionnés au préalable. Ainsi, les points qui se trouvent à l'intérieur d'un triangle dans l'image de gauche doivent avoir leurs correspondants dans un triangle dual formé par les points caractéristiques mis en correspondance dans l'image de droite. Ensuite, dans chaque triangle, on cherche de nouveaux points caractéristiques. On les met en correspondance et on recrée de nouveaux triangles [Rab05]. Dans cet algorithme, la triangulation de Delaunay permet de déterminer au final une mise en correspondance dense.

1.6.2.5 Mise en correspondance hiérarchique

Cette méthode a été développée pour la construction de modèles numériques de terrain à partir de deux vues aériennes. À partir d'une paire d'images, on produit une hiérarchie d'images. Cette hiérarchie est constituée de deux pyramides d'images et chacune d'elle contient l'image initiale de taille $N \times M$ (voir figure 1.8), ainsi que des images de tailles $N/2 \times M/2$, $N/4 \times M/4$, $N/8 \times M/8$ et $N/16 \times M/16$, etc. Chaque pyramide a donc plusieurs niveaux de résolution. Dans ces pyramides, on réduit une image par le calcul de la moyenne des pixels dans une fenêtre de 2.2 pixels. Ces quatre pixels sont remplacés par un nouveau pixel dans l'image de résolution inférieure, et ainsi de suite. Partant de l'image de droite à la plus haute résolution, on recherche des points d'intérêts (coins, points lumineux, etc.). On suit ensuite les coordonnées de ce point en allant jusqu'à l'image à la plus basse résolution. Ensuite, il faut trouver le correspondant de ce point dans l'image de gauche à la plus basse résolution. Pour ce faire, l'utilisation de la géométrie épipolaire s'avère utile. On cherche sur la droite épipolaire le point de l'image de gauche le plus en adéquation avec le point de l'image de droite. Et, enfin, il suffit de parcourir en sens inverse la pyramide de droite de la plus basse résolution vers la plus haute résolution. Le point sélectionné dans l'image à la plus basse résolution permet d'initialiser une zone de recherche pour trouver le correspondant recherché du point à cette résolution. Cette procédure est poursuivie jusqu'à atteindre l'image à la plus haute résolution. Des contraintes sont par ailleurs utilisées pour garantir une certaine cohérence des appariements déjà existants.

Hoff et Ahja [HA89] utilisent l'approche hiérarchique dans une méthode de mise en correspondance de points de contour. Une pyramide de trois niveaux est utilisée. L'approche de Cochran et Medioni [CM92] est aussi basée sur une pyramide composée de trois images. Elle intègre la mise en correspondance de deux types de primitives, les segments et les régions. Les auteurs mettent en correspondance les régions par corrélation et obtiennent ainsi une carte de disparité dense en utilisant les contraintes d'ordre et de la continuité de disparité. La coopération entre mises en correspondance de contours et de régions permet ainsi d'aboutir à un résultat correct d'appariement. Néanmoins, la fusion des mises en correspondance des régions et des segments est très coûteuse en temps de

calcul.

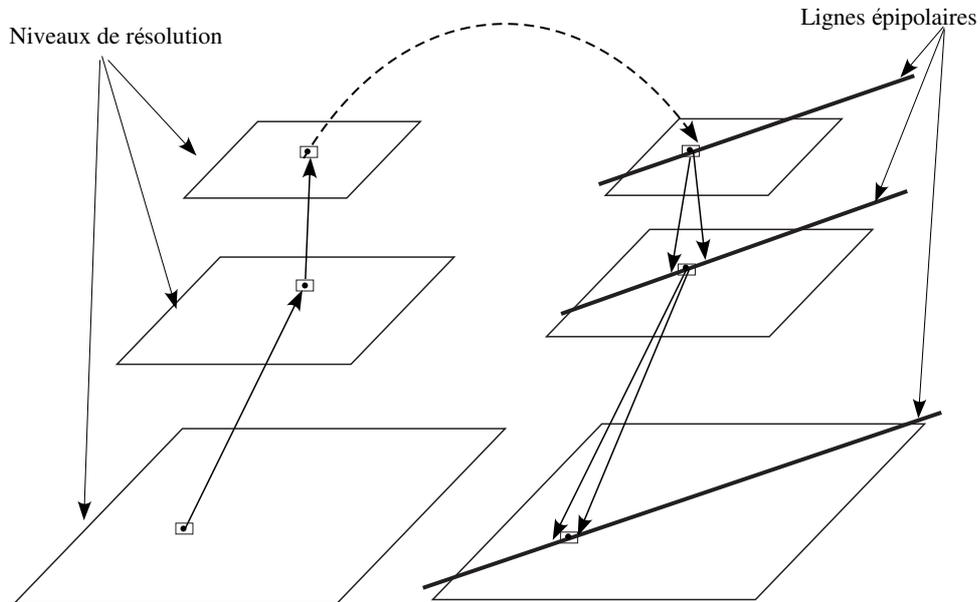


FIG. 1.8 – Principe de l'appariement hiérarchique

Les approches que nous venons de voir, comme par exemple l'appariement par corrélation ou l'appariement hiérarchique, n'utilisent pas de primitives à proprement parlé puisque l'élément utilisé est le pixel. Ces approches mènent à une reconstruction non structurée de la scène : l'ensemble des objets de la scène est inclus dans la carte de profondeur résultant de la reconstruction.

1.6.2.6 Approches basées sur des primitives

La primitive est une représentation fidèle et utile du contenu d'une image. Les primitives ont en général les propriétés suivantes : unicité, répétitivité et sens physique. Dans le contexte de la stéréoscopie, le but des approches d'appariement par primitives est d'obtenir des correspondances correctes, même en présence d'un certain bruit. Dans le cas où l'on dispose de données de plus haut niveau extraites des images, telles que des points de contour ou des segments, le problème de la mise en correspondance est légèrement différent. En effet, les données de type primitives sont moins nombreuses et plus riches en information que des données de type pixel sélectionnées dans les images. Ces primitives sont plus fiables et possèdent en particulier des attributs géométriques qui peuvent être utilisés pour l'appariement.

Nous avons présenté quelques méthodes de mise en correspondance qui ont été largement utilisées dans des travaux connus sur la stéréovision. Ces méthodes exploitent une telle diversité de contraintes qu'elles utilisent, en fait, des types de primitive très divers. Cependant, ces méthodes restent essentiellement combinatoires, tant dans leur mise en œuvre, que dans leur principe. Certaines d'entre elles nécessitent des étapes de traitement très coûteuses en temps de calcul. De plus, la recherche combinatoire des appariements par introduction de plusieurs contraintes s'avère souvent être très complexe. Pour palier ce fait, plusieurs travaux de recherche ont été menés pour réduire cette complexité algorithmique [Som97].

La plupart des approches pour la mise en correspondance que nous avons exposées suivent une même stratégie, reposant sur la comparaison de primitives entre elles, tout en essayant de réduire les ambiguïtés et le nombre de comparaisons en exploitant les contraintes exposées dans la section 1.6.1. La liste des méthodes d'appariement que nous venons de citer est loin d'être exhaustive. Il existe bien d'autres méthodes et d'autres techniques. Néanmoins, il n'existe pas actuellement de méthode assez générale et unique permettant d'effectuer l'appariement quel que soit le type d'images utilisé.

Les différentes approches d'appariement se distinguent essentiellement par les contraintes imposées aux objets présents dans la scène (objets lisses ou texturés, segments, contours d'occultation, etc.), ainsi que par les applications (p.ex., l'inspection dimensionnelle).

L'appariement fournit des couples de points images appariées. La position du point 3D correspondant à cet appariement est déterminé par triangulation.

1.7 Reconstruction 3D

L'étape de reconstruction tridimensionnelle consiste à calculer la géométrie de la scène observée par le système de vision. Cette phase de reconstruction utilise les résultats des algorithmes de mise en correspondance. Ainsi, on doit, pour chaque paire de primitives (points, segments, régions, etc.) mis en correspondance entre les images stéréoscopiques, trouver dans l'espace la position de la primitive 3D correspondante.

En vision stéréoscopique, différentes approches permettant la reconstruction tridimensionnelle, à partir de deux images, ont été suggérées [Eva98, Hor00]. Ces approches sont liées au type d'étalonnage utilisé. Le calcul de la position d'un point 3D à partir de ses deux projections dans les deux images se fait par les méthodes dites de triangulation. Cette dernière consiste à estimer l'intersection dans l'espace des deux droites de vue passant par les deux projections dans les deux images et les centres de projection modélisant les caméras.

Plusieurs types d'étalonnage existent et conduisent à des reconstructions à différents niveaux d'information. Pour obtenir une reconstruction euclidienne, c'est-à-dire une reconstruction qui contient des informations métriques sur la scène, il faut avoir réalisé un étalonnage euclidien ou métrique. Si tous les paramètres d'étalonnage ne peuvent être dé-

1.7 Reconstruction 3D

terminés, la structure 3D reconstruite est moins riche en information. Dans le cas d'un étalonnage partiel, on peut distinguer l'étalonnage affine et l'étalonnage projectif. Ces deux types d'étalonnage conduisent respectivement à une reconstruction affine et projective de la scène. Ces deux reconstructions peuvent malgré tout être utiles. La reconstruction affine préserve le parallélisme et permet de mesurer des rapports de distance entre des points sur des segments de droites parallèles, tandis que la reconstruction projective préserve la colinéarité des points. On peut ainsi mesurer des birapports de points colinéaires.

1.7.1 Le problème de la triangulation

Soit un point P dans l'espace, visible dans les deux images, gauche et droite, et dont les matrices de projection sont M_g et M_d . Soient p_g et p_d les deux projections du point tridimensionnel P . La détermination des deux rayons de projection correspondant aux deux points image ne pose pas de difficulté particulière. Par contre, le problème réside dans la détermination de l'intersection de ces deux droites de vue, gauche et droite, passant par les centres optiques O_g et O_d et par les deux points images p_g et p_d (voir figure 1.9. (a)). En pratique, les deux droites de vue ne se coupent pas dans l'espace et il devient nécessaire de trouver une solution approchée. Une méthode classique consiste à choisir le point minimisant la somme des distances aux rayons de projection non incidents, c'est-à-dire à trouver le point milieu de la perpendiculaire commune aux rayons.

La méthode du point milieu et d'autres approches classiques Les travaux classiques de reconstruction à partir de deux images stéréoscopiques d'une scène s'inscrivent dans un contexte de caméras étalonnées. L'étape de triangulation ou reconstruction se fait en se basant sur la méthode du point milieu qui est couramment utilisée (choisir le point milieu des deux rayons de projection) [Stu97]. Toscani et Faugeras proposent deux

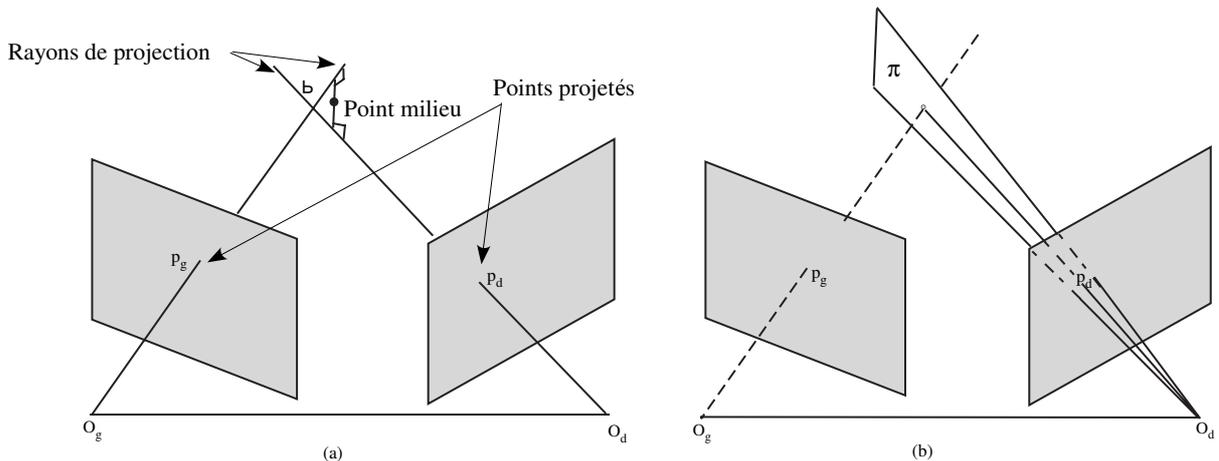


FIG. 1.9 – Deux méthodes classiques pour la triangulation. (a) Point milieu de la perpendiculaire commune aux rayons de projection. (b) Méthode proposée par Toscani et Faugeras.

solutions pour la triangulation. Dans la première, le point reconstruit est celui qui minimise l'erreur de reprojection. Cette approche a été appliquée par plusieurs chercheurs dans

une formulation du type « ajustement de faisceaux », avec l'objectif d'estimer les points 3D. Pour cette estimation, on utilise des méthodes itératives. La deuxième solution, des mêmes auteurs, suppose que le point reconstruit est choisi comme étant l'intersection du rayon de projection de l'un des points image en correspondance avec le plan π contenant le deuxième rayon de projection et perpendiculaire au plan épipolaire du premier rayon (voir figure 1.9. (b)).

Par ailleurs, des méthodes de reconstruction projective à partir de deux vues *non étalonnées* ont été proposées par Faugeras [Fau92], Hartley *et al* [HGC92]. Ces travaux ont pour principe la détermination des matrices de projection dans un repère projectif du monde. La triangulation est obtenue par la résolution d'un système d'équations linéaires, en minimisant une erreur algébrique. Quant à Beardsley [BMZ92], il propose une méthode qui permet d'utiliser un étalonnage approximatif des caméras, et qui conduit à une reconstruction *quasi-euclidienne*. De ce fait, il serait envisageable d'appliquer la méthode du point milieu. Il propose ainsi une solution approximative. Un premier point image est laissé fixe, tandis que la position de son correspondant est corrigée en projetant cette position orthogonalement sur la droite épipolaire du premier point. Ainsi, on obtient deux points qui se correspondent exactement, ce qui permet aux deux rayons de projection de se couper sans erreur. Franz Sturm [Stu97], en collaboration avec Hartley, a décrit dans ses travaux une méthode de reconstruction applicable aux cas projectif, affine et euclidien et qui fournit un résultat optimal, compte tenu d'un critère de minimisation des erreurs de reprojection.

Types de reconstruction La reconstruction repose sur *la fusion* des résultats d'étalonnage et de mise en correspondance. Il existe à l'heure actuelle différentes méthodes pour effectuer la reconstruction 3D. Dans cette partie, nous présentons les types de reconstruction usuels.

Soit P un point de la scène se projetant en p_g et p_d . Reconstruire P consiste à déterminer sa position dans l'espace, c'est-à-dire à déterminer l'intersection des lignes de vues passant par p_g et p_d (triangulation). Soient M_g et M_d les deux matrices de projection des images gauche et droite. Alors :

$$p_g = M_g \cdot P \tag{1.29}$$

$$p_d = M_d \cdot P \tag{1.30}$$

Connaissant M_g et M_d , la résolution des deux équations précédentes (1.29, 1.30) donne la position de P . Ce qui correspond à la reconstruction euclidienne présentée dans la section 1.7.2. Par contre, si M_g et M_d ne sont pas connues, la solution P ne constitue pas une solution unique, et la reconstruction est faite à une homographie ⁴ H près. Nous avons :

$$p_g = M_g \cdot H^{-1} \cdot H \cdot P \tag{1.31}$$

$$p_d = M_d \cdot H^{-1} \cdot H \cdot P \tag{1.32}$$

⁴une homographie est une matrice de transformation linéaire 4×4 de P^3 dans P^3 . P^3 est un plan projectif défini comme l'espace R^4 sans le point $(0\ 0\ 0\ 0)$, avec la relation d'équivalence suivante : $(x_1, x_2, x_3, x_4) \approx (x'_1, x'_2, x'_3, x'_4)$ si et seulement si $\exists \lambda \neq 0$ tel que $(x_1, x_2, x_3, x_4) = \lambda(x'_1, x'_2, x'_3, x'_4)$ [Hor00].

1.7 Reconstruction 3D

Le point H . P est aussi une solution des équations de projection. Cette transformation peut cependant être estimée en fonction des informations sur la scène dont on dispose. On peut ainsi citer deux techniques qui sont la reconstruction affine [Stu97] et la reconstruction projective [DF96]. Dans cette thèse, nous avons opté pour la reconstruction euclidienne car on dispose des paramètres d'étalonnage des deux caméras. Dans ce qui suit, nous abordons ce type de reconstruction.

1.7.2 Reconstruction euclidienne

Dans ce type de reconstruction, on considère une scène avec un nombre connus de points de référence, comme pour l'étalonnage classique à partir de mires d'étalonnage. On peut par exemple utiliser des objets avec des cibles (points marquants) facilement détectables dans les images. Les coordonnées tridimensionnelles de ces cibles sont supposées connues par rapport à un repère de référence avec une grande précision. Si les positions d'au moins six points de référence sont connues, il est possible de déterminer, pour chaque image, les matrices suivantes, par résolution des équations de projection appliquées aux points de référence :

$$\tilde{M}_g = K_g \cdot [R_g \hat{t}_g] \quad (1.33)$$

$$\tilde{M}_d = K_d \cdot [R_d \hat{t}_d] \quad (1.34)$$

où \tilde{M}_g et \tilde{M}_d sont les matrices de projection estimées des images gauche et droite. Dans ce type de reconstruction, nous connaissons le modèle des caméras. Supposons que $p_g(u_g, v_g)$ et $p_d(u_d, v_d)$ sont deux points images mis en correspondance. Le point $P(X, Y, Z)$ est le point 3D que l'on veut calculer (voir figure 1.10). Alors :

$$\begin{pmatrix} su_g \\ s\nu_g \\ s \end{pmatrix} = \begin{pmatrix} m_{g11} & m_{g12} & m_{g13} & m_{g14} \\ m_{g21} & m_{g22} & m_{g23} & m_{g24} \\ m_{g31} & m_{g32} & m_{g33} & m_{g34} \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \tilde{M}_g \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1.35)$$

$$\begin{pmatrix} su_d \\ s\nu_d \\ s \end{pmatrix} = \begin{pmatrix} m_{d11} & m_{d12} & m_{d13} & m_{d14} \\ m_{d21} & m_{d22} & m_{d23} & m_{d24} \\ m_{d31} & m_{d32} & m_{d33} & m_{d34} \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \tilde{M}_d \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1.36)$$

Ce système fournit quatre équations pour déterminer les trois inconnues X , Y et Z :

$$u_g = \frac{m_{g11}X + m_{g12}Y + m_{g13}Z + m_{g14}}{m_{g31}X + m_{g32}Y + m_{g33}Z + m_{g34}} ; \quad \nu_g = \frac{m_{g21}X + m_{g22}Y + m_{g23}Z + m_{g24}}{m_{g31}X + m_{g32}Y + m_{g33}Z + m_{g34}} \quad (1.37)$$

$$u_d = \frac{m_{d11}X + m_{d12}Y + m_{d13}Z + m_{d14}}{m_{d31}X + m_{d32}Y + m_{d33}Z + m_{d34}} ; \quad \nu_d = \frac{m_{d21}X + m_{d22}Y + m_{d23}Z + m_{d24}}{m_{d31}X + m_{d32}Y + m_{d33}Z + m_{d34}} \quad (1.38)$$

La résolution de ce système est effectuée en utilisant la méthode des moindres carrés.

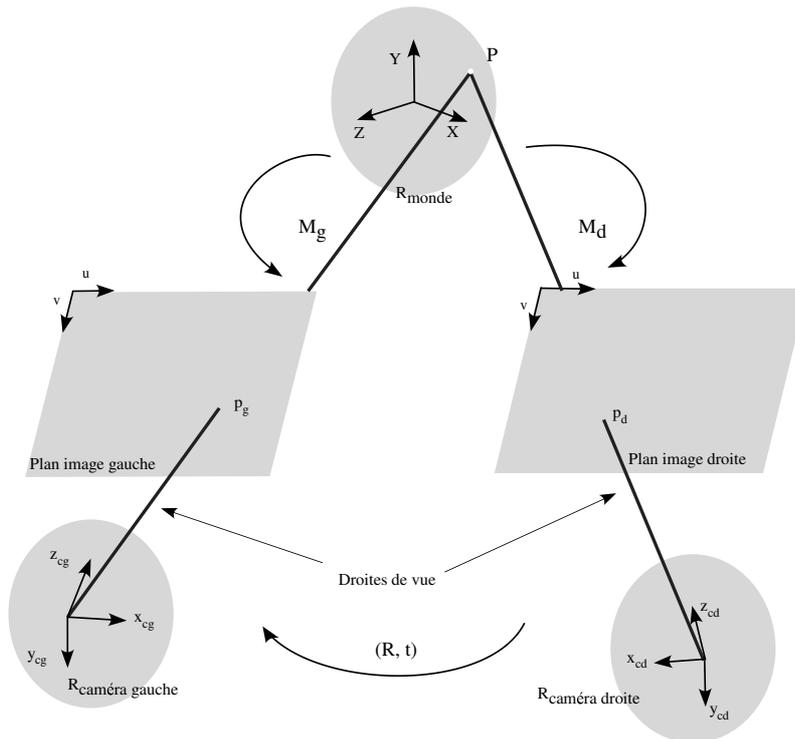


FIG. 1.10 – Reconstruction euclidienne

1.8 Conclusion

La fusion d'indices visuels observés simultanément à partir de plusieurs points de vues est une faculté largement utilisée par les systèmes de perception biologiques. En ce qui concerne les systèmes de vision artificielle passive, la stéréoscopie reste la méthode la plus intuitive et la plus probante pour reconstruire et évaluer une scène en 3 dimensions.

Dans ce chapitre, nous avons présenté un état de l'art sur les concepts fondamentaux de la stéréovision. Nous avons commencé par présenter la modélisation, ainsi que les différentes techniques d'étalonnage d'un capteur stéréoscopique. Puis, nous avons étudié la géométrie épipolaire, ainsi que différents indices visuels permettant de représenter des entités extraites des images et qui peuvent ensuite servir à l'appariement entre données de deux images stéréoscopiques. Nous avons enfin abordé les principales techniques d'appariement développées dans la littérature. À la fin de ce chapitre, nous avons vu que, selon le type d'étalonnage effectué, plusieurs types de reconstruction 3D sont possibles.

La perception du relief peut s'opérer à partir de plusieurs indices. Plusieurs approches cherchent ainsi à exploiter la prise d'images stéréoscopiques. Ces approches sont généralement classées selon [VG91] en deux catégories. Premièrement, les approches dites implicites, qui ne nécessitent pas une connaissance *a priori* exacte des paramètres d'étalonnage du capteur stéréoscopique, ne visent pas à obtenir un résultat *exact* pour la reconstruction d'une scène et génèrent simplement une représentation basée sur la géométrie projective. Deuxièmement, les approches explicites, quant à elles, mettent en correspondance des pri-

1.8 Conclusion

mitives (points, segments, régions, etc.) dans un couple d'images stéréoscopiques, et, grâce à une connaissance des conditions de prise de vue (paramètres d'étalonnage du capteur), permettent d'aboutir à une information 3D métrique pour les primitives par triangulation. Nous nous sommes ainsi essentiellement intéressés à ce type d'approches, étant donné que la précision de la reconstruction 3D joue un rôle primordial dans le domaine du contrôle qualité et de la mesure dimensionnelle des pièces industrielles.

À partir de notre étude, on peut distinguer deux grandes familles d'algorithmes d'appariement utiles dans les approches explicites pour la reconstruction :

- Les algorithmes de bas niveau qui consistent à essayer, pour chaque petite zone de l'image, de trouver dans l'autre image celle qui lui correspond. Cette approche est naturellement bien adaptée au cas de surfaces relativement texturées ; à l'exclusion des textures régulières qui risquent de multiplier les ambiguïtés d'appariement,
- Les algorithmes de haut niveau, où l'on extrait de chaque image les objets qui composent l'image grâce, par exemple, à des techniques d'extraction de contours, puis où l'on met en correspondance ces objets. Ces algorithmes nécessitent des méthodes d'extraction de contours performantes. Ils sont très bien adaptés aux images contenant des formes artificielles comme, par exemple, des pièces manufacturées.

Dans le cas de notre étude, nous avons ainsi utilisé l'approche stéréoscopique pour aboutir à une reconstruction 3D de pièces industrielles à partir d'images stéréoscopiques, en vue d'une inspection dimensionnelle et d'un contrôle qualité. Pour ce faire, nous allons décrire dans le chapitre 3 de ce mémoire une nouvelle méthode d'appariement, conduisant ensuite à une reconstruction 3D de contours efficace.

Étant donné notre objectif et le cadre de notre application, dans laquelle on cherche à effectuer une analyse dimensionnelle d'un objet par comparaison avec son modèle conceptuel, nous avons fait les choix suivants :

- **Étalonnage du capteur.** Parmi les différentes techniques proposées dans la littérature, nous avons retenu la méthode de Tsai. Cette dernière est disponible dans notre équipe et permet d'obtenir une bonne estimation des paramètres intrinsèques et extrinsèques des caméras. Cette technique est décrite dans l'annexe A de ce mémoire.
- **Mise en correspondance.** Les différentes méthodes de mise en correspondance présentées ne sont pas adaptées à notre type d'images. Nos images d'un objet manufacturé sont spécifiques. Cependant, nous nous sommes inspirés des méthodes présentées pour développer une nouvelle approche. Cette dernière est basée sur une étape cruciale dite de prétraitement, qui conduit à la classification des contours extraits des images selon des critères géométriques simples. De plus, notre technique utilise le modèle CAO disponible de la pièce à reconstruire (voir chapitre 3), ce qui lui confère une robustesse vis-à-vis des faux appariements.
- **Indices visuels.** Les indices visuels choisis sont les contours. Ces derniers sont en effet des éléments descriptifs essentiels de l'objet à évaluer.
- **Géométrie épipolaire.** Notre algorithme d'appariement est basé sur l'estimation de la géométrie épipolaire. Parmi les méthodes étudiées, nous avons retenu l'algorithme normalisé dit des huit points pour estimer la matrice fondamentale décrivant cette géométrie.
- **Reconstruction 3D.** Dans notre environnement de travail contrôlé, les caméras

utilisées pour l'acquisition des images sont étalonnées avec une grande précision, ce qui nous a conduit à mettre en place une technique de reconstruction euclidienne.

Chapitre 2

Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques

2.1 Introduction

Pour des applications industrielles de la vision par ordinateur, comme par exemple le contrôle industriel et la métrologie, il est essentiel de développer des outils permettant une reconstruction 3D précise du contenu de l'image. L'objet reconstruit peut alors être comparé avec le modèle CAO de l'objet à inspecter.

Dans cette optique, et dans une première contribution de cette thèse, nous proposons une approche pour l'appariement et la reconstruction utilisant la représentation tridimensionnelle des contours. La stéréovision est une technique très adaptée à l'extraction d'information 3D en utilisant deux images 2D représentant le même objet. Dans la majorité des processus de vision, la notion d'appariement est fondamentale, particulièrement pour une reconstruction tridimensionnelle par stéréovision. L'identification du même objet projeté sur les deux rétines des yeux de l'être humain se fait instantanément, mais la même opération d'identification semble être très difficile à réaliser par une machine. Nous avons constaté qu'il n'existe pas actuellement de méthode d'appariement suffisamment générale, capable d'être appliquée à une paire d'images stéréoscopiques indépendamment de leurs contenus. Cela nous a amené à ne pas tenter d'améliorer les performances des approches déjà existantes, mais à nous tourner vers une nouvelle approche permettant d'apparier des images de type industrielles. Xie [Xie95] a élaboré, au cours de ses travaux, une méthode d'appariement coopérative pour mettre en correspondance des points de contours. Nous nous sommes ainsi inspirés de son approche en utilisant deux types de représentation d'indices visuels, à savoir les points de contour et les chaînes de points de contour. Cependant, son approche est plus particulièrement adaptée aux images texturées. Ce qui n'est pas le cas de nos images d'objets de type industriel qui sont lisses.

Dans ce chapitre, nous allons décrire une approche originale, adaptée à l'appariement de contours, qui se fonde sur le calcul de la géométrie épipolaire dans la paire d'images stéréoscopiques, ainsi que sur la comparaison de données réelles et conceptuelles. Ce calcul se fait de façon bidirectionnelle, en calculant la contrainte épipolaire deux fois, de gauche vers la droite et vice-versa, pour éliminer les faux appariements. De plus, la comparaison des images réelles et conceptuelles permet de sélectionner les contours à apparier. La technique proposée se base sur une étape cruciale de classification de contours en segments de droites et segments curvilignes, selon un critère géométrique. Cette classification permet de restreindre la recherche du correspondant d'un type de segments aux segments du même groupe géométrique. Cela revient à introduire une contrainte géométrique à notre méthode d'appariement.

Dans ce chapitre, nous commençons par présenter les primitives utilisées pour l'appariement et les raisons pour lesquelles nous avons choisi ces types de primitive. Puis, nous décrivons les différentes étapes de notre méthode d'appariement. La comparaison des données réelles et synthétiques est ensuite introduite. Plusieurs expériences ont été réalisées afin de valider les différentes étapes de l'approche d'appariement proposée.

2.2 Représentation des primitives contours utilisées

On peut évaluer l'efficacité d'une primitive par son pouvoir descriptif, sa capacité à lever les ambiguïtés et par son invariance par rapport au point de vue, ce que l'on appelle

aussi sa stabilité. On peut ainsi affecter aux primitives les caractéristiques suivantes :

- *Dimensionnalité*. Les primitives de type point sont insensibles aux déformations perspectives, mais possèdent de faibles caractéristiques discriminatoires, contrairement aux primitives de type contour qui sont cependant sensibles aux distorsions perspectives.
- *Taille*. Les primitives de grandes tailles possèdent davantage de caractéristiques distinctives, mais sont souvent sensibles aux variations perspectives.
- *Contraste*. Les primitives qui ne se situent pas sur des variations photométriques notables sont plus difficiles à distinguer entre elles.
- *Contenu sémantique*. Plus le contenu sémantique d'une primitive est fort, plus il sera aisé de différencier entre eux les différents candidats à l'appariement.
- *Densité d'occurrence*. Les primitives ayant une faible occurrence dans la paire d'images permettront un appariement plus aisé, du fait que les appariements possibles sont moins nombreux ; mais cela conduit également à un nombre plus restreint de points 3D reconstruits.

Les primitives que nous allons traiter dans cette étape sont les chaînes de points de contour, classés selon un critère géométrique simple (segments de droites et segments curvilignes). Dans le cas de l'analyse dimensionnelle d'une pièce, ces primitives sont celles qui fournissent la meilleure information relative à la géométrie de l'objet.

Chaînage des contours Avant de réaliser la mise en correspondance, nous procédons à une étape de chaînage des points de contour, pour construire des listes primaires de points de contour. Certains auteurs, comme [GHR92], pensent que cette étape de chaînage présente un inconvénient, qui peut affecter le déroulement de l'appariement, notamment, en présence de faux contours (liés aux effets d'illumination). Ce problème ne se posera pas dans notre approche. En effet, des images de synthèse correspondant aux images réelles de l'objet à évaluer sont disponibles. Ces images ne contiennent que des contours correspondant à des contours réels de l'objet. Elles vont être utilisées pour éliminer les effets indésirables (bruit, effet d'illumination, etc.) des images réelles. De cette manière, seuls les contours réels sont conservés dans les images réelles. Pour toutes ces raisons, nous avons maintenu cette étape de chaînage de points de contour avant d'appliquer notre approche d'appariement.

Pour ce faire, nous commençons par réaliser une approximation des contours en éléments géométriques simples, à savoir des segments de droites et des segments curvilignes. Ces derniers sont les éléments descriptifs principaux du type d'objets à analyser (pièces manufacturées). On peut ainsi aboutir à une bonne approximation des contours des images de pièces manufacturées, si l'on arrive à construire précisément ces éléments géométriques. Contrairement aux points de contour, les segments ont des spécificités géométriques qui facilitent leur discrimination. Néanmoins, il se trouve que cette discrimination n'est pas simple si la scène contient plusieurs objets identiques. Cependant, ce problème n'est pas traité dans notre étude. Dans notre application, nous n'utilisons qu'un seul objet pour chaque traitement.

Pourquoi des segments de contour ? Nous avons choisi de construire une approximation des contours en segments de droite et en segments curvilignes. Ces segments sont

2.3 Détection, construction et classification des chaînes de points de contour

les primitives qui vont être mis en correspondance. Ce choix est justifié par le fait que l'utilisation de segments en stéréovision a montré que cela pouvait être une alternative intéressante à l'appariement de points de contour individuels. Cela s'explique notamment par le fait que les attributs géométriques mesurés sur les segments de contour sont plus riches et donc plus discriminants que ceux mesurés sur des points [Aya89].

Ces segments ont aussi l'avantage d'être moins sensibles aux erreurs de position que les points traités séparément. Ainsi, ils conduisent à des contraintes globales plus fortes (unicité, ordre, continuité), et chaque primitive est porteuse d'information pertinente. La direction et la longueur des segments sont deux caractéristiques effectivement pertinentes dans notre algorithme. Il est clair que, dans cette approche utilisant des segments, on cherche à simplifier la description de départ des images en la décomposant en primitives simples.

Appariement entre contours de la même famille La solution, apportée par notre approche à la mise en correspondance, consiste à faire des appariements entre primitives de contour. La primitive utilisée est de type segment de droite ou segment curviligne. Ainsi, chaque primitive de la première image sera comparée avec les primitives de la même classe de la deuxième image. Il en sera de même pour chaque primitive de la deuxième image. Nous cherchons ainsi à apparier séparément ces types de primitive en introduisant des contraintes géométriques et des critères de compatibilité.

Une primitive est composée d'un ensemble de points de contour connectés qui donnent un *a priori* sur sa longueur. On suppose que le réglage des paramètres d'illumination, préalablement aux acquisitions des images, ainsi que des paramètres de détection de contours sont fiables.

Les étapes qui vont nous permettre de réaliser la segmentation/classification des contours en formes géométriques simples sont les suivantes : détection et chaînage des points de contour pour construire une représentation symbolique de l'image, segmentation puis classification de ces chaînes en segments de droite et segments de forme curviligne.

Nous commençons ainsi par décrire l'étape de prétraitement des images pour extraire et classer les primitives que l'on va chercher à mettre en correspondance. Ensuite, nous aborderons l'approche d'appariement proposée.

2.3 Détection, construction et classification des chaînes de points de contour

L'approximation des contours à l'aide de primitives est une étape très importante pour l'obtention d'une description symbolique de l'image. Cette opération s'effectue en deux étapes : construction de chaînes de points de contour (listes primaires), suivie d'une classification de ces dernières selon un critère géométrique après segmentation.

La détection et l'amincissement des contours sont respectivement les deux premières étapes de notre algorithme. Daul [Dau94] avait présenté et comparé dans son travail de thèse différentes techniques de détection et d'amincissement de contours. Son objectif était

Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques

de retenir celles qui sont le mieux adaptées à notre type d'applications. En effet, de nombreux travaux ont été consacrés au développement d'opérateurs de détection de points de contour. On peut distinguer deux grandes familles de détecteurs : les opérateurs linéaires et non-linéaires. Les opérateurs Laplaciens et gradients sont les opérateurs les plus utilisés. Les opérateurs gradients permettent de mettre en évidence les fortes variations de niveaux de gris correspondant aux contours pouvant apparaître dans une image. Ils sont habituellement constitués de deux filtres directionnels. Ces directions sont orthogonales et généralement données par les deux axes du repère dans lequel est donnée l'image. Chacun de ces filtres permet de déterminer une composante de la variation du niveau de gris selon la direction dans laquelle il est orienté. Ainsi, un vecteur est obtenu pour chaque point de l'image, dont le module et l'argument représentent respectivement l'amplitude de la variation du niveau de gris et la direction dans laquelle cette variation est observée. Les opérateurs de type gradient fournissent une information sur la direction de la variation maximale du niveau de gris en tout point de l'image. Ces opérateurs sont directionnels. On peut, lorsqu'on connaît cette direction pour un point de contour, déterminer la tangente du contour en ce point (la tangente en un point du contour et la direction du gradient formant un angle droit). Les opérateurs gradients sont, en outre, moins sensibles au bruit que d'autres opérateurs (p. ex., les opérateurs Laplaciens).

Nous avons ainsi choisi le détecteur de points de contour de type gradient introduit par Canny et qui représente un bon compromis entre robustesse, performances et facilité d'implémentation.

On peut ensuite calculer le module et la direction du gradient des niveaux de gris en chaque point de l'image. Par application de ce détecteur aux images de la figure 2.1, on obtient les images résultats pour les directions horizontale, voir figure 2.2, et verticale, voir figure 2.3, du gradient, pour la direction du gradient, voir figure 2.4, et l'amplitude du gradient, voir figure 2.5.

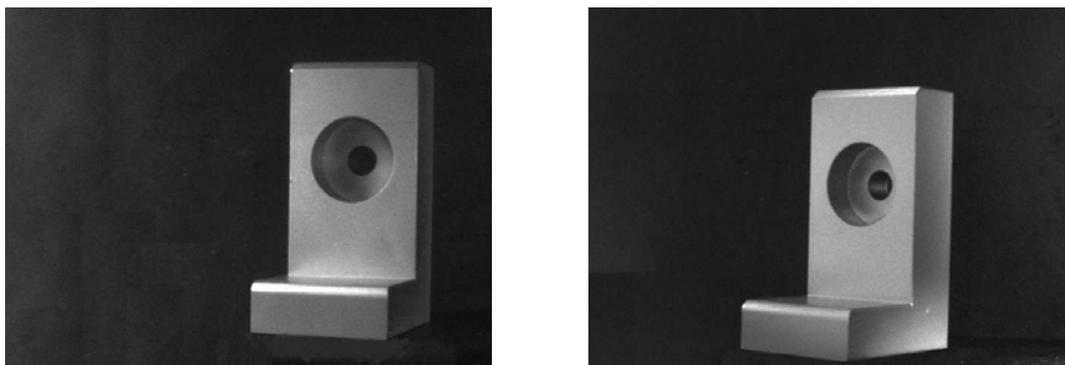


FIG. 2.1 – Images stéréoscopiques réelles

La détection de contours se fait aisément en utilisant un seuillage approprié dans l'image du module. Cette méthode de détection de points de contour nécessite une étape supplémentaire d'amincissement pour obtenir les coordonnées des points les plus représentatifs des contours détectés. Ainsi, les contours, résultant de la convolution d'une image

2.3 Détection, construction et classification des chaînes de points de contour

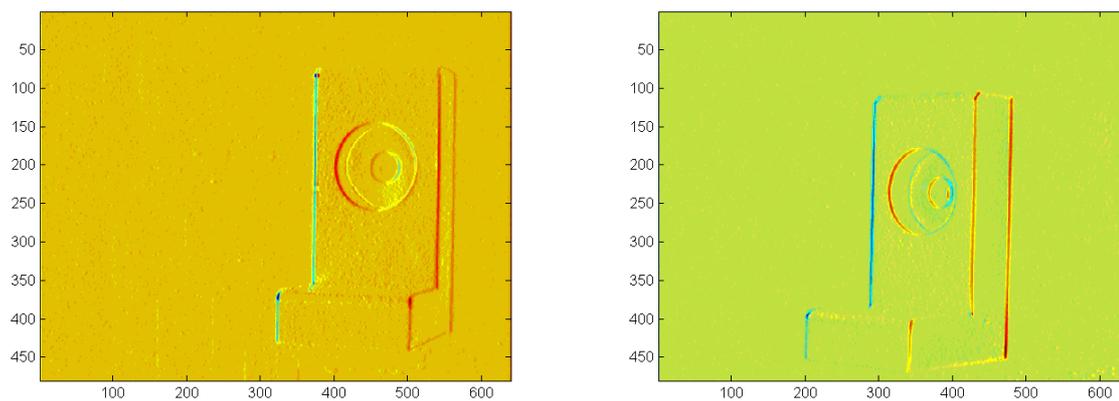


FIG. 2.2 – Images du gradient horizontal

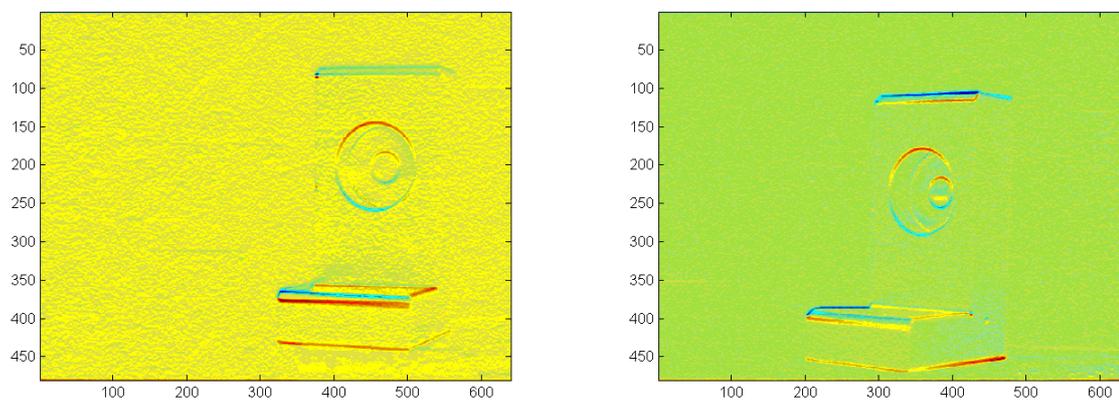


FIG. 2.3 – Images du gradient vertical

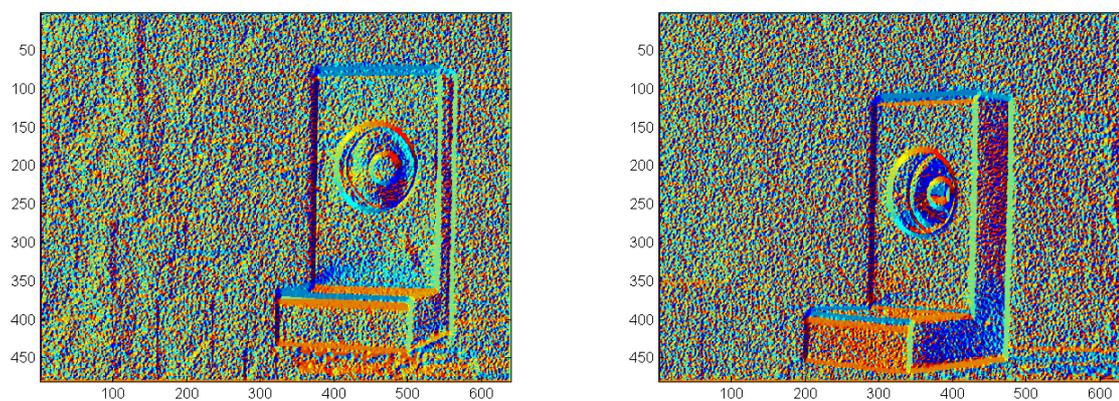


FIG. 2.4 – Images de la direction du gradient

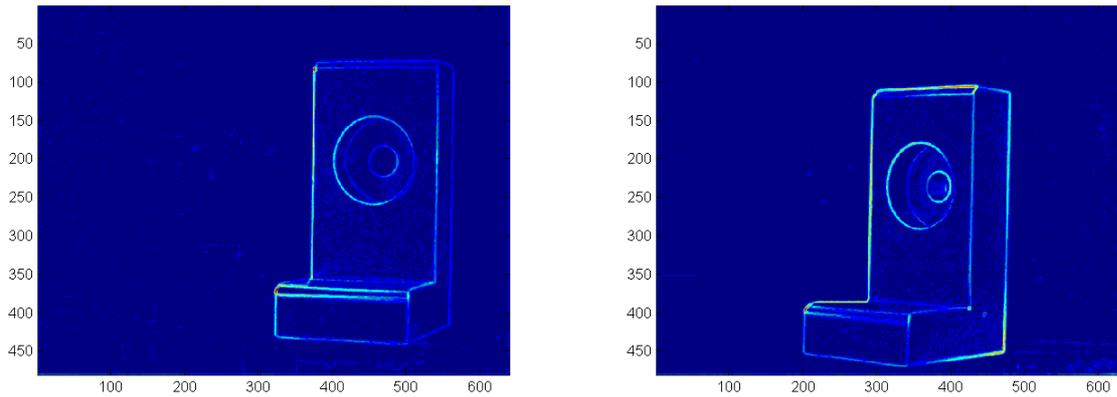


FIG. 2.5 – Images de la norme du gradient

avec un opérateur de détection de points de contour, présentent dans certaines régions une épaisseur qui peut s'étendre sur plusieurs pixels. Il est donc nécessaire pour ces régions d'appliquer une procédure d'amincissement de contour. Plusieurs méthodes d'amincissement de contours sont proposées dans la littérature. Nous nous sommes limités aux approches d'amincissement de contours détectés par des opérateurs gradients et plus particulièrement par le détecteur de Canny. Parmi ces méthodes d'amincissement, on peut citer la *technique de squelettisation*, qui est une méthode simple. Elle présente cependant des performances peu élevées, parce qu'elle est basée sur un algorithme itératif. Une autre approche d'amincissement de contours développée par Onken [Onk89], porte sur l'étude des différents arrangements possibles des pixels situés dans une fenêtre de taille 3×3 centrée sur le point image à analyser. Il existe 512 arrangements possibles de pixels pour ce voisinage. Ce chiffre peut être ramené à 56 combinaisons différentes pour des raisons de symétrie. Onken vérifie, pour toutes ces combinaisons, si la suppression du point central de la fenêtre 3×3 change la connectivité des huit voisins. Un point est supprimé lorsque la connectivité de ses voisins reste inchangée.

Korn [Kor88] a proposé une méthode d'amincissement de contours qui consiste à détecter des maxima dans l'image des amplitudes du gradient. Pour rechercher ces maxima, il suffit de vérifier si un pixel constitue un maximum local dans la direction du gradient lui correspondant. Pour cela, on considère un voisinage 3×3 autour du pixel examiné et on associe un secteur angulaire à chaque voisin (voir figure 2.6). Chacun de ces secteurs est délimité par deux segments de droites formant les angles α et α' avec l'horizontale. Si la direction du gradient du pixel examiné appartient à l'intervalle $[\alpha, \alpha']$ on sélectionne le secteur formé par α et α' , et, par suite, les voisins lui correspondant, pour sélectionner les points servant à la recherche du maximum. Le pixel $p(i, j)$, dont l'amplitude du gradient vaut $A_{i,j}$, est retenu comme point de contour s'il vérifie une des inéquations (2.1), en

2.3 Détection, construction et classification des chaînes de points de contour

fonction du secteur S_1, S_2, S_3 ou S_4 auquel appartiennent ses voisins :

$$\begin{aligned}
 A_{i,j} &\geq A_{i,j-1} \text{ et } A_{i,j} \geq A_{i,j+1} \text{ dans } S_1 \\
 A_{i,j} &\geq A_{i-1,j-1} \text{ et } A_{i,j} \geq A_{i+1,j+1} \text{ dans } S_2 \\
 A_{i,j} &\geq A_{i-1,j} \text{ et } A_{i,j} \geq A_{i+1,j} \text{ dans } S_3 \\
 A_{i,j} &\geq A_{i-1,j+1} \text{ et } A_{i,j} \geq A_{i+1,j-1} \text{ dans } S_4
 \end{aligned}
 \tag{2.1}$$

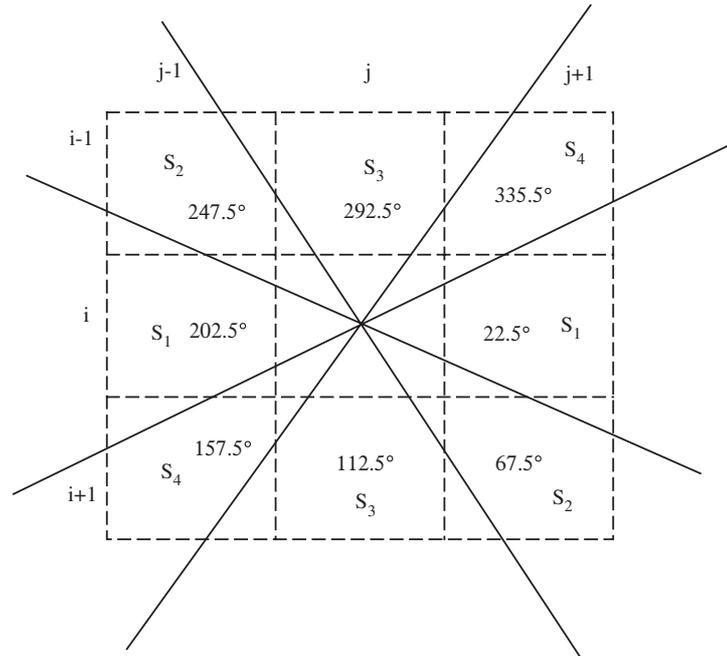


FIG. 2.6 – Détection des maxima dans l’image des amplitudes du gradient (d’après Daul [Dau94])

La méthode de Korn est plus précise que celle d’Onken, pour la sélection des points à retenir pour définir les contours [Dau94]. La méthode de Korn présente cependant l’inconvénient de ne pas amincir les plateaux pouvant apparaître dans l’image des amplitudes du gradient. De tels plateaux correspondent, par exemple, à des variations linéaires du niveau de gris (*i.e.*, une transition de niveaux de gris dont le profil suit une rampe) dans l’image d’origine. Daul a évité l’utilisation d’un algorithme supplémentaire pour traiter le cas des plateaux en modifiant légèrement les tests formulés par Korn. Appelons p_1 et p_2 les voisins situés dans la direction du gradient du pixel examiné et A_{p1} et A_{p2} leurs amplitudes du gradient respectives. Les tests de Korn sont réécrits pour le point p à examiner en changeant les signes des inégalités de l’expression (2.1).

$$\begin{aligned}
 A_{p1} &\leq A_{p2} \text{ et } A_{p2} < A_p \\
 \text{ou} \\
 A_{p1} &< A_p \text{ et } A_{p2} \leq A_p
 \end{aligned}
 \tag{2.2}$$

On obtient, à l’aide de ces tests, un contour d’un pixel d’épaisseur. La figure 2.7 donne un exemple d’images de contours amincis.

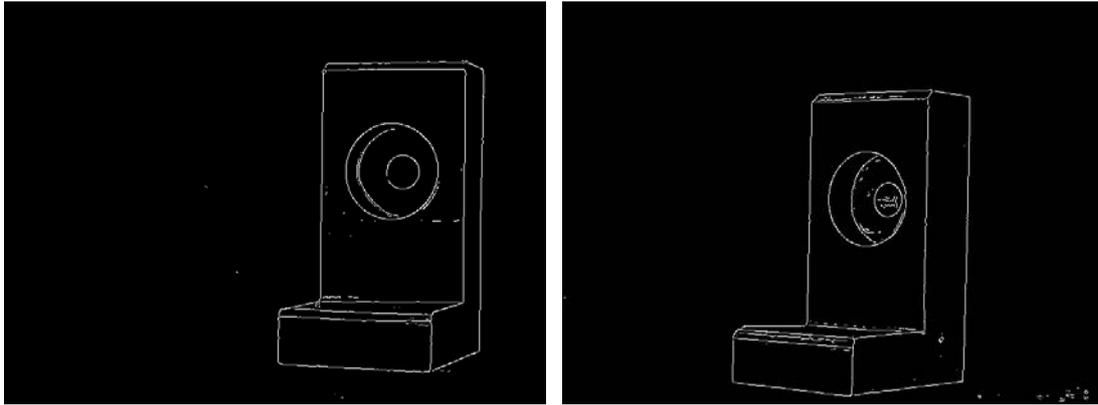


FIG. 2.7 – Images de contours après amincissement

Une fois les contours amincis, l'étape suivante consiste à obtenir des listes connexes des points de contour (figure 2.8). Le principe consiste à regrouper les voisins connexes dans une même liste, sans aucune discontinuité. Ces listes forment des chaînes de points de contour appelées listes primaires de points de contour.

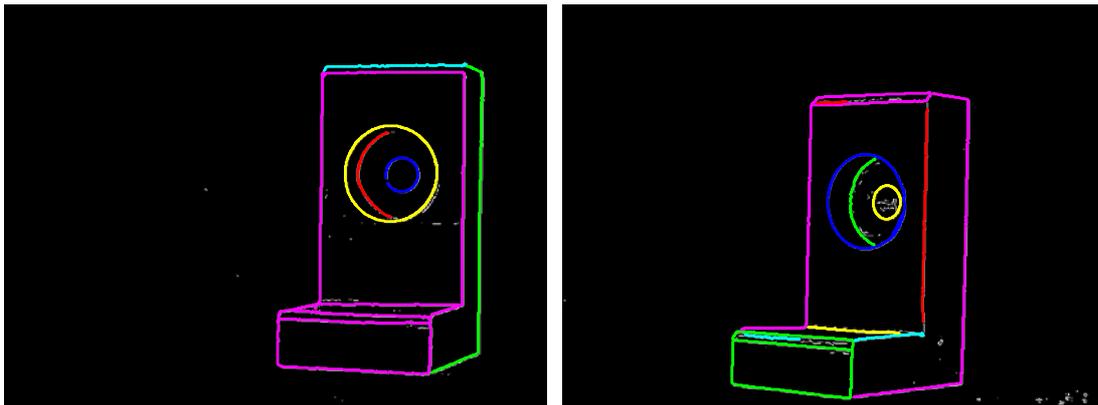


FIG. 2.8 – Construction des chaînes de points de contour ou listes primaires de points de contour. Les points de contour d'une même couleur dans un voisinage proche appartiennent à la même liste primaire de points de contour

Classification des listes primaires de points de contours en primitives

Nous considérons qu'à ce stade les traitements de bas niveau sont effectués. Nous disposons d'images prétraitées et segmentées dans lesquelles apparaissent les contours de l'objet. Une fois les listes primaires de points de contour construites, une nouvelle tâche est la segmentation et la classification de ces listes en primitives. Il s'agit de subdiviser les contours en une ou plusieurs parties appartenant chacune à une forme géométrique simple. Notre algorithme utilise la direction du gradient propre à chaque point de contour

2.3 Détection, construction et classification des chaînes de points de contour

et repose sur le tracé d'histogrammes locaux des directions du gradient pour rechercher les segments de droite [Dau94]. La présence d'un éventuel pic dans un histogramme permet de détecter un segment de droite orienté selon l'angle correspondant à ce pic. Cette méthode permet de classer ainsi tous les points appartenant à des segments de droite, en traitant séparément toutes les listes primaires de points de contour construites au préalable (voir figure 2.9). Tous les autres ensembles connexes de points sont classés comme appartenant à des formes quelconques (formes curvilignes) et leurs histogrammes de directions du gradient ont une allure plus ou moins étalée (voir figure 2.10).

Nous avons donc classé les contours extraits des images réelles avec cette méthode utilisant des histogrammes locaux. Les résultats obtenus sont suffisants pour préparer l'étape de mise en correspondance. Un exemple de résultats obtenus avec notre méthode de classification des listes de points de contour est donné dans la figure 2.11. Dans cet exemple, la classification des listes primaires de points de contour conduit à la détection de primitives verticales, horizontales ou de forme curviligne. Mais cette méthode permet aussi de détecter des segments de droite orientés dans des directions autres que verticales et horizontales.

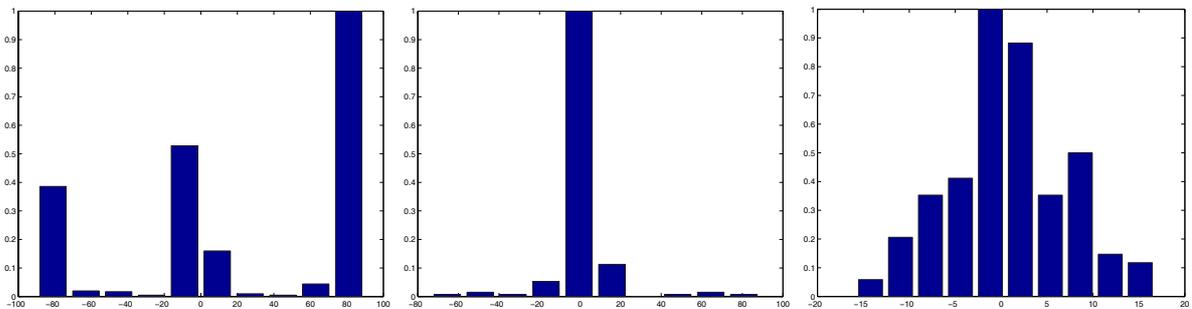


FIG. 2.9 – Exemples d'histogrammes de la direction du gradient de points appartenant aux listes primaires de points de contour de la figure 2.8. La détection des pics dans ces histogrammes permet de définir les segments de droite

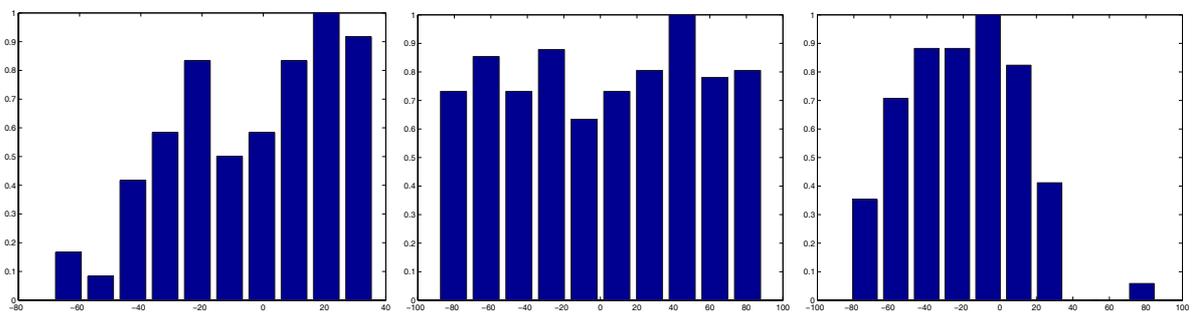


FIG. 2.10 – Exemples d'histogrammes de la direction du gradient de points appartenant aux listes primaires de points de contour de la figure 2.8. L'absence de pics dans ces histogrammes permet de définir les primitives de forme curviligne

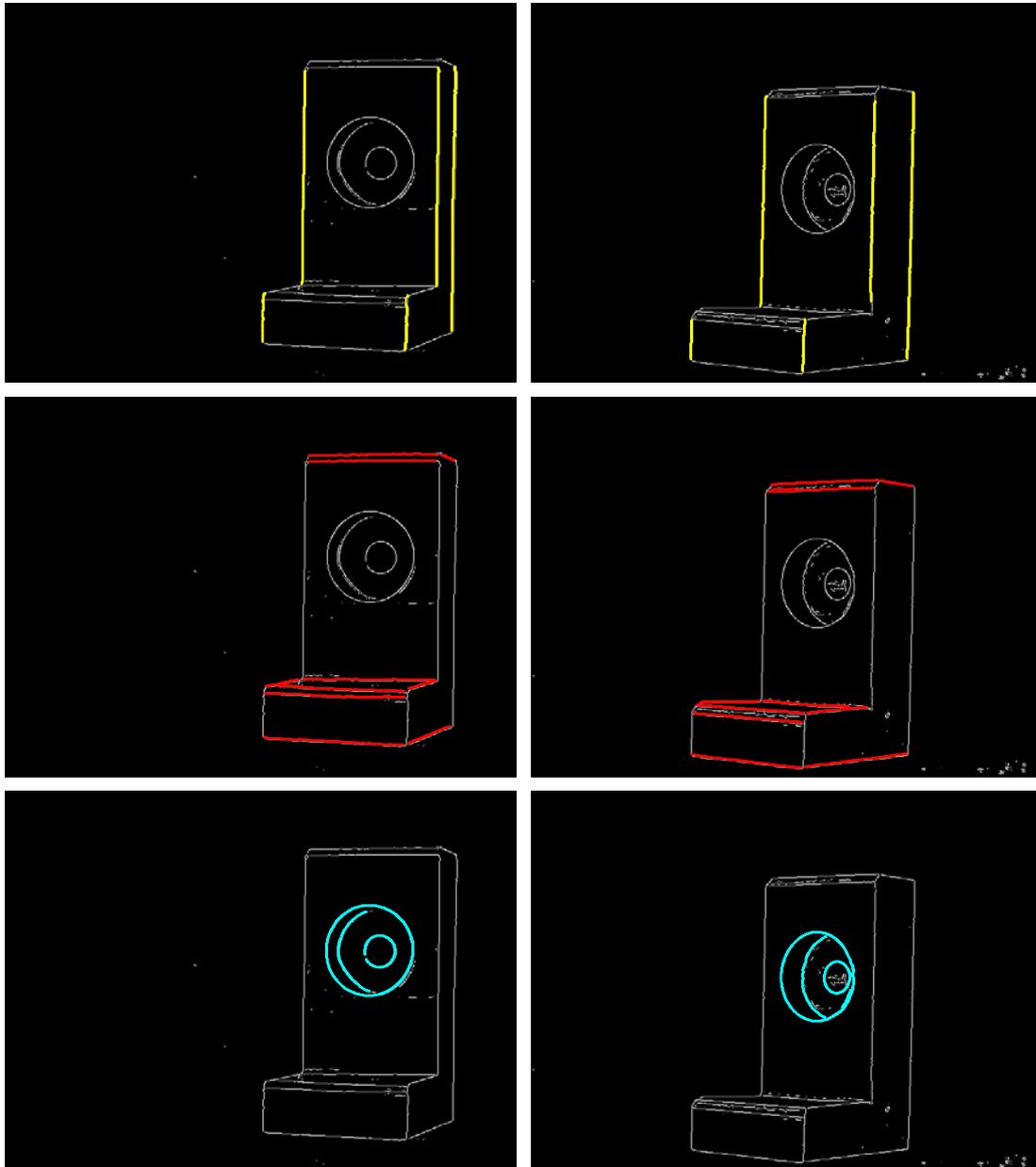


FIG. 2.11 – Classification des listes primaires de points de contour en primitives verticales (images du haut), horizontales (images du milieu) ou de forme curviligne (images du bas). Dans chaque images les primitives classées apparaissent en couleur

2.4 Procédure d'appariement basée sur le calcul bidirectionnel de la contrainte épipolaire

Le capteur est composé de deux caméras, fixées solidairement sur un banc expérimental, et conduit à deux points de vue distincts O_g et O_d . Chaque caméra est représentée par un modèle sténopé et est associée à une matrice de projection, M_g pour la caméra gauche et M_d pour la caméra droite. Le point tridimensionnel P est représenté par sa projection $p_d = M_d.P$ dans l'image droite et sa projection $p_g = M_g.P$ dans l'image gauche. La géométrie épipolaire décrit ensuite la géométrie projective entre les deux images, indépendamment de la structure de la scène. Cette géométrie dépend seulement des paramètres internes propres à chaque caméra ainsi que de leur position relative l'une par rapport à l'autre. La matrice fondamentale F représente la seule contrainte géométrique disponible pour résoudre le problème de la mise en correspondance. Plus précisément, le correspondant p_g de p_d est contraint d'appartenir à une droite épipolaire dans l'image droite. Algébriquement, l'équation suivante doit être satisfaite :

$$p_g^t . F . p_d = 0 \quad (2.3)$$

avec

$$F = K_g^{-1} t . R . K_d^{-1} \quad (2.4)$$

(R, t) est la transformation rigide (rotation et translation) reliant le système de coordonnées de la caméra droite à celui de la caméra gauche [HZ00]. De même, l'équation équivalente :

$$p_d . F^t . p_g = 0 \quad (2.5)$$

où F^t est la matrice transposée de F , permet de tracer les droites épipolaires à partir de l'image droite vers l'image gauche. Pour estimer les matrices fondamentales, nous avons d'abord utilisé l'algorithme classique des huit points [LH87] appliqué sur nos images stéréoscopiques (figure 2.12). Cet algorithme a été perfectionné par Hartley [Har97] et nous l'avons appliqué pour effectuer l'estimation de la matrice fondamentale. Un exemple de droites épipolaires obtenues par l'application de cet algorithme dit des 8 points normalisé, sur les images de la figure 2.1, est présenté sur la figure 2.13.

2.4.1 Principe d'appariement

Notre approche pour l'appariement des contours, basée sur la géométrie épipolaire, repose également sur le calcul d'un critère de similarité. L'étape précédente de classification des points de contour en segments de droite et en listes de points de forme quelconque, a deux avantages majeurs. Le premier réside dans le fait que le contenu de chaque image est décrit par un nombre bien moins important de caractéristiques que le nombre de points de contour disponibles. Deuxièmement, la recherche des correspondances de primitives est sensiblement réduite. En conséquence, pour chaque primitive d'une classe donnée dans l'image gauche, la recherche de son correspondant s'effectue dans la même classe de l'image droite. Dans notre méthode, deux représentations d'éléments de contour sont prises en compte : les points de contour et les listes de points de contour, en accord avec [Xie95]. Nous détaillons dans le paragraphe suivant l'algorithme d'appariement.

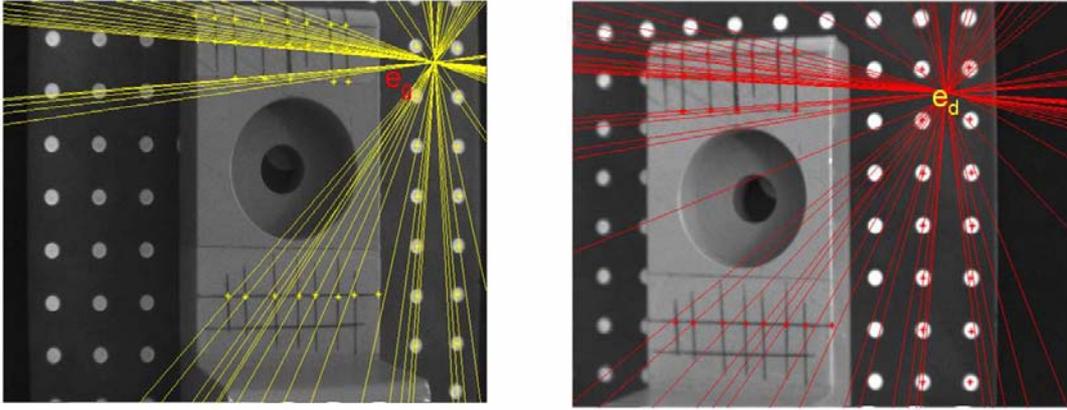


FIG. 2.12 – Exemple de droites épipolaires estimées à partir de l'algorithme des huit points ; e_g et e_d sont les épipoles gauche et droit respectivement

Soient les ensembles de points de contour $p_{g,i}$, appartenant à un contour $C_{g,i}$, pour i allant de 1 à m où m est le nombre de primitives dans l'image gauche et $p_{d,j}$, appartenant aux contours $C_{d,j}$, pour j allant de 1 à n , n étant le nombre de primitives dans l'image droite. Nous recherchons les correspondances entre ces deux groupes de points. L'utilisation de la matrice fondamentale F permet de guider les correspondances entre les deux groupes de points de contour. Mais, habituellement, les droites épipolaires conduisent à plusieurs intersections avec un contour donné. Ainsi, pour tous les points $p_{g,i}$ appartenant à $C_{g,i}$, nous calculons les droites épipolaires correspondantes et nous déterminons le nombre de leurs intersections avec toutes les primitives droites. L'ensemble des intersections $Int_{g,i}$, pour i allant de 1 à m , fournit les candidats initiaux pour un appariement avec les points $p_{g,i}$. Les contours gauches $C_{g,i}$, auxquels appartiennent les points candidats, sont aussi de bons candidats pour l'appariement avec les contours droits $C_{d,j}$. Mais l'appariement obtenu de cette manière entre deux points ne traduit qu'une correspondance possible entre les deux contours auxquels ils appartiennent. Ceci n'exclut donc pas les faux appariements de contours. Il est ainsi nécessaire d'éliminer les faux appariements dans l'ensemble des appariements potentiels entre primitives.

2.4.2 Élimination des faux appariements

L'étape exposée ci-dessus permet de construire une matrice d'appariement contenant le nombre d'intersections entre droites épipolaires et primitives d'une image, calculées pour tous les points constituant les primitives de l'autre image. Les points des primitives de l'image gauche servent à calculer, sur l'image droite, les droites épipolaires qui forment un certain nombre d'intersections avec des primitives de l'image droite. Ces nombres d'intersections, obtenus à partir de l'ensemble des primitives de l'image gauche, permettent de construire la matrice d'appariement gauche/droite $App(g, d)$. Elle est de dimension $m \times n$. Dans cette matrice, comme on peut l'observer dans la figure 2.14, certains contours conduisent à un nombre plus élevé d'intersections, qui traduit une probabilité plus grande

2.4 Procédure d'appariement basée sur le calcul bidirectionnel de la contrainte épipolaire



FIG. 2.13 – Ensembles de droites épipolaires, sur les images de gauche, calculées à partir des primitives de forme curviligne des images de droite

Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques

d'appariement entre les primitives concernées par ce nombre.

Pour minimiser les éventuels faux appariements, la même procédure est appliquée une deuxième fois, en partant des points constituant les primitives droites. Cela signifie que nous effectuons le calcul des intersections avec les droites épipolaires, en prenant l'image droite comme point de départ de la procédure (voir figure 2.17). La procédure mène ainsi à la construction d'une deuxième matrice $App(d, g)$ qui résume à nouveau tous les appariements possibles (voir figure 2.14). Cette matrice est de dimension $n \times m$ [FKH05b].

\nearrow	c_{d1}	c_{d2}	c_{d3}	c_{d4}	c_{d5}	c_{d6}	c_{d7}
c_{g1}	342	116	0	0	0	106	97
c_{g2}	75	110	0	0	0	35	35
c_{g3}	0	0	0	141	90	0	26
c_{g4}	0	0	123	0	0	5	5
c_{g5}	0	0	0	94	135	9	26
c_{g6}	300	107	121	0	42	172	210
c_{g7}	304	100	12	76	115	211	218

$App(g, d)$

\nearrow	c_{g1}	c_{g2}	c_{g3}	c_{g4}	c_{g5}	c_{g6}	c_{g7}
c_{d1}	326	114	0	0	0	105	93
c_{d2}	73	105	0	0	0	38	33
c_{d3}	0	0	0	133	0	6	1
c_{d4}	0	0	135	0	87	0	15
c_{d5}	0	0	92	0	129	11	22
c_{d6}	292	99	0	118	35	173	209
c_{d7}	297	106	119	119	116	213	219

$App(g, d)$

FIG. 2.14 – Exemple de matrices d'appariement

Dans l'exemple de la figure 2.14, 342 intersections sont obtenues entre les droites épipolaires, calculées à partir des points correspondant au contour C_{g1} de l'image gauche, et le contour C_{d1} de l'image droite dans $App(g, d)$. Cela permet d'effectuer immédiatement l'appariement entre C_{g1} et C_{d1} . On élimine ensuite les colonnes et les lignes correspondant aux contours appariés. La recherche du maximum suivant (219) dans la nouvelle matrice issue de $App(d, g)$ permet d'apparier C_{d7} avec C_{g7} , puis C_{g6} avec C_{d6} (le maximum est de 173 dans $App(d, g)$). On peut observer que, vers la fin, la procédure détecte un maximum de 141 pour $App(g, d)$ et de 135 pour $App(d, g)$, ce qui conduit à un appariement incorrect de C_{g3} avec C_{d4} . Notre étape préalable de classification des primitives en segments de même type permet cependant de minimiser et éventuellement d'éliminer ces risques de faux appariements entre primitives de type différent. Le calcul de $App(g, d)$ et $App(d, g)$ est ainsi fait pour chaque classe de primitives.

L'appariement peut enfin être finalisé en calculant un critère de similarité impliquant la recherche des maxima qui se correspondent dans les deux matrices, comme indiqué dans l'algorithme ci-dessous. Cet algorithme récapitule aussi l'ensemble des étapes de la méthode, y compris la reconstruction 3D à partir des points correctement appariés [FKH05a].

- 1 Prétraitement des données images.
 - a. Estimation de la géométrie épipolaire (matrices F et F^t).
 - b. Segmentation des images et construction des listes de points de contour.
 - c. Classification des listes de points de contour en primitives verticales, primitives horizontales et primitives de formes curvilignes. Voir figure 2.11.
- 2 Détermination des primitives géométriques observées dans le couple d'images stéréoscopiques en utilisant une connaissance *a priori* (voir section 2.5). Voir figure 2.26.
- 3 Détermination des appariements potentiels.

2.4 Procédure d'appariement basée sur le calcul bidirectionnel de la contrainte épipolaire

- d. Pour chaque primitive gauche $C_{g,i}$, calcul du nombre d'intersections $Int_{g,i}$ des droites épipolaires $d_{g,i}$ avec les primitives droites $C_{d,j}$. Obtention de la matrice $App(g, d)$. Voir figures 2.14, 2.18, 2.19, 2.20.
 - e. Pour chaque primitive droite $C_{d,j}$, calcul du nombre d'intersections $Int_{d,j}$ des droites épipolaires $l_{d,j}$ avec les primitives gauches $C_{g,i}$ dans l'autre image. Obtention de la matrice $App(d, g)$. Voir figures , 2.13, 2.14, 2.15, 2.16.
- 4 Finalisation/validation des appariements.
- f. Évaluation du critère de similarité entre les deux matrices, impliquant la recherche des maxima qui se correspondent dans les deux matrices.
 - g. Élimination des contours appariés à la suite de l'étape (f) dans les deux matrices.
 - h. Itération des étapes (d) à (g) jusqu'à ce que toutes les primitives soient traitées. Voir figure 2.21.
- 5 Reconstruction 3D de l'objet. Voir figure 2.23.

2.4.3 Appariement point à point

Une fois les primitives appartenant à chaque groupe de forme (classe) appariées, nous avons effectué l'appariement point à point pour chaque couple de primitives mis en correspondance en nous basant sur des approches classiques. En effet, le correspondant de chaque point appartenant à une primitive gauche est bien à l'intersection de la droite épipolaire, associée à ce point, avec la primitive droite correspondante. Nous avons tenu compte d'une deuxième contrainte pour l'appariement point à point des points appartenant à une forme curviligne. En effet, il peut exister deux intersections pour la droite épipolaire associée à un point appartenant à un élément curviligne. La contrainte utilisée est la direction du gradient. Deux points sont appariés s'ils ont approximativement la même direction du gradient (voir figure 2.22).

Une fois la liste des points de l'image gauche correctement appariés obtenue, nous pouvons calculer l'information 3D par l'intermédiaire d'une reconstruction euclidienne (voir section 1.7.2). En effet, les paramètres d'étalonnage sont déterminés par l'utilisation d'un logiciel disponible dans notre laboratoire et mettant en œuvre la méthode d'étalonnage de Tsai [Tsa86]. Un exemple de résultat de reconstruction est donné dans la figure 2.23.

La technique d'appariement suggérée suppose que, pour chaque groupe géométrique, le nombre de primitives est le même dans les deux images, ce qui est très restrictif. Pour traiter le cas général, pour lequel le nombre de primitives dans les deux images peut être différent, une approche classique consiste à introduire des contraintes supplémentaires, telles que la contrainte de similarité, d'exclusivité ou de proximité [HS89]. Dans le cadre de notre étude, comme nous considérons des applications industrielles telles que le contrôle de qualité et qu'un modèle CAO de l'objet est disponible (voir figure 2.25), il semble naturel d'utiliser cette connaissance *a priori* fournie par le modèle CAO comme contrainte. Ce point est développé dans la section suivante.

Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques

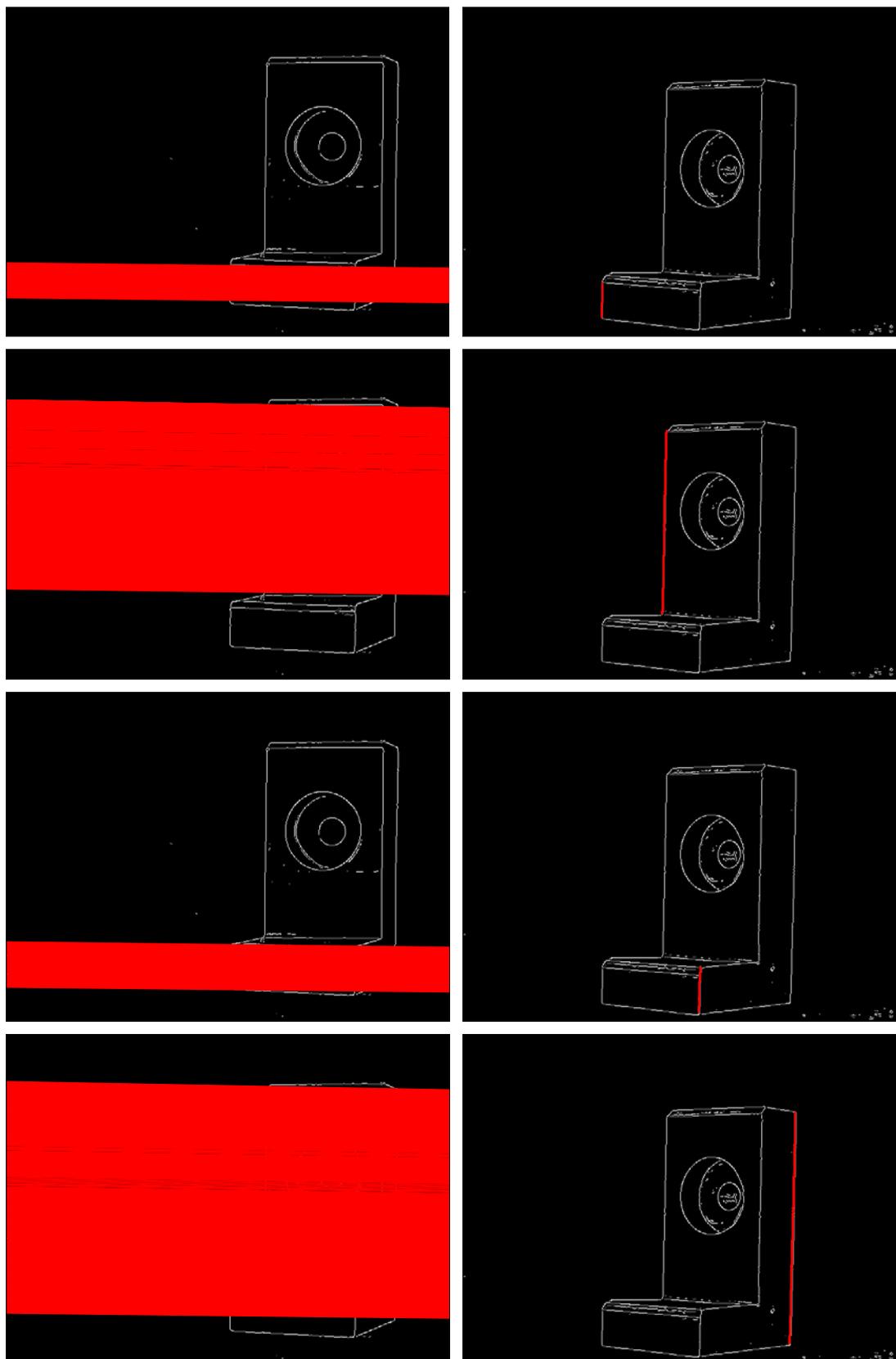


FIG. 2.15 – Ensembles de droites épipolaires, sur les images de gauche, calculées à partir des segments verticaux des images de droite

2.4 Procédure d'appariement basée sur le calcul bidirectionnel de la contrainte épipolaire

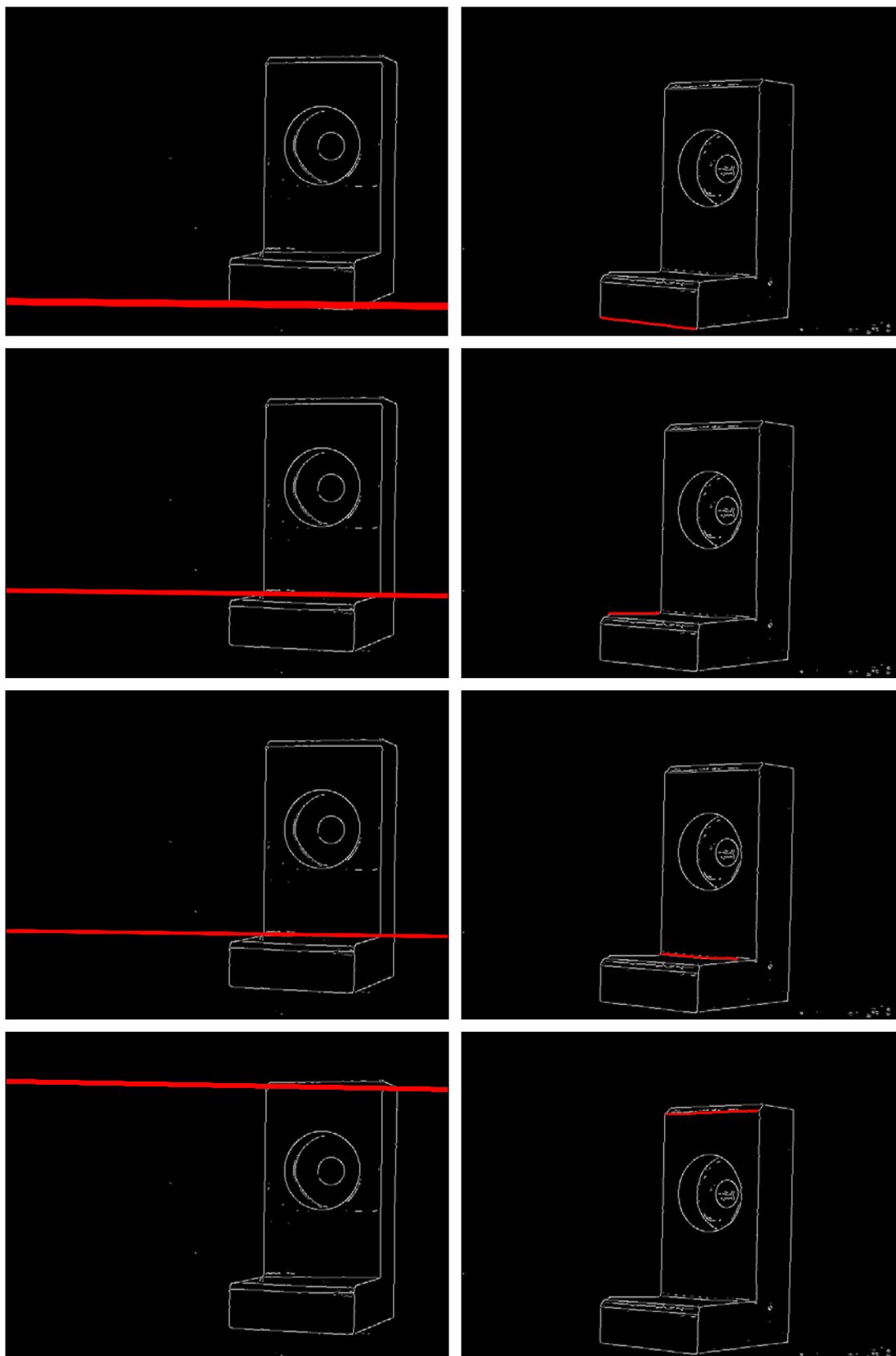


FIG. 2.16 – Ensembles de droites épipolaires, sur les images de gauche, calculées à partir des segments horizontaux des images de droite

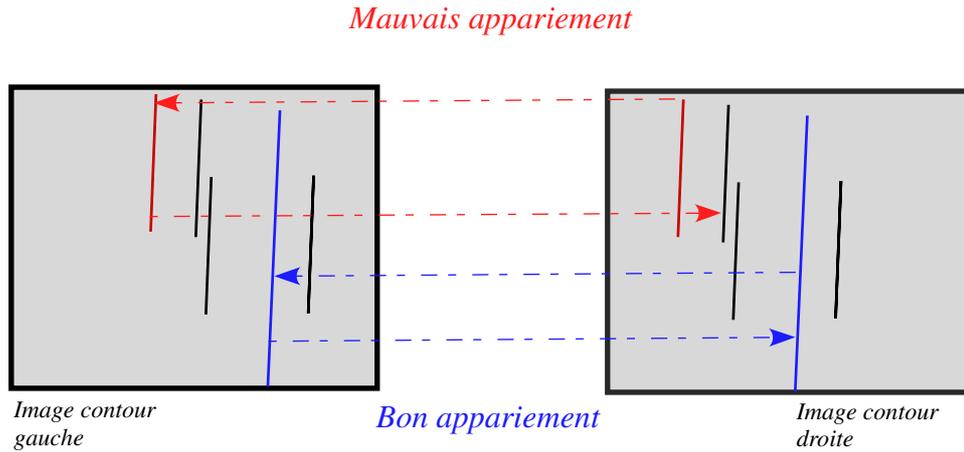


FIG. 2.17 – Principe de l'appariement bidirectionnel

2.5 Comparaison de données images réelles et conceptuelles

Un modèle complet 3D est établi en utilisant l'outil CAO CATIA V5 (Dassault Systems) (voir figure 2.25 et chapitre 3). L'approche proposée pour sélectionner les primitives qui doivent d'abord être appariées, puis reconstruites, consiste à comparer les images réelles traitées (voir figure 2.24) aux images conceptuelles. Ces dernières sont obtenues en tenant compte des points de vue des deux caméras par l'intermédiaire des paramètres d'étalonnage [GLZH02]. Les images conceptuelles gauche et droite résultantes sont ensuite utilisées comme référence (voir figure 2.26). En introduisant un codage spécifique (par le biais du fichier descriptif associé au modèle conçu sous CATIA V5) et en tenant compte du point de vue de la caméra (spécifié par les paramètres d'étalonnage), ces représentations accentuent les segments visibles dans les deux images. Par conséquent, seuls ces contours accentués vont subir la procédure d'appariement décrite en section (2.4.1). Les caractéristiques non visibles des points de vue courants des caméras déterminés par étalonnage sont éliminées. Ceci conduit à l'obtention d'un modèle réduit de la pièce qui ne contient que les caractéristiques visibles dans les images réelles (voir figure 2.26). Les deux images synthétiques correspondantes peuvent alors être superposées aux images réelles pour les deux points de vue des caméras composant la tête stéréoscopique. Cette comparaison est effectuée point par point, avec les images de synthèse et les images réelles traitées. On ne conserve ainsi que les primitives communes aux deux images. De cette façon, l'utilisation des modèles réduits CAO permet de mettre en correspondance sans erreur les primitives des deux images utiles à la reconstruction de l'objet, même si les images ne contiennent pas le même nombre de contours. On vérifie ainsi, après application de notre méthode d'appariement, la justesse des mises en correspondance obtenues. En outre, ce modèle CAO fournit une information *a priori* en accentuant les primitives visibles dans les deux images (voir chapitre 3, section 3.2.4.2), ce qui est particulièrement utile lors de l'évaluation dimensionnelle de la pièce.

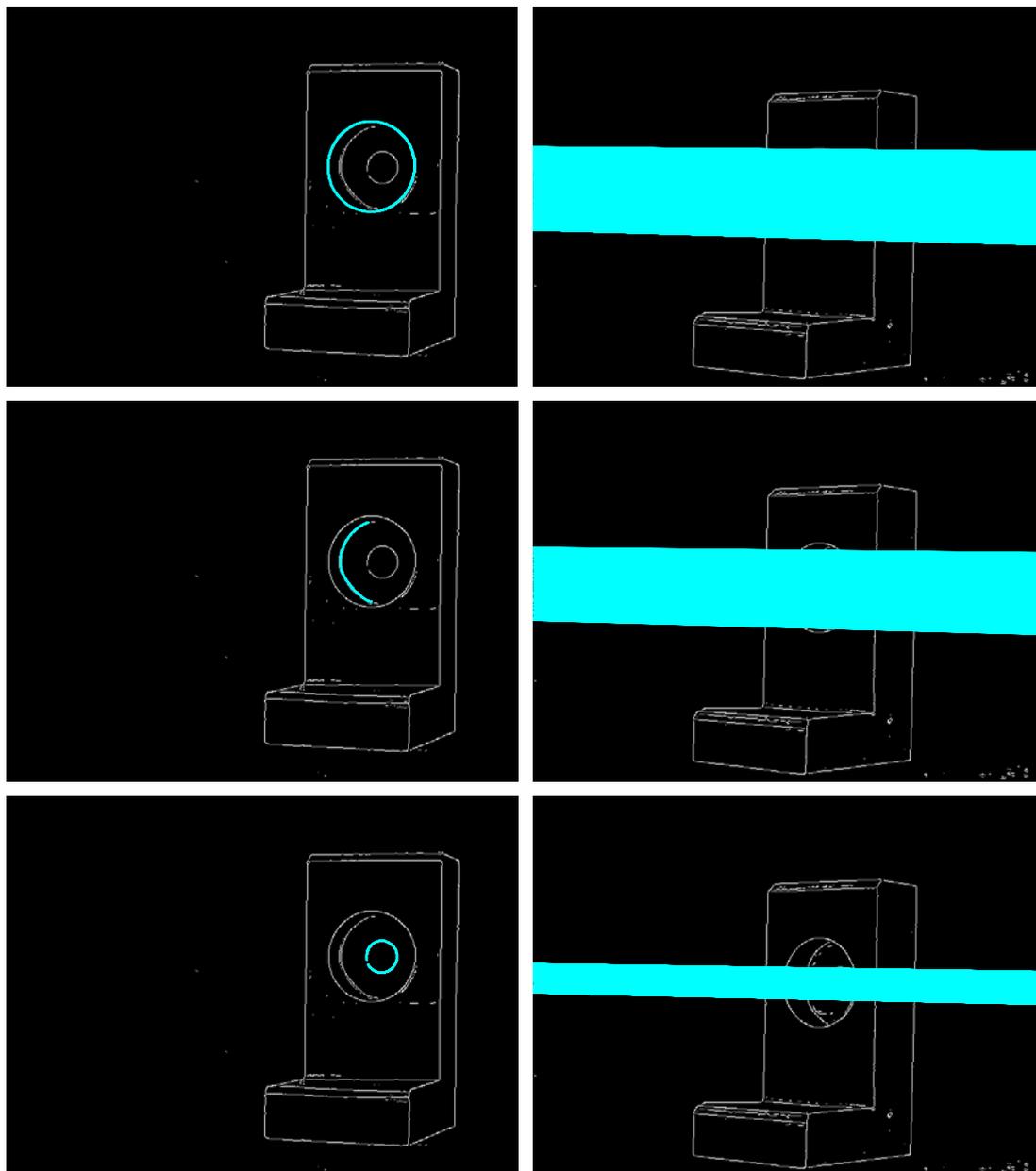


FIG. 2.18 – Ensembles de droites épipolaires, sur les images de droite, calculées à partir des primitives de forme curviligne des images de gauche

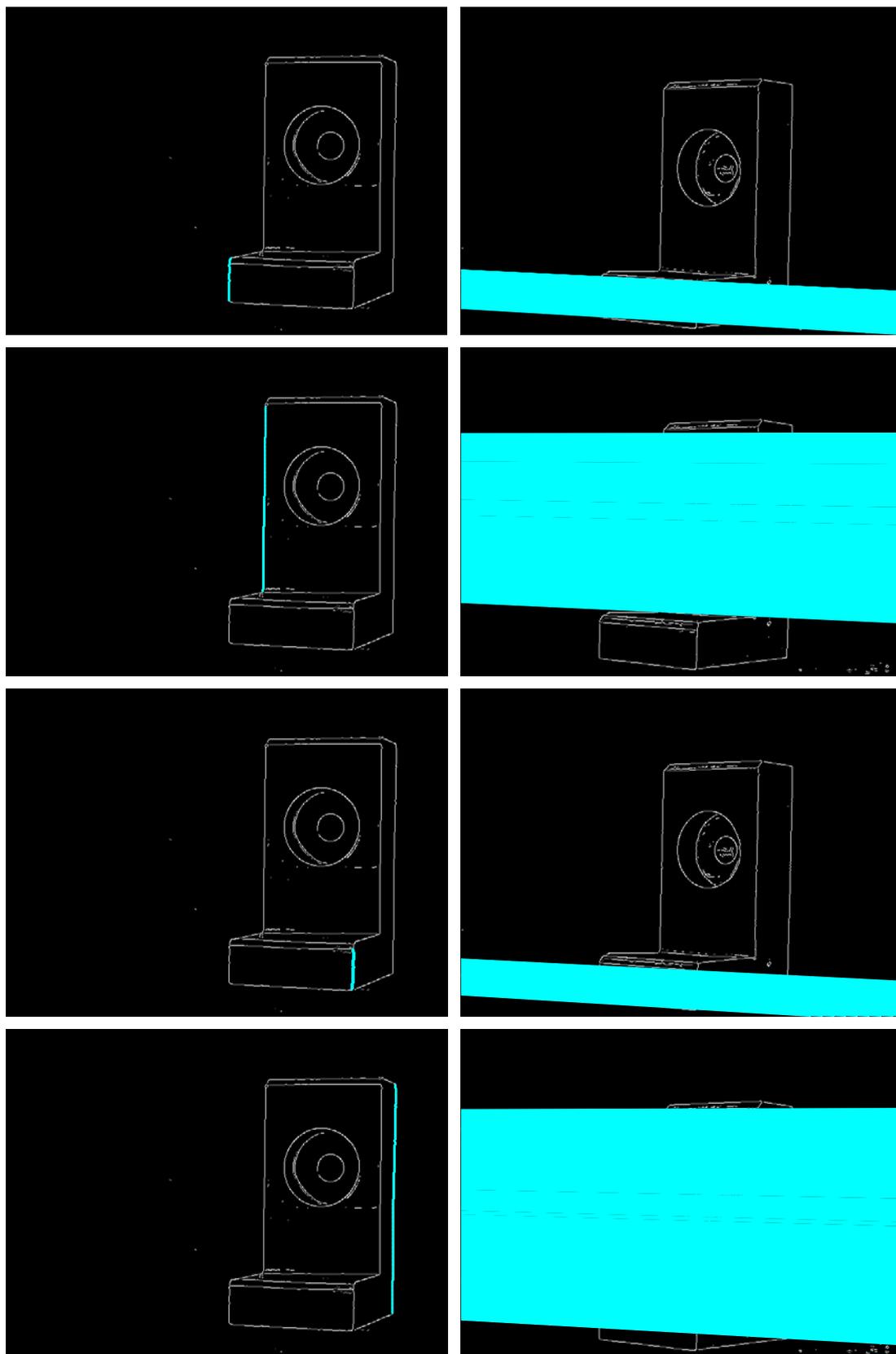


FIG. 2.19 – Ensembles de droites épipolaires, sur les images de droite, calculées à partir des segments verticaux des images de gauche

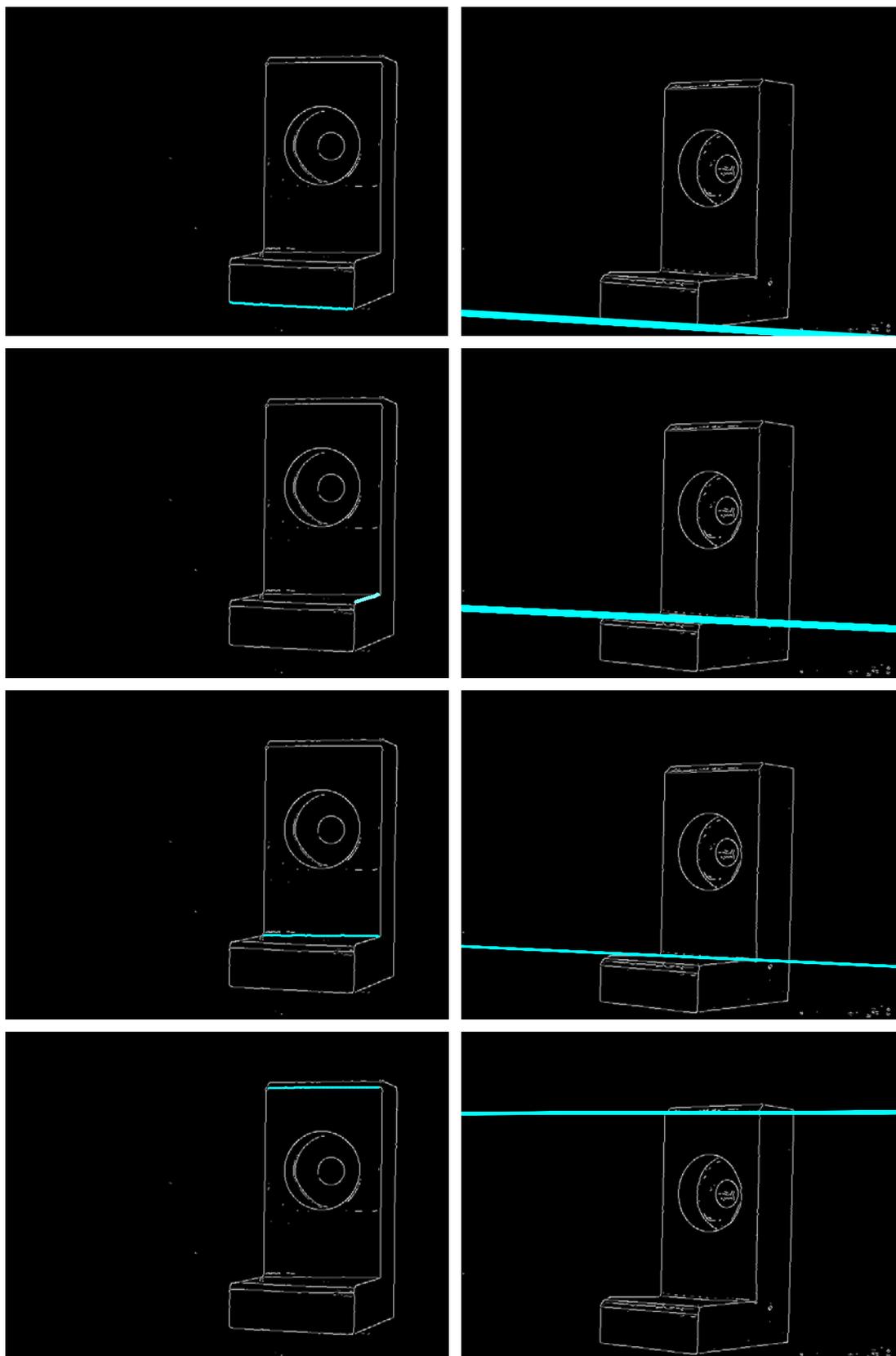


FIG. 2.20 – Ensembles de droites épipolaires, sur les images de droite, calculées à partir des segments horizontaux des images de gauche

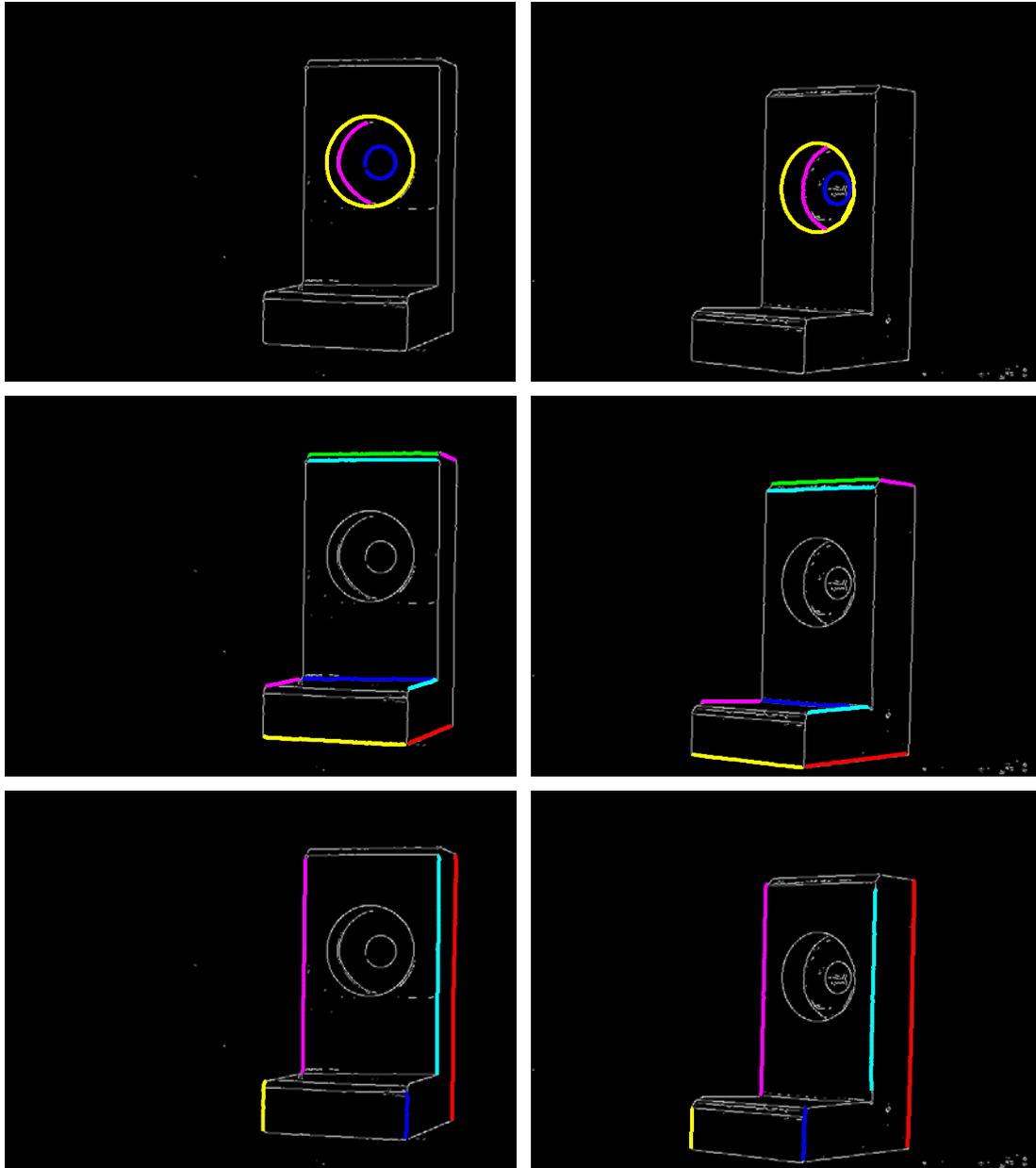


FIG. 2.21 – Appariement des contours extraits des images de la figure 2.1. Les primitives appariées de chaque classe sont de couleur identique dans les images gauche et droite

2.5 Comparaison de données images réelles et conceptuelles

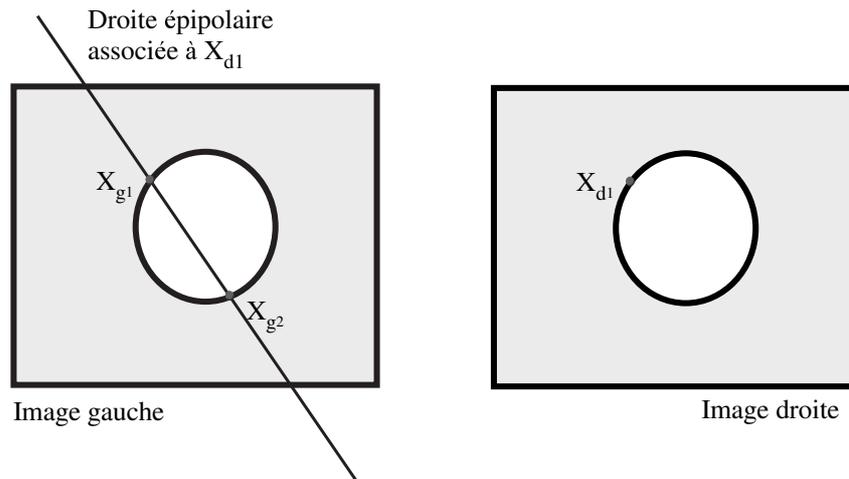


FIG. 2.22 – Ambiguïté éventuelle, observée lors de l'appariement point à point pour les formes curvilignes

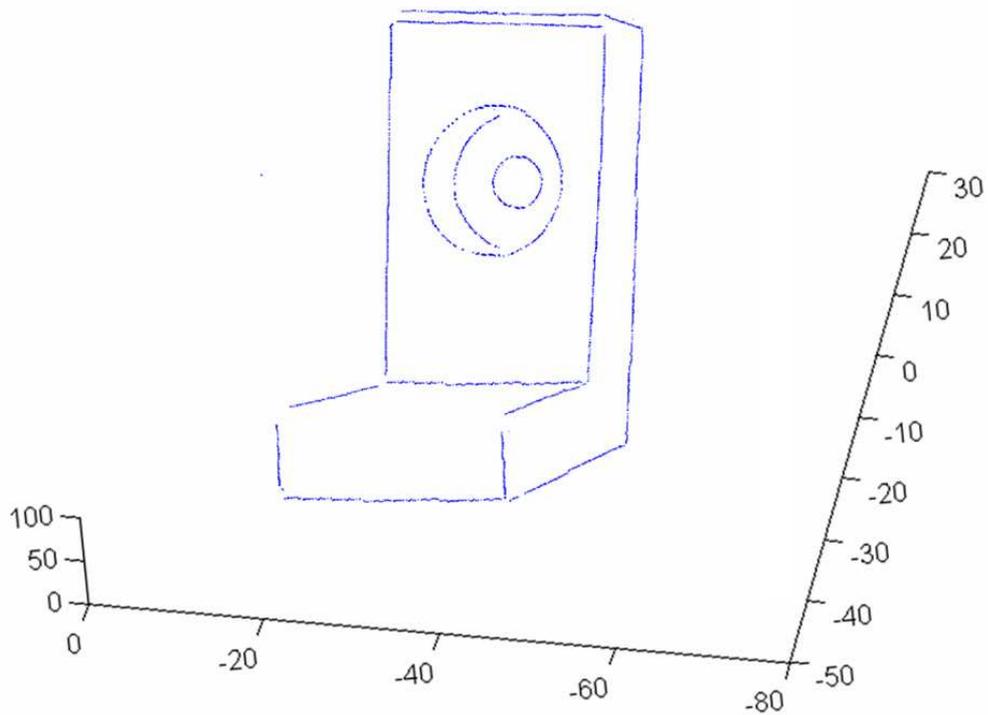


FIG. 2.23 – Reconstruction 3D de l'objet vu dans les images de la figure 2.1

Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques

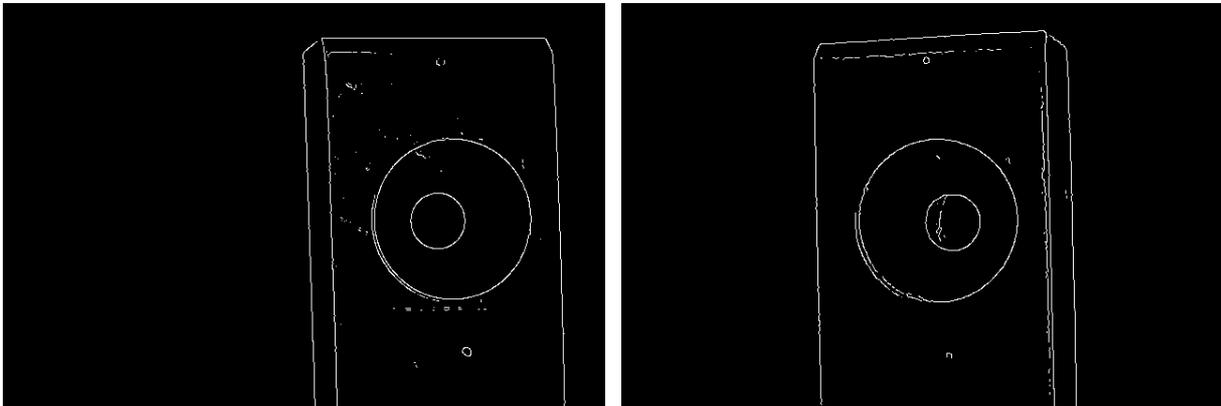


FIG. 2.24 – Exemple de points de contour obtenus à partir d'une paire d'images réelles

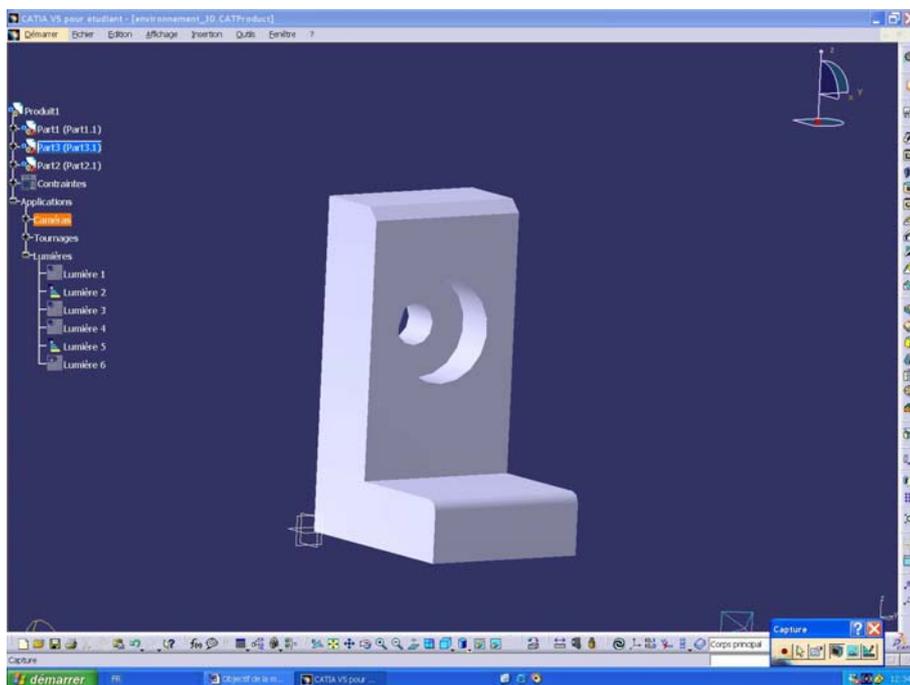


FIG. 2.25 – Représentation CAO CATIA V5 du modèle de la pièce utilisée dans nos essais

2.6 Conclusion

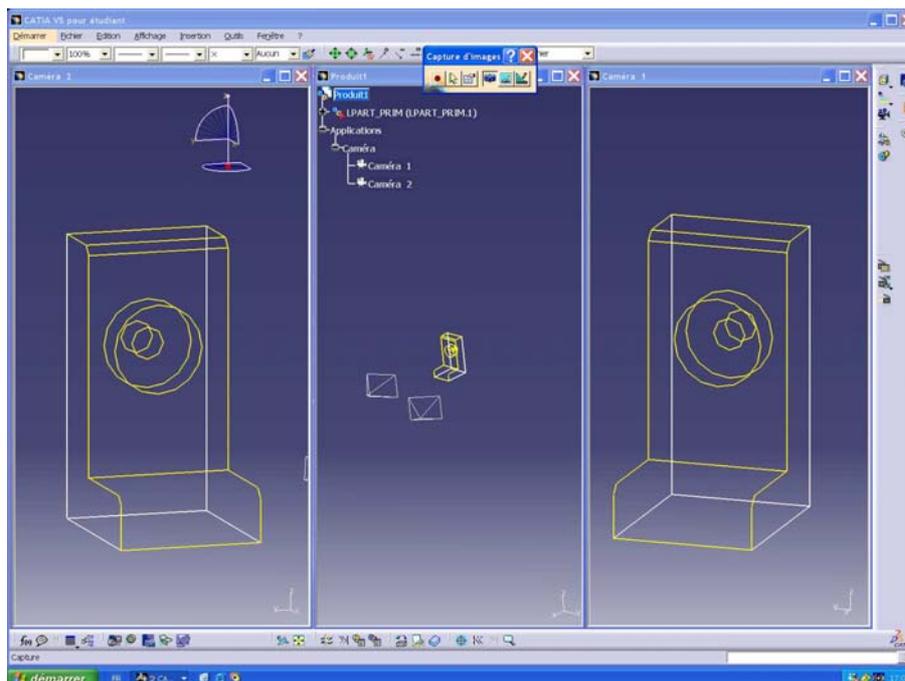


FIG. 2.26 – Modèles fil de fer générés en tenant compte des conditions de prise de vue. Les primitives jaunes sont communes aux deux images et vont subir la procédure d'appariement

2.6 Conclusion

Nous avons décrit dans ce chapitre une approche permettant la mise en correspondance de primitives de type contours. Nous avons décrit les critères de similarité utilisés dans notre approche d'appariement, en précisant leur pertinence et leur dépendance vis-à-vis de notre application, ainsi que la manière dont ils sont construits et utilisés (avant et pendant le processus d'appariement). Nous avons classé ces primitives en fonction de leurs particularités géométriques, ce qui conduit à un appariement correct. À la fin de ce travail sur l'appariement, nous pouvons conclure que, malgré les avancées considérables faites dans ce domaine durant ces dernières années, l'objectif de disposer d'une méthode unique de mise en correspondance, suffisamment générale pour qu'elle soit applicable à tout type d'images, n'est pas encore atteint.

Enfin, dans cette contribution, nous avons introduit une technique originale d'appariement des points de contour. La performance de cet algorithme d'appariement dépend fortement de l'estimation de la matrice fondamentale et de la qualité de la segmentation. L'approche que nous avons adoptée permet de construire une carte de segments 2D pour les deux images gauche et droite pour reconstruire, après mise en correspondance, leur équivalent dans l'espace 3D. La chaîne algorithmique utilisée détecte d'abord les points de contour dans chacune des deux images. Puis, les points de contour obtenus sont associés en chaînes. Les listes de points de contour sont ensuite classés selon un critère géométrique, après leur segmentation. L'étape suivante consiste en la mise en correspondance

Reconstruction 3D de pièces manufacturées à partir d'un appariement robuste d'images stéréoscopiques

des primitives avant de les reconstruire par triangulation. L'information véhiculée par un segment est de nature purement géométrique. Celle-ci renforce la contrainte épipolaire, afin de mieux orienter l'appariement.

Développée spécifiquement pour des pièces manufacturées quasi-polyédriques, en partant d'une classification des points de contour en formes géométriques simples (segments de droites et de forme curviligne), la procédure d'appariement évalue les correspondances établies de la gauche vers la droite et vice-versa, en utilisant ce que nous avons appelé des matrices d'appariement. L'approche repose aussi sur une sélection des listes de points de contour à apparier obtenue en comparant les images acquises avec des images conceptuelles générées à l'aide d'un outil CAO. L'approche a été testée sur différentes paires d'images réelles acquises en laboratoire et les résultats expérimentaux montrent que notre approche conduit à un appariement robuste et à une reconstruction 3D satisfaisante. Il serait néanmoins intéressant, pour augmenter la robustesse de notre appariement en vue d'une reconstruction d'objets manufacturés plus complexes, d'ajouter d'autres contraintes comme, par exemple, l'unicité, l'ordre, etc.

Cette reconstruction de points de contour de l'objet à évaluer, va être comparée avec son modèle fourni par un outil CAO. Cette comparaison permettra d'évaluer l'objet, par exemple dans une chaîne d'inspection dimensionnelle. En effet, les principales applications considérées dans notre travail effectuent une analyse dimensionnelle des objets par comparaison avec leurs modèles conceptuels. Avec cet objectif, nous allons voir dans le chapitre suivant comment la modélisation CAO de pièces industrielles à l'aide d'un logiciel CAO (dans notre cas, CATIA V5) permet de guider le travail du système de vision. Grâce à la conception de ce modèle, nous allons, en effet, pouvoir exploiter une information *a priori* pendant l'évaluation de la reconstruction tridimensionnelle, ou pendant la correction des traitements en ligne conduisant à cette reconstruction.

Chapitre 3

Modélisation CAO et modélisation de l'illumination

3.1 Introduction

Un système de contrôle qualité faisant appel à la vision par ordinateur utilise souvent une connaissance *a priori* sur le contenu des scènes imagées. Cette connaissance peut prendre la forme d'un modèle CAO représentant des caractéristiques particulières d'un objet. Cette modélisation 3D est capable de fournir la géométrie idéale de l'objet à étudier. Pour obtenir une représentation réaliste du modèle, par exemple pour être capable de prédire quelle va être la nature des images acquises, il peut être intéressant, lors de la construction des modèles, de prendre en compte un certain nombre d'éléments décrivant les conditions d'acquisition (nature et position des sources d'éclairage) et le capteur de mesure mis en œuvre (p. ex. les paramètres d'étalonnage des caméras).

La qualité de l'éclairage et de la prise de vue est fondamentale dans un système de vision par ordinateur. Pour une application de contrôle qualité, le choix des paramètres de prise de vue (p. ex. les caractéristiques de la source lumineuse) ne relève pas d'une méthodologie prédéfinie. Cependant, la conception d'un système d'éclairage adapté à un type d'objets nécessite souvent un travail expérimental lourd [AMDG02], [GWCP05].

Dans ce chapitre, nous abordons d'abord la définition et le développement d'une procédure de modélisation adéquate pour la pièce à évaluer, ainsi que la génération des images de synthèse. Puis, dans un second temps, nous revenons sur l'importance de l'illumination qui doit être adaptée aux besoins du système de mesure développé. La qualité de l'acquisition et notamment la mise en valeur de l'information pertinente est entièrement dépendante du réglage des paramètres de la source lumineuse. Nous présentons ainsi de premiers résultats obtenus pour une modélisation de l'illumination, qui pourront être utilisés ultérieurement lors de la replanification dynamique mise en œuvre dans notre système de reconstruction 3D.

3.2 Modélisation CAO

Dans le domaine du contrôle qualité, les caractéristiques dimensionnelles exactes et les formes sont en général connues *a priori* pour une pièce manufacturée. Il est donc possible de modéliser une pièce dont les dimensions sont idéales. Pour cela, il s'avère indispensable de définir et de développer une procédure de modélisation adéquate pour la pièce à évaluer. L'outil CATIA V5 de Dassault Systems est un logiciel de conception assistée par ordinateur très perfectionné (voir Annexe B). Plus particulièrement, la dernière version de CATIA V5 permet de concevoir des modèles 3D très complets et donne aussi la possibilité de simuler l'environnement d'acquisition des images, tout en autorisant un contrôle des paramètres d'illumination, des paramètres du capteur ainsi que de la position de l'objet. Grâce à ses ateliers très spécialisés, la version V5 de CATIA permet de générer des images de synthèse avec un contrôle précis des sources lumineuses. L'utilisation de ce logiciel permet aussi d'accéder à un fichier descriptif du modèle conçu. Ce fichier contient l'ensemble des primitives (points, segments de droites, segments curvilignes, etc.) de la pièce à évaluer.

Pour le développement de notre système d'inspection dimensionnelle, nous avons besoin d'une information *a priori*. Cette information est basée sur la conception du modèle

CAO de la pièce à inspecter. Celui-ci ne contient pas de défauts et peut être utilisé comme référence pour l'évaluation de la pièce industrielle.

En se basant sur la comparaison d'images réelles et d'images conceptuelles de la pièce à inspecter, une technique organisant la coopération entre images acquises et représentations conceptuelles (modèle CAO) est proposée, en vue d'obtenir une meilleure description du contenu des images. L'analyse des images doit pouvoir, en premier lieu, éliminer les effets indésirables (effets d'éclairage, bruit, etc.). La méthode mise au point pour réaliser l'extraction contrôlée des indices intéressants consiste à comparer les images réelles, incluant les nuisances, avec leur modélisation issue de l'outil CAO et représentée par des images de synthèse ne reproduisant pas ces défauts. Pour ce faire, un modèle complet 3D est établi en utilisant l'outil CAO CATIA V5. Des images conceptuelles sont obtenues en tenant compte des points de vue des deux caméras par l'intermédiaire des paramètres d'étalonnage. Elles sont alors comparées aux images réelles pour sélectionner les primitives à apparier, qui serviront à la reconstruction.

Dans cette partie du mémoire, nous allons préciser comment nous avons modélisé une pièce manufacturée à l'aide de CATIA V5, en vue d'obtenir une description utilisable pendant les phases de simulation et d'automatisation de la reconstruction.

3.2.1 Méthodes de modélisation d'un objet

La modélisation a pour objectif principal de déterminer et de structurer les données numériques décrivant un objet, afin de permettre, par la suite, l'application d'algorithmes spécialisés (p.ex. de fabrication, de visualisation, de rendu, etc.). Nous allons dans la suite distinguer trois principales familles de modélisation couramment utilisées pour des objets industriels (comme les pièces mécaniques, p. ex.), à savoir les modélisations de type fil de fer, surfacique et volumique [PAGM90].

3.2.1.1 Modélisation filaire

Le modèle filaire a été historiquement le premier à être mis en œuvre [BB82]. Dans cette modélisation, on représente l'objet par un graphe dont les nœuds sont des points 3D, des coins ou d'autres points remarquables de l'objet et les liens les arêtes physiques de l'objet. Ce type de représentation ne contient pas d'informations sur les surfaces. La représentation filaire reste la plus utilisée pour manipuler des objets polyédriques et des objets non-polyédriques de formes simples.

Modélisation filaire polyédrique. Cette modélisation est largement répandue dans le domaine de la vision par ordinateur [FG93]. Un modèle polyédrique est construit à partir des coins et des arêtes. Les faces sont des polygones délimités par une liste ordonnée d'arêtes ou de coins.

Représentation filaire. Une représentation filaire appelée représentation filaire étiquetée a été proposée par Gigus [GM90]. Une arête convexe est représentée par un signe

3.2 Modélisation CAO

« + », une arête concave par un signe « - », et une arête occluse par une « flèche » (lorsqu'on se déplace sur l'arête dans la direction de la flèche, les faces qui rejoignant l'arête sont sur la droite de celle-ci ; voir figure 3.7). Cette représentation est surtout utilisée pour les objets polyédriques, mais peut aussi s'appliquer à des objets constitués de portions de sphère, de cylindre et de cône. Cependant l'addition de nouvelles étiquettes augmente considérablement le nombre de types de jonction possibles. Ces techniques d'étiquetage sont cependant des outils puissants pour interpréter des représentations filaires.

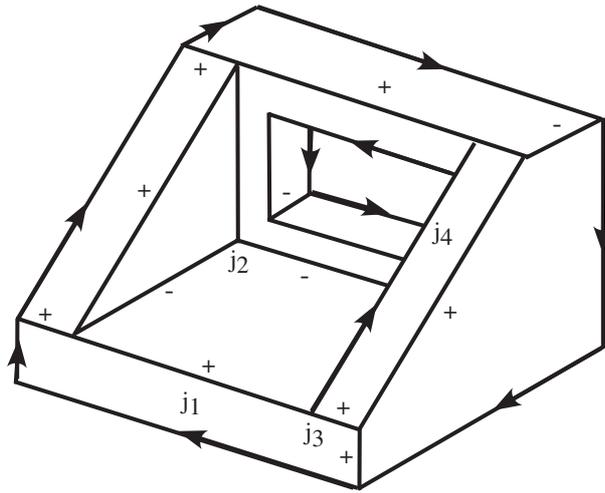


FIG. 3.1 – Une représentation filaire, où j_1 et j_2 sont des jonctions d'arêtes et j_3 est une jonction en T d'après MORN [Mor96]

3.2.1.2 Modélisation surfacique

Le modèle surfacique est très utile pour la conception de formes. Celui-ci doit comporter deux types d'informations : des informations géométriques, comme des coordonnées de points ou des équations de faces, et des informations topologiques comme par exemple la façon dont les informations géométriques sont liées les unes aux autres. Cette représentation a l'avantage de faciliter les descriptions et la visualisation des transformations géométriques, mais elle a aussi ses faiblesses. En particulier, on peut observer des problèmes de cohérence dus au fait que chaque élément de surface de l'objet est décrit indépendamment des autres éléments. Ainsi cette modélisation n'est pas adaptée à l'inspection de pièces manufacturées car les relations entre les différents éléments de la pièce ne sont pas spécifiés.

3.2.1.3 Modélisation volumique

Dans ce type de modélisation, l'objet est défini directement comme étant un volume. Dans cette approche, on prend en compte aussi bien des propriétés géométriques que des propriétés topologiques de l'objet.

Il existe trois modélisations volumiques. Nous pouvons d'abord citer l'énumération de

l'occupation spatiale [BH87]. Dans ce modèle, l'objet à modéliser est représenté par l'ensemble de cellules occupées par le volume de l'objet dans l'espace. Ces cellules peuvent être par exemple des cubes de tailles fixes appelés *voxels* (volume element). Ils sont fréquemment localisés par les coordonnées du sommet du cube le plus proche de l'origine du système d'axes (voir figure 3.2).

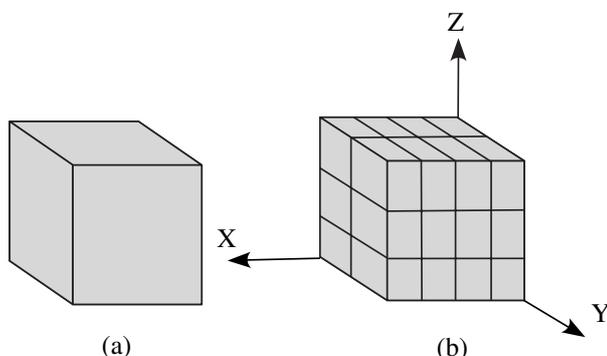


FIG. 3.2 – Modélisation par énumération spatiale d'après Morn [Mor96]. (a) objet spatial à modéliser, (b) approche par énumération spatiale

Cette technique nécessite le stockage en mémoire de très nombreuses informations, ce qui a conduit à proposer d'autres techniques [TG95]. Cette approche possède cependant l'avantage de réaliser simplement les opérations booléennes. Néanmoins, l'inconvénient de cette technique réside dans le fait que ces représentations ne permettent pas de maintenir la continuité entre éléments de surface, notamment à cause des effets de discrétisation inhérents au processus de construction des descriptions. Ceci pénalise la visualisation réaliste des objets.

Le deuxième type de modélisation volumique est la représentation par volumes élémentaires [Kho91]. Le domaine d'application de cette représentation est limité à une famille d'objets, les objets étant indexés à l'intérieur d'une famille par un certain nombre de paramètres. Ce modèle ne se prête pas à la description et au traitement d'objets complexes [Mor96]. Par exemple, (« cube,3 ») permet de désigner l'élément de taille trois unités dans la famille des cubes.

Enfin le troisième type de modélisation volumique repose sur la modélisation par arbre de construction (*Constructive Solid Geometry* ou CSG). Celle-ci est très utilisée et permet de modéliser un objet par des opérations ensemblistes portant sur des primitives volumiques. On dispose d'un ensemble de volumes élémentaires tels que des sphères, des parallélépipèdes, des cônes, etc. Leurs dimensions (longueur, largeur, hauteur, rayon, etc.) sont paramétrées. On dispose de transformations géométriques (translations, rotations, etc.) et d'un ensemble d'opérateurs de composition comme l'union, l'intersection ou la soustraction qui vont permettre de combiner les volumes de base afin de construire des objets 3D complexes. La figure 3.3 donne un exemple de description du type CSG d'un objet.

L'avantage de cette technique est qu'elle permet de définir un solide sans ambiguïté.

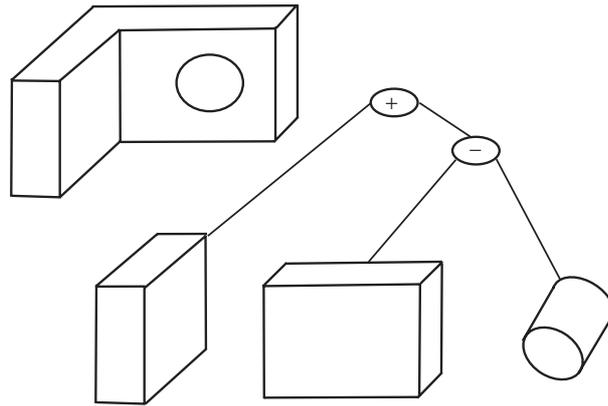


FIG. 3.3 – Construction d'un objet et son arbre CSG associé

Mais l'inconvénient réside dans la difficulté à calculer des propriétés géométriques et physiques pour un objet complexe, comme par exemple un objet incluant des surfaces gauches.

Nous pouvons ainsi constater qu'aucune technique de modélisation n'est entièrement satisfaisante. Nous avons cependant retenu la technique de modélisation par arbre de construction (CSG). En effet, cette dernière permet de faire une modélisation sans ambiguïté d'un objet. Par ailleurs, cette technique est applicable dans le système CAO CATIA V5.

Après avoir présenté l'objectif d'une modélisation CAO dans notre application, ainsi que les différentes techniques utilisables pour la modélisation, nous allons décrire quelques spécificités du système CAO CATIA V5 et nous exposerons la méthode mise en œuvre pour la conception du modèle CAO de notre pièce modèle. Ensuite, nous verrons comment modéliser la scène à analyser, son contenu et les conditions de prises de vue. Puis nous présenterons la mise en œuvre de la génération des images de synthèse par l'introduction des paramètres de prise de vue et nous finirons par exposer l'approche adoptée, permettant l'utilisation des images de synthèse et du modèle CAO, pour diriger les traitements nécessaires à la reconstruction et prévus par un système de planification.

3.2.2 Modélisation de pièces industrielles et système CATIA V5

CATIA V5 est un système qui a été développé par la société Marcel Dassault Bréguet Aviation. Il repose sur une modélisation des objets à l'aide de facettes et de surfaces libres. Ce système était initialement destiné à la conception de fuselages d'avion, mais il est actuellement utilisé dans tous les secteurs industriels faisant appel à la CAO. C'est un logiciel de description et de construction de formes 3D. Il peut être utilisé interactivement ou de manière programmée pour définir des objets. Pour spécifier des formes tridimensionnelles, on suppose disponibles des formes élémentaires. Ce logiciel permet alors la construction de nouvelles formes en utilisant des transformations topologiques (intersection, union ou soustraction) et/ou des contraintes de dimension et de position appliquées aux formes élémentaires. La figure 3.4 présente un exemple de pièce utilisé

dans notre laboratoire. Cette pièce de forme en « L » est constituée de primitives géométriques correspondant à des formes géométriques simples, à savoir des segments de droites et des cercles. Ces primitives peuvent être décrites précisément à l'aide du fichier descriptif du modèle généré par CATIA V5. En tenant compte des données techniques fournies par la figure 3.5, nous disposons des informations nécessaires pour la modélisation de cette pièce sous CATIA V5. Nous avons conçu le modèle en utilisant l'atelier « Part Design » et nous avons choisi pour cela une approche du type CSG ; ce qui nous permet de générer directement une description géométrique de la pièce qui va être exploitée par la suite pour le développement d'un système de planification permettant de guider une reconstruction 3D.

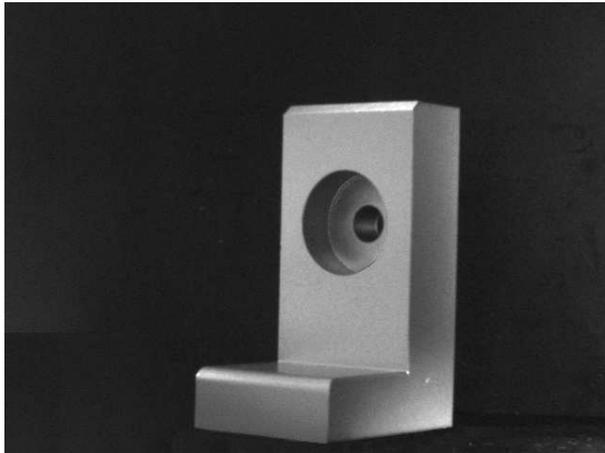


FIG. 3.4 – Pièce réelle de forme « L »

Le modèle 3D de la pièce réelle ainsi obtenu est montré dans la figure 3.6. Nous pouvons aussi générer le modèle en mode filaire comme montré dans la figure 3.7. Cette représentation met en avant les arêtes de la pièce, plus particulièrement utiles pour effectuer l'étape d'inspection dimensionnelle et guider l'appariement.

3.2.2.1 Modélisation de la scène (conditions de prise de vue)

Pour mettre en œuvre notre système d'inspection basé sur une comparaison d'images réelles et conceptuelles, nous devons reproduire avec suffisamment de fidélité les conditions d'acquisition d'une image contenant un objet quasi-polyédrique. À l'aide de caméras virtuelles, on peut simuler le processus d'acquisition. Les images de synthèse qui en résultent peuvent significativement être comparées aux images réelles de la pièce à inspecter. Cette étape de simulation des données réelles suppose disponibles le modèle conçu auparavant dans l'atelier de conception, ainsi que les paramètres réels d'acquisition de l'image, à savoir les paramètres d'étalonnage de la tête stéréoscopique et les paramètres d'illumination. Ces paramètres vont permettre de définir le point de vue de l'image ainsi que les conditions d'éclairage qui vont agir sur les surfaces de la pièce à analyser.

3.2 Modélisation CAO

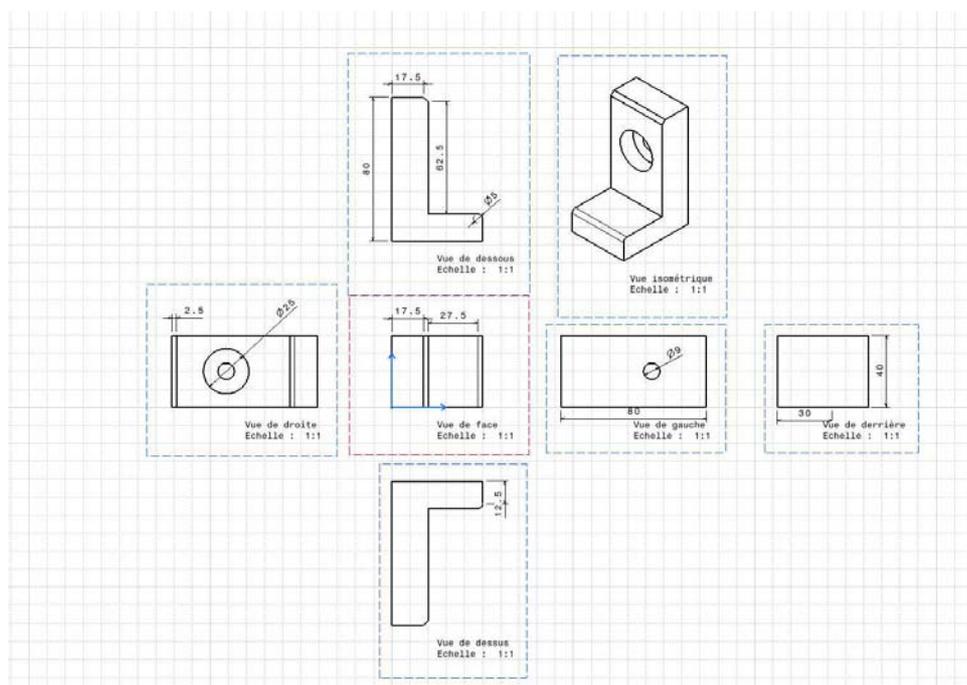


FIG. 3.5 – Données techniques de la pièce « L »

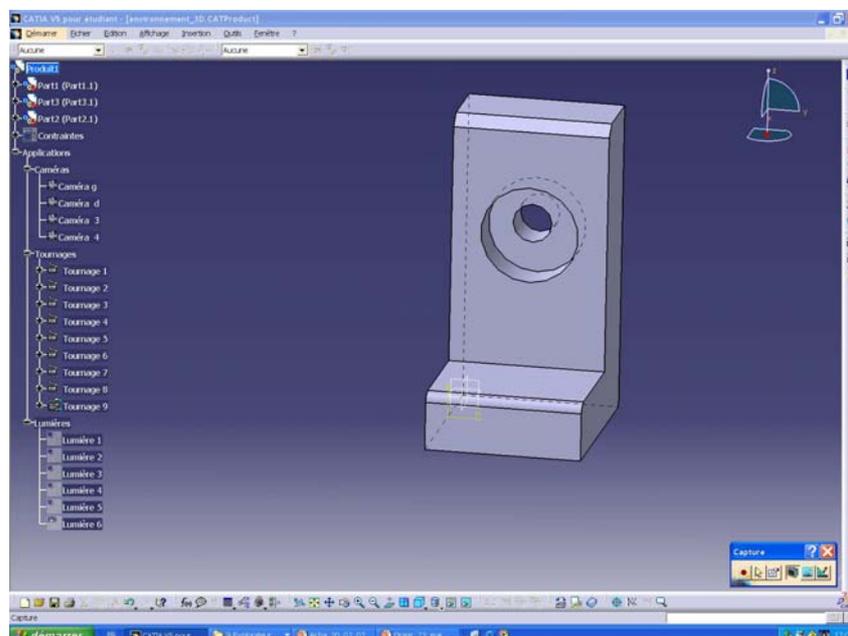


FIG. 3.6 – Modèle de la pièce « L » visualisé en mode rendu réaliste

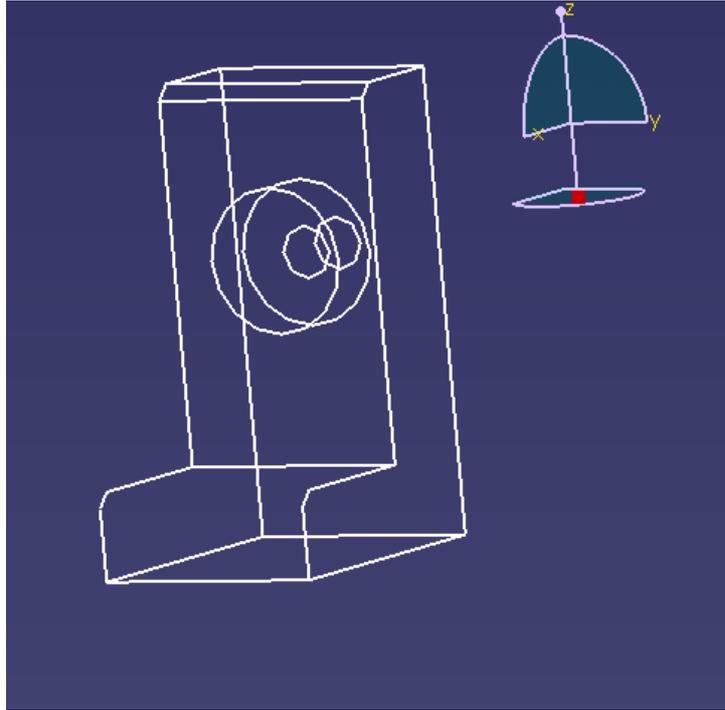


FIG. 3.7 – Modèle de la pièce « L » visualisé en mode fil de fer

3.2.2.2 Génération des images conceptuelles et images filaires

Le modèle géométrique résultant de la conception est exploité avec le même logiciel CATIA V5, mais en utilisant un atelier spécifique de visualisation. Cet atelier, appelé « Photo Studio », permet de simuler des prises de vues en tenant explicitement compte des conditions réelles d'acquisition. À l'aide de cet atelier, nous avons été capables de simuler complètement la procédure d'acquisition à partir de la description des caractéristiques de l'objet et des conditions de prise de vue. Nous avons ainsi généré un couple d'images stéréoscopiques conceptuelles de la pièce, et ceci avec les mêmes points de vue que ceux du capteur du système réel de mesure. Nos images réelles sont acquises au moyen d'un capteur stéréoscopique parfaitement étalonné. Les images de synthèse générées à l'aide de ces paramètres d'étalonnage estimés reflètent alors très précisément la réalité, notamment la géométrie de la scène réelle.

Comme nous utilisons une approche stéréoscopique pour la reconstruction 3D des contours de pièces industrielles, notre intérêt s'est porté sur les primitives du modèle représentant les contours de la pièce réelle. De ce fait, nous nous sommes aussi focalisés sur la génération d'un couple d'images de synthèse stéréoscopiques en mode fil de fer.

Création des caméras virtuelles

Deux caméras de type CCD ont été créées virtuellement dans l'atelier de visualisation. Ces caméras sont réglées de telle sorte qu'elles soient similaires aux caméras réelles utilisées pour l'acquisition des images. Les paramètres que nous avons introduits pour modéliser ces caméras virtuelles sont les suivants :

3.2 Modélisation CAO

- l'angle de vue des deux caméras,
- la position et l'orientation des caméras par rapport au repère CATIA (voir section 3.2.3),
- la projection du centre optique sur le plan image pour les deux caméras,
- le type de la projection (parallèle ou perspective),
- les dimensions de l'image.

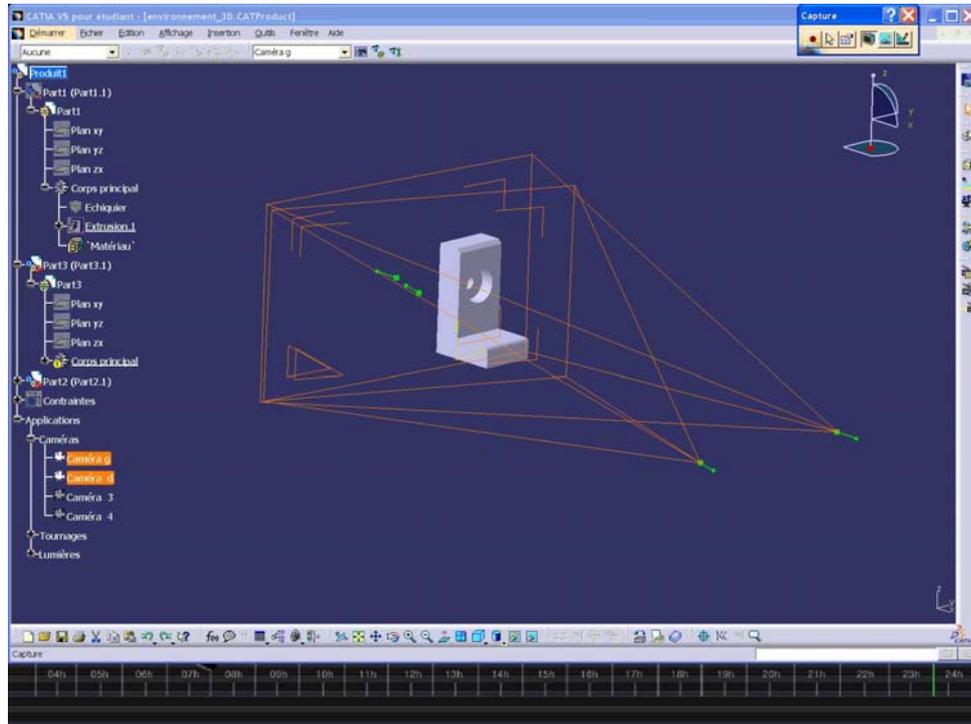


FIG. 3.8 – Création des caméras virtuelles

Ces paramètres sont obtenus à l'issue de la procédure d'étalonnage du capteur stéréoscopique pour un point de vue réel donné.

Création des sources lumineuses

CATIA V5 est un logiciel qui permet de simuler plusieurs types de source lumineuse : source ponctuelle, rectangulaire, surfacique, sphérique, etc. Nous avons ainsi la possibilité de régler l'intensité, la position ou encore le type des sources utilisées. Une autre fonctionnalité de CATIA V5 permet d'éviter les ombres et les phénomènes de réflexions multiples. Pour notre application, nous avons utilisé une seule source ponctuelle. En effet, nous souhaitons pouvoir modéliser les paramètres de cette source lumineuse afin de pouvoir agir sur certains d'entre eux. Après avoir validé cette approche pour une source ponctuelle, il sera ensuite nécessaire d'étudier son extension aux cas de sources lumineuses multiples, surfaciques, voire sphériques. La position de la source lumineuse dans le repère de la pièce modélisée a été choisie après avoir réalisé plusieurs expériences. Nous disposons, sur notre banc expérimental, de plusieurs sources lumineuses ponctuelles et surfaciques (rectangulaires). Des essais réalisés de façon empirique ont conduit à positionner la source ponctuelle aux coordonnées (($X=295.6\text{mm}$, $Y=141.5\text{mm}$, $Z=80.02\text{mm}$) dans

le repère CATIA). Ainsi, nous évitons d'avoir des phénomènes de saturation ou d'ombre trop marqués. La figure 3.9 donne un exemple de modélisation de source lumineuse (ici, ponctuelle).

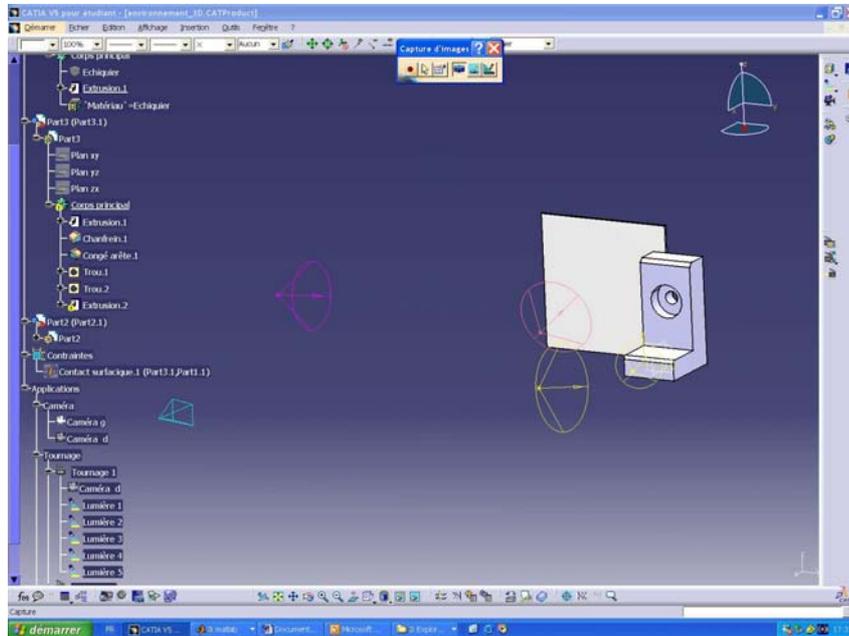


FIG. 3.9 – Exemple de création de quatre sources lumineuses avec CATIA

Nature des surfaces des objets

Afin de modéliser l'interaction physique surface/source lumineuse, ainsi que pour obtenir un rendu de surface réaliste, il est possible de choisir le matériau constituant la surface du modèle conçu (voir figure 3.10). Le matériau que nous avons choisi se rapproche de la texture de la pièce réelle, à savoir une peinture « gris moyen » brillante.

Après avoir modélisé la scène à analyser, son contenu et les conditions d'acquisition, nous avons à effectuer un changement de repère pour exprimer les relations d'étalement, dans le repère CATIA (ce repère est représenté en jaune sur la figure 3.11).

3.2.3 Modélisation du capteur et changement de repère

Le processus de formation d'image peut être divisé en cinq étapes, comme nous l'avons illustré dans la figure (3.11).

Transformation monde/caméra

Cette transformation permet le passage du repère de référence lié à la scène (repère rouge sur le schéma de la figure 3.11) au repère caméra (repère vert), cette transformation est

3.2 Modélisation CAO

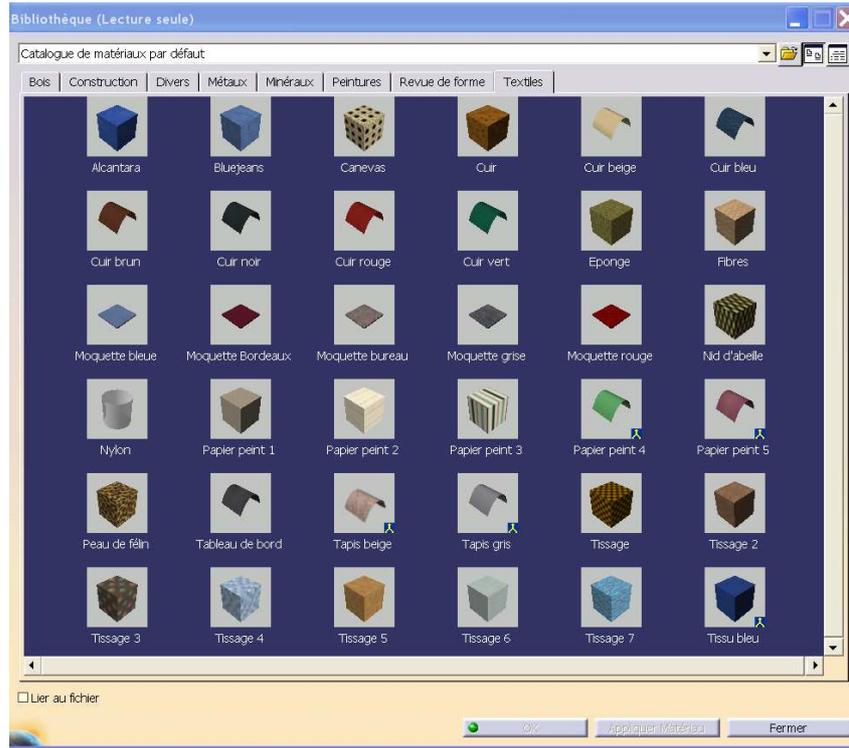


FIG. 3.10 – Choix du matériau

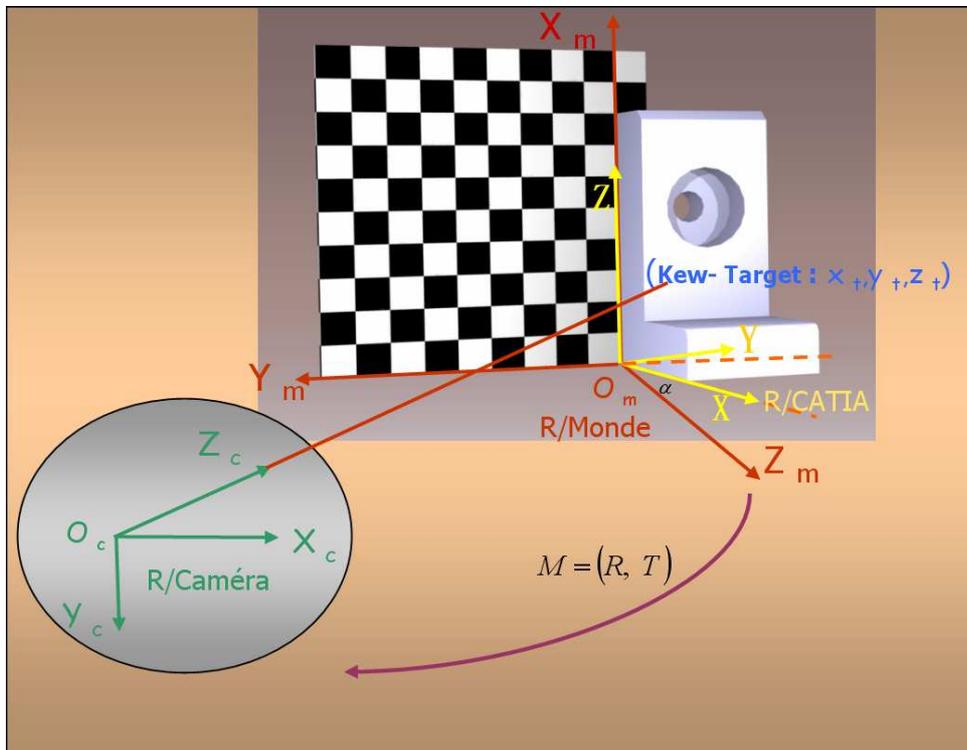


FIG. 3.11 – Modélisation du capteur et changements de repère associés

donnée par la relation suivante :

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_m \\ y_m \\ z_m \end{pmatrix} \quad (3.1)$$

Avec R la matrice de rotation et T le vecteur de translation. Ces paramètres sont obtenus par la procédure d'étalonnage mise en œuvre en utilisant la méthode de Tsai [Tsa86].

Transformation CATIA/monde

Cette transformation permet le passage du repère de modélisation de la pièce sous CATIA V5 vers le repère de référence (ici le repère utilisé pour l'étalonnage du capteur). Ce changement de repère s'avère être très utile. En effet, la position des primitives prédites va en dépendre fortement. Cette transformation est décrite par la relation suivante :

$$\begin{pmatrix} x_m \\ y_m \\ z_m \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_{cat} \\ y_{cat} \\ z_{cat} \end{pmatrix} \quad (3.2)$$

avec α l'angle compris entre l'axe X du repère CATIA et l'axe Z_m du repère du monde (voir figure 3.11). Cet angle exprime une rotation du repère CATIA autour de l'axe Z_m du monde.

Transformation CATIA/caméra

Si on observe la figure (3.12), on peut déduire facilement la transformation entre le repère de la modélisation et le repère caméra.

Cette transformation est représentée par la rotation et la translation suivantes :

$$(R_t, T_t) = (R, T) \times (R', T') \quad (3.3)$$

avec (R, T) la transformation monde/caméra et (R', T') la transformation CATIA/monde.

Transformation caméra/image

Cette transformation est contrôlée par la matrice des paramètres intrinsèques :

$$K = \begin{pmatrix} k_u \cdot f & 0 & C_x \\ 0 & k_v \cdot f & C_y \\ 0 & 0 & 1 \end{pmatrix} \quad (3.4)$$

et est donnée par la relation suivante :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = K \cdot \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} \quad (3.5)$$

Transformation CATIA/image

Enfin, pour obtenir les coordonnées image à partir des coordonnées 3D, nous utilisons la relation suivante :

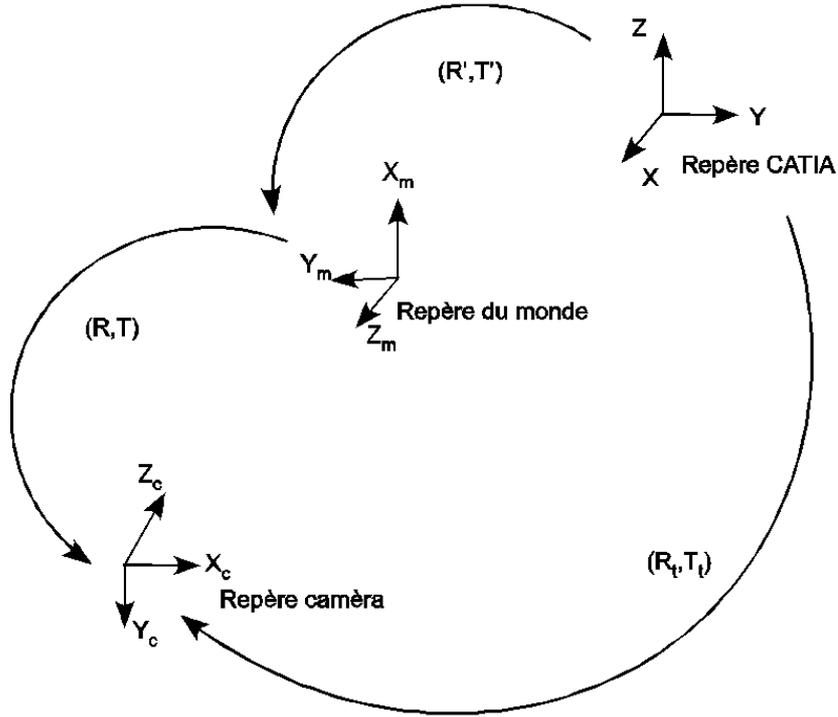


FIG. 3.12 – Transformation CATIA/caméra

$$M = K \cdot A \quad (3.6)$$

avec

$$A = (R_t, T_t) \quad (3.7)$$

On peut en déduire qu'un point dans le repère CATIA est donné dans le repère image par :

$$\begin{pmatrix} su \\ s\nu \\ s \end{pmatrix} = M \cdot \begin{pmatrix} x_{cat} \\ y_{cat} \\ z_{cat} \end{pmatrix} \quad (3.8)$$

avec

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \quad (3.9)$$

Une fois les changements de repère réalisés, nous pouvons générer les images conceptuelles associées au point de vue du capteur stéréoscopique, par introduction des paramètres d'étalonnage obtenus préalablement. Un exemple d'images résultantes est donné par la figure 3.13.

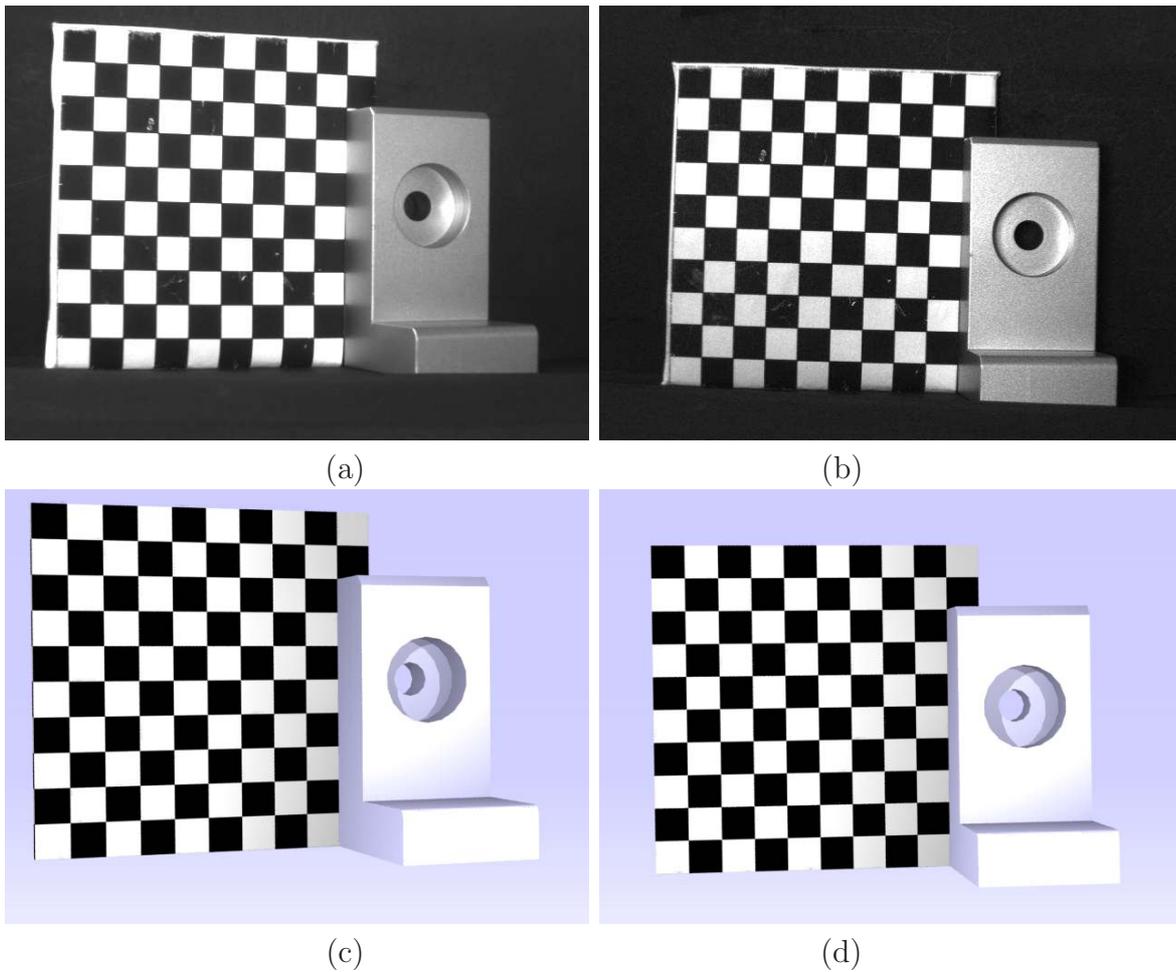


FIG. 3.13 – Exemple d'images stéréoscopiques conceptuelles obtenues ((c), (d)) par introduction des conditions de prise de vue des images réelles (a) et (b) dans le logiciel CATIA

3.2 Modélisation CAO

Ces images conceptuelles peuvent servir par ailleurs pour l'étude des interactions lumière-surfaces (voir section 3.3).

Nous pouvons avec la même procédure générer des images « fil de fer » qui ne contiennent que les primitives du modèle de type contours qui sont simultanément perçues dans les deux images de la paire stéréoscopique. En effet, seules ces primitives sont utiles pour la reconstruction 3D de l'objet. Ces images tiennent aussi compte des conditions d'acquisition des images réelles. Un exemple d'images en mode fil de fer est présenté dans la figure 3.14. À ce stade de l'étude, nous avons terminé de modéliser la scène d'acquisition

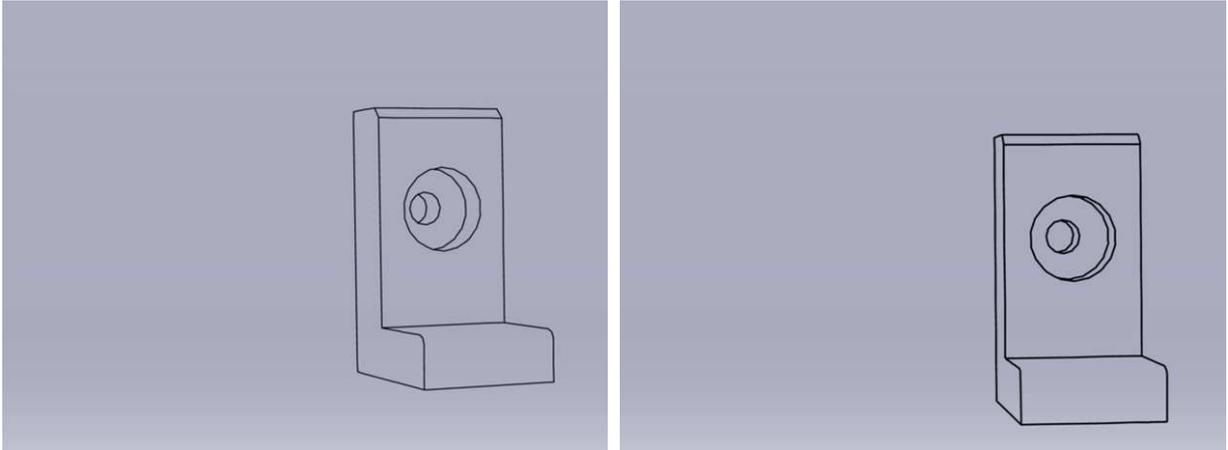


FIG. 3.14 – Images stéréoscopiques filaires obtenues par introduction des conditions de prise de vue (voir les images (a) et (b) de la figure 3.13)

dans CATIA V5. Nous avons ainsi généré les images de synthèse, ainsi que les images filaires, associées au point de vue du capteur stéréoscopique. Les images de type fil de fer sont similaires aux images contours réelles. Nous verrons, dans la section 3.2.4, comment exploiter cette information *a priori* dans le système de planification d'une reconstruction 3D développé dans notre laboratoire.

3.2.4 Exploitation de l'information *a priori*

La chaîne de mesure que nous voulons développer est construite autour d'un système de planification dédié à la reconstruction 3D par stéréovision d'objets industriels. Ce dernier doit exploiter un ensemble de paramètres décrivant la scène (capteur, objet). Ces paramètres comportent notamment les primitives géométriques visibles dans l'image de la pièce à évaluer. Ces primitives doivent être définies par rapport au point de vue du capteur stéréoscopique. Nous déterminons ainsi dans une phase appelée « phase de traitements hors ligne », l'ensemble des informations nécessaires pour décrire par la suite les caractéristiques géométriques de la pièce réelle, pour effectuer ensuite une mesure dimensionnelle de celles-ci.

Pour sélectionner les primitives géométriques qui vont être reconstruites et qui sont associées au point de vue du capteur, nous avons procédé comme suit :

- récupération des informations sur les primitives du modèle, à partir du fichier descriptif généré par CATIA V5. Ce fichier contient la description des primitives de type contours de la pièce. Il donne ainsi les informations 3D de chaque segment de la pièce (p. ex., un segment est représenté par deux points 3D, un cercle est donné par un rayon et un centre, etc.),
- génération des images de synthèse en introduisant les paramètres d'étalonnage du capteur (images conceptuelles gauche et droite),
- génération des images gauche et droite en mode fil de fer, en respectant le point de vue choisi pour le capteur stéréoscopique réel. La version V5 de CATIA nous permet d'éliminer automatiquement les parties cachées de la pièce (ces parties sont celles que la caméra ne peut pas voir) (voir figure 3.15),
- utilisation, par ailleurs, du fichier descriptif pour projeter sur les plans images CATIA, avec les paramètres d'étalonnage, toutes les primitives CATIA,
- enfin, superposition des images fil de fer et des images de l'ensemble des primitives CATIA pour accentuer les primitives communes aux deux images fil de fer.

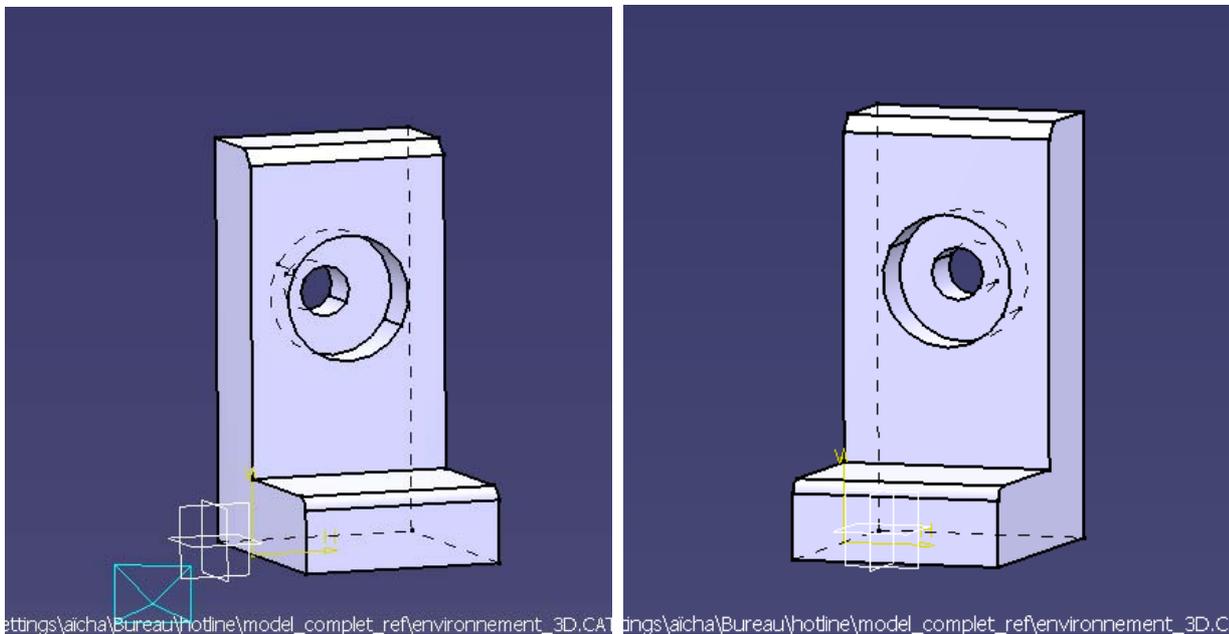


FIG. 3.15 – Parties cachées, en pointillé, du modèle de la pièce « L »

Le schéma descriptif de la figure 3.16, illustre clairement l'approche que nous avons adoptée pour la sélection des primitives pertinentes à l'aide de CATIA V5. Ces images des primitives communes CATIA vont être ensuite superposées aux images réelles pour guider l'appariement des primitives dans les images réelles. Ces opérations sont détaillées dans ce qui suit.

3.2.4.1 Projection des primitives obtenues par CATIA V5

Les primitives (segments de droite, cercles) du modèle de la pièce sont définies dans l'espace 3D et ces définitions sont récupérables en exploitant le fichier descriptif du modèle. Les pièces industrielles, et plus particulièrement notre pièce quasi-polyédrique en L,

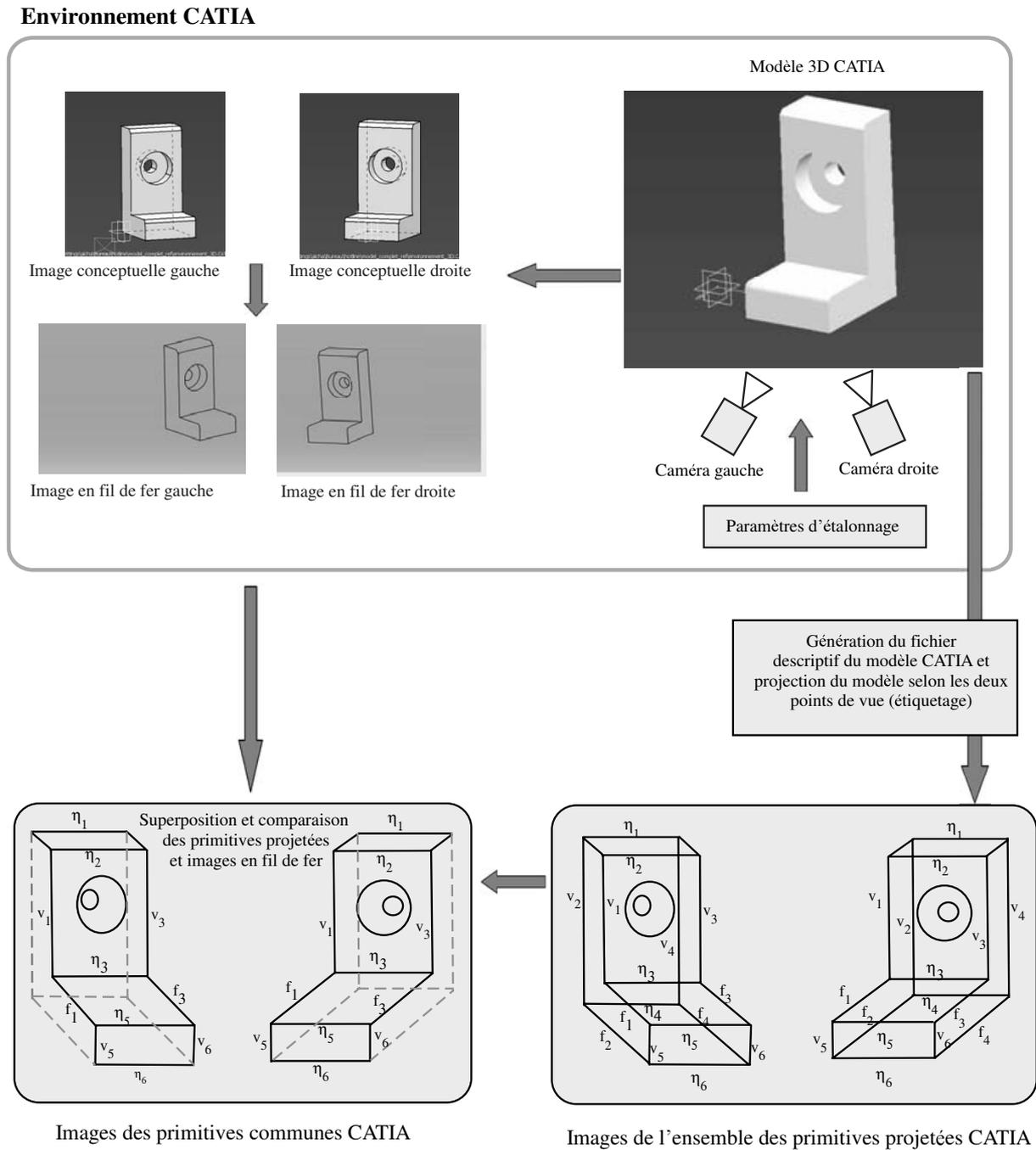


FIG. 3.16 – Récupération des primitives prévues pour un point de vue donné, en vue d’une reconstruction 3D, et en utilisant les images « fil de fer » et le fichier descriptif CATIA V5

comportent souvent des formes géométriques de type segments de droites et elliptiques. Nous construisons ainsi avec CATIA V5 deux projections des primitives sur les plans images : une projection de segments et une projection d'ellipses.

Projection des segments. La définition d'un segment est donnée par les coordonnées de ses deux extrémités. Après projection de ces deux points dans le plan image, les coordonnées 2D de ces derniers permettent de définir le segment de droite correspondant sans aucune ambiguïté. Cette projection utilise aussi les valeurs estimées des paramètres d'étalonnage des caméras. De ce fait, nous pouvons exploiter la relation liant un point 3D de l'espace à son correspondant dans le plan image exprimé en coordonnées images.

Projection des ellipses. Pour les caractéristiques autres qu'un segment, la projection dans le plan image a été faite pour des ellipses et des arcs d'ellipse. Pour généraliser l'approche, les cercles et les arcs de cercles sont considérés comme étant des ellipses ou des arcs d'ellipses dont le grand axe et le petit axe sont identiques. Le problème se ramène donc à la projection d'une ellipse 3D dans un plan 2D.

Dans sa thèse, Zhou [Zho00] a décrit la manière de transformer analytiquement les équations de segments de droite, d'ellipses ou d'arcs d'ellipse de la scène (espace 3D), pour obtenir les expressions correspondantes dans le plan image après projection. L'application de ces transformations aux éléments géométriques constituant la pièce quasi-polyédrique permet d'obtenir les descriptions géométriques analytiques 2D de cette pièce. En nous basant sur ce résultat, nous pouvons exploiter les primitives ainsi prédites pour évaluer la reconstruction 3D des contours réels. La figure 3.17 montre un exemple de projection de primitives géométriques de l'objet « L » décrites dans CATIA, sur le plan des images réelles.

3.2.4.2 Superposition et comparaison

Nous avons simulé la prise de vue à partir des paramètres des caméras. Ensuite, nous avons procédé à une projection des primitives CATIA V5 qui décrivent analytiquement les éléments géométriques constituant la pièce sur le plan image. Ceci permet d'aboutir à une représentation 2D du modèle, c'est-à-dire deux images qui pourront être comparées aux images réelles acquises. La figure 3.17 illustre un exemple de projection des primitives CATIA sur le plan de l'image réelle. À partir de ces deux images de synthèse, nous effectuons une superposition avec les représentations fil de fer, pour accentuer les primitives du modèle. Celles-ci permettent de guider les différentes tâches de traitement pour les données des images réelles qui seront finalement reconstruites et évaluées. Cette comparaison effectuée pixel par pixel permet de ne conserver que les primitives communes aux deux images et permet aussi à la chaîne d'inspection de ne traiter qu'une partie des images stéréoscopiques.

Ces primitives sélectionnées ainsi et le modèle 3D vont être exploités dans le système de planification de la reconstruction 3D sous la forme de paramètres estimés *a priori* (segments de droite et cercles décrits à partir du fichier descriptif associé au modèle).

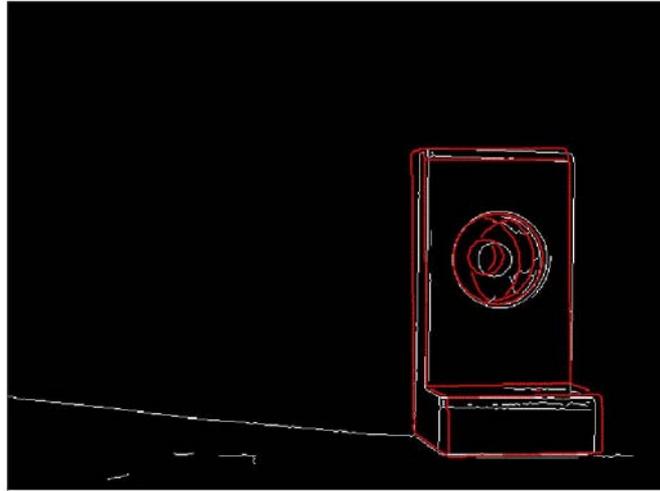


FIG. 3.17 – Exemple de projection des primitives CATIA V5 sur le plan de l'image réelle

3.2.5 Conclusion

Dans notre projet, on modélise la pièce à inspecter à l'aide d'un outil CAO. Le modèle géométrique résultant est ensuite exploité pour simuler la prise de vue. Pour ce faire, on prend explicitement en compte les conditions réelles d'acquisition. Le logiciel CATIA V5 permet de générer une ou plusieurs images conceptuelles de la pièce, et ceci avec le point de vue du capteur du système réel de mesure. Une étape de superposition des images filaires et des images contenant la projection des caractéristiques géométriques du modèle 3D CATIA V5, nous a permis d'accentuer *a priori* les primitives prédites de la pièce. La comparaison des images résultantes aux images réelles facilite notamment l'extraction des caractéristiques pertinentes qui vont être exploitées par la suite dans un système d'inspection dimensionnelle de pièces manufacturées.

3.3 Modélisation de l'illumination

Nous avons vu que le logiciel CATIA V5 permet de régler les paramètres de sources lumineuses. Par contre, il est impossible, à partir des images de synthèse obtenues, de revenir aux paramètres de la source lumineuse en vue, par exemple, de les modifier pour améliorer le contenu de l'image lors d'une nouvelle acquisition. Il est aussi par exemple impossible, si l'on constate que des primitives ne sont pas perçues dans les images réelles, de déterminer et maîtriser les modifications à apporter au système d'illumination pour apporter une correction. Seuls des essais successifs réalisés de façon empirique peuvent conduire à percevoir ces primitives en améliorant les conditions d'acquisition. Une telle démarche n'est pas envisageable si l'on a pour objectif une replanification dynamique de la reconstruction 3D, capable d'automatiquement ajuster des opérations en

fonction des conditions d'acquisitions réelles. Cette approche nécessite la construction de stratégies qui déterminent, par exemple, un changement de réglage des paramètres de la source lumineuse.

Comme étape préalable, il est donc indispensable d'être capable de modéliser des effets lumineux observés (ombre, surintensité, etc.), en particulier lorsque les paramètres de la source lumineuse (position, forme, intensité, etc.) ne sont pas optimisés compte tenu de l'objet particulier à évaluer. La modélisation des effets lumineux joue donc un rôle clé dans la vision par ordinateur, mais la prise en compte de ces phénomènes est d'une telle complexité que peu d'études ont été menées pour le moment.

Dans une première section, nous revenons sur l'importance de l'éclairage dans un système de vision par ordinateur. Puis nous présentons succinctement les différents types d'éclairages usuels. La modélisation théorique de la propagation et de la réflexion de la lumière ainsi que les différents modèles d'éclairage font l'objet de la troisième section. Une partie expérimentale utilisant la modélisation de l'illumination est présentée à la fin de ce chapitre.

3.3.1 Importance de l'éclairage

De nombreuses études [BHH85, Gev00] mentionnent le fait que la conception d'une application de la vision par ordinateur comprend systématiquement une première étape de mise au point des conditions d'illumination. En effet, l'expérience montre que le plus performant des traitements ne peut pas « restaurer » une image dégradée par une acquisition imparfaite. Il est donc nécessaire de développer un système d'éclairage spécifique à chaque application. Ces développements restent encore la plupart du temps empiriques, même si la modélisation de certains phénomènes tend à se généraliser [Cou97, Gev00]. Dans ce cas, les paramètres considérés pour l'élaboration du dispositif d'éclairage sont, entre autres, les suivants :

- la distribution spatiale de l'éclairage (uniformité, directivité, surface, diffusion, etc.),
- la distribution spectrale,
- les perturbations dues à la lumière ambiante,
- la réflectance de l'objet inspecté,
- l'intensité et la stabilité de l'éclairage,
- etc.

Ces paramètres ont également une influence sur les caractéristiques des têtes de mesure utilisées. En effet, par exemple, le réglage de l'intensité d'une source lumineuse est souvent fonction de l'optique utilisée et de la sensibilité de la caméra. Lorsque l'étude de ces paramètres n'est pas réalisable, le développement d'un système d'illumination adapté se fait souvent en suivant des règles bien connues propres à chaque type d'éclairage.

3.3.2 Différents types d'éclairages

Selon le matériau constituant l'objet à inspecter, la lumière n'est pas réfléchiée de la même façon par une surface. En effet, la lumière réfléchiée par une surface en bois est complètement différente de celle réfléchiée par une surface métallique, une surface rugueuse ou une surface lisse.

La prise en compte du matériau des pièces à inspecter et de l'environnement de l'applica-

3.3 Modélisation de l'illumination

tion a conduit à classer les techniques d'éclairage, ensuite classiquement employées pour résoudre des problèmes particuliers liés aux conditions d'illumination. On peut notamment citer [BHH85] :

Éclairage uniforme : utilisation d'une source de lumière circulaire. Le système est composé d'un tube fluorescent circulaire et d'une surface diffusante hémisphérique comportant une ouverture au centre pour permettre la prise de vue. Cette technique va permettre à la caméra de percevoir des creux qui apparaissent sombres avec un système d'éclairage classique. On minimise ainsi des phénomènes d'éblouissement sur d'autres parties de l'objet, car il n'est pas nécessaire d'augmenter exagérément l'intensité lumineuse de la source.

Éclairage sans éblouissement : le problème de l'éblouissement est souvent observé avec des surfaces très lisses (par exemple, des surfaces métalliques brillantes). Pour contrôler cet effet, on utilise généralement des lampes directionnelles ou « spots ». On peut également placer un écran noir muni d'une ouverture devant la caméra pour la prise de vue afin d'éliminer l'image de la caméra réfléchi sur la surface.

Éclairage rasant : cet éclairage permet d'illuminer des surfaces à faible granularité (papier, textile, surfaces métalliques peintes). On utilise une lumière parallèle ou bien une source ponctuelle. Les creux vont apparaître en sombre et les bosses vont apparaître en clair. Cet éclairage permet la détection des défauts de granularité.

Éclairage par l'arrière : ce type d'éclairage est utilisé pour éclairer des objets transparents ou translucides. Il est composé de ce qu'on appelle une « boîte lumineuse » (par exemple, une plaque de verre dépoli éclairé par un néon ou par des fibres optiques). L'objet est placé entre la caméra et l'éclairage.

Toutes ces techniques d'éclairage ne reposent pas sur des études théoriques précises mais sont très fréquemment appliquées pour le développement d'un dispositif d'éclairage spécifique. Il est évident que l'approche qui est la nôtre ne peut pas envisager de se servir de telles techniques empiriques. Elles ont néanmoins l'avantage de permettre de guider le choix des paramètres sur lesquels il convient de se focaliser selon le type d'éclairage développé. Il est cependant indispensable, pour adopter une approche moins empirique, d'arriver à modéliser au mieux les phénomènes d'illumination.

3.3.3 Modélisation de la réflexion de la lumière

Dans une application de vision par ordinateur, le phénomène d'illumination le plus important à prendre en compte est généralement la réflexion de la lumière sur la pièce à analyser. Il convient donc d'étudier plus en détails les différentes modélisations de la réflexion de la lumière disponibles. La réflexion lumineuse sur un objet comporte deux composantes, qui permettent de distinguer les surfaces rugueuses et les surfaces lisses (voir figure(3.18)).

La composante diffuse : elle est due à la pénétration de la lumière incidente sous la surface du matériau. Cette lumière est ensuite réfléchiée par les atomes ou molécules qui composent la matière. Cette composante obéit à la loi de Lambert et est donc identique dans toutes les directions. Ce mode de réflexion est typique des surfaces rugueuses.

La composante spéculaire : observée avec le verre ou les surfaces métalliques bien polies, elle traduit la réflexion de la lumière par un miroir. Étant donné que la surface d'un objet est rarement semblable à un miroir, la réflexion ne comporte pas qu'une composante spéculaire. Des réflexions dites mixtes ou « spéculaires rugueuses » apparaissent ainsi.

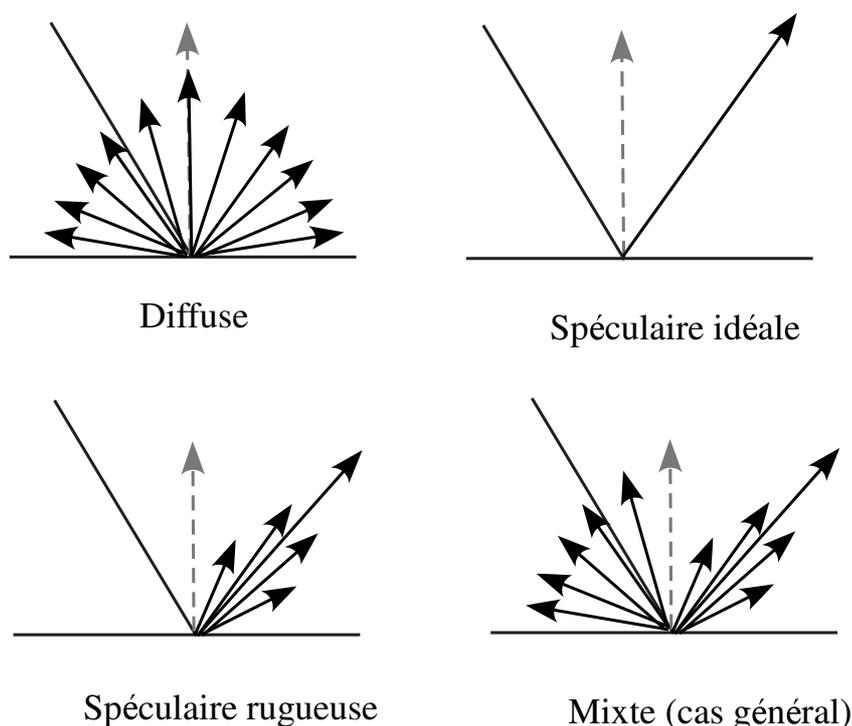


FIG. 3.18 – Différents modes de réflexion

Ces composantes diffuse et spéculaire sont à la base des deux modèles couramment utilisés, le modèle de Beckmann et Spizzichino [BS63] et le modèle de Torrance et Sparrow [TS67], pour modéliser la propagation de la lumière d'une source ponctuelle et la réflexion de celle-ci sur un objet. Le premier repose sur l'optique physique et la théorie des ondes électromagnétiques et analyse la réflexion de la lumière incidente sur la pièce en tenant compte de la rugosité de la surface. Le deuxième utilise l'optique géométrique et ne s'applique que lorsque la longueur d'onde de la lumière est faible comparée aux irrégularités du revêtement de la surface. Ces modèles dits géométriques ont une formulation mathématique beaucoup plus simple que les modèles dits physiques. Mais seuls les modèles de réflexion obtenus par l'optique physique sont capables de décrire la réflexion de la lumière sur des surfaces lisses et sont suffisamment génériques pour tenir compte de rugosités variables [NIK91].

3.3 Modélisation de l'illumination

3.3.3.1 Modèle de Beckmann et Spizzichino

La théorie des ondes électromagnétiques utilisée dans cette modélisation conduit à exprimer l'interaction entre la surface de l'objet et l'onde incidente par les équations de Maxwell. La résolution de ces équations est obtenue en prenant en compte des conditions aux limites imposées par la surface. Ces conditions sont souvent liées aux propriétés électriques de la surface.

Pour généraliser leur approche aux surfaces de rugosité variable, les auteurs ont considéré que la lumière réfléchie par une surface comporte trois composantes primaires : un lobe diffus, un lobe spéculaire et un pic spéculaire. On retrouve ainsi les réflexions dites mixtes (voir 3.19).

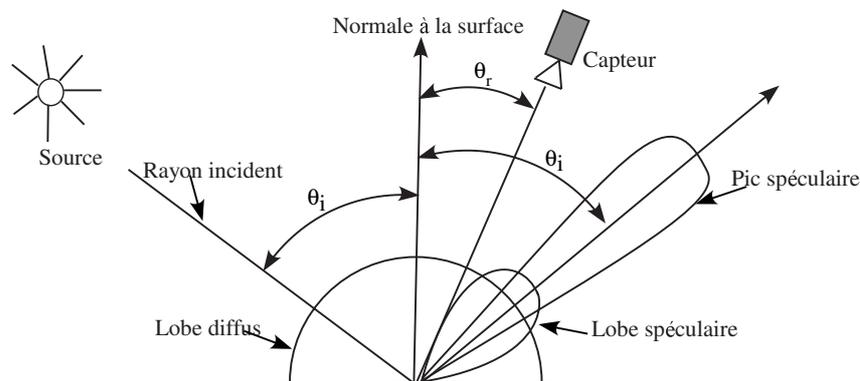


FIG. 3.19 – Modèle standard d'un système d'illumination avec trois composantes de réflexion [KH03]

Le lobe diffus représente le mécanisme de rayonnement interne et est distribué dans la direction normale à la surface. Le lobe spéculaire représente la réflexion unique de la lumière incidente et est orienté dans la direction spéculaire symétrique à l'angle incident. Enfin, le pic spéculaire traduit la réflexion de type miroir qui est généralement l'effet dominant sur les surfaces lisses.

Avec cette approche, la luminance est alors une fonction assez complexe des angles d'incidence et de réflexion et de la rugosité. Les diagrammes de luminance représentent la variation de la luminance en fonction de l'angle de réflexion (position du capteur) et permettent de prédire l'intensité lumineuse reçue par un capteur en fonction des paramètres de la source lumineuse, de l'orientation et de la rugosité de la surface.

Plusieurs généralisations de cette approche ont été proposées, notamment des extensions à des sources lumineuses diffuses [LKGC99] ou à des formes de sources lumineuses différentes [LTYM99, Kru97].

3.3.3.2 Modèle de Torrance et Sparrow

Les modèles géométriques obtenus par simplification des modèles physiques sont beaucoup plus simples à mettre en œuvre mais ne peuvent s'appliquer qu'aux surfaces rugueuses car on suppose que la longueur d'onde incidente de la lumière est petite devant

la rugosité (rugosité d'au moins 1 micromètre).

On considère que la lumière incidente est constituée d'un ensemble de rayons se propageant en ligne droite dans un milieu homogène. Ainsi, on suppose que chaque rayon est réfléchi par la surface dans la direction spéculaire. À partir de ces hypothèses, le modèle de réflexion peut être obtenu en utilisant une analyse purement géométrique. Mais les modèles géométriques ne sont que des cas limites des modèles physiques et ont l'inconvénient de ne pas pouvoir s'appliquer aux surfaces lisses. De ce fait, ce modèle ne convient pas à notre application.

3.3.4 Premiers essais de modélisation de l'illumination

L'approche retenue pour modéliser les phénomènes d'illumination repose sur l'utilisation d'un modèle physique de type Beckmann-Spizzichino. Cela signifie que, dans un premier temps, seules les sources ponctuelles sont abordées, une généralisation aux autres formes étant envisageable. Sur le marché des logiciels de modélisation de l'illumination, il existe un nombre très important de logiciels permettant d'ajuster les paramètres liés aux sources lumineuses. Bon nombre d'entre eux ne reposent pas sur un modèle optique physique et ont pour unique objectif la création d'images de synthèse. Ces logiciels ne nous intéressent pas puisqu'ils ne permettent pas d'avoir une approche physique permettant éventuellement d'exprimer les modifications souhaitables des paramètres liés à une source.

Cependant, parmi ces logiciels, il en existe trois dont la modélisation repose sur un modèle optique physique : CORNELL BOX, APILUX et RADIANCE [COR, APL, RAD]. Mais le but de ces logiciels n'est pas lié directement à nos objectifs. En effet, RADIANCE est destiné à aider les concepteurs de systèmes d'éclairage et les architectes à prédire les niveaux et les effets d'illumination dans l'espace avant la construction, par exemple, d'immeubles. APILUX est un logiciel de conception optique dédié à l'étude des transferts d'énergie lumineuse dans des systèmes opto-mécaniques. Ce logiciel évalue des intégrales de Kirchoff dérivées des équations de Maxwell. Enfin, l'objectif de CORNELL BOX est de simuler globalement l'illumination dans des scènes de synthèse. Ce dernier a été développé par des infographistes et utilise des techniques de lancer de rayons et de radiosité.

Ces trois logiciels analysent la propagation de l'énergie lumineuse et prennent en compte l'ensemble des paramètres physiques d'un système.

CORNELL BOX et RADIANCE sont capables de générer des images de synthèse en simulant physiquement la propagation de la lumière, tandis qu'APILUX délivre des images de nombreux paramètres radiométriques et photométriques. Les paramètres évalués comportent, entre autres, la luminance et l'éclairement et sont évalués sur des surfaces à définir dans le logiciel. Ceci revêt un grand intérêt pour notre application, d'autant plus que le modèle CAO généré par CATIA V5 peut être directement utilisé par ce logiciel.

En effet, dans notre application, nous avons l'objectif d'accéder à des paramètres liés à l'illumination et mesurables sur les images réelles, afin de pouvoir obtenir des informations sur la manière de modifier l'éclairage. Étant données les nombreuses distributions de paramètres calculées par ce logiciel (éclairement, luminance, intensité, énergie, absorp-

3.3 Modélisation de l'illumination

tion, ...), il a semblé intéressant d'en étudier l'évolution lorsque, par exemple, la source lumineuse ponctuelle est déplacée. Des distributions de luminance et d'éclairement ont ainsi été générées à la fois sur la pièce en L à analyser et sur les fonds de l'image.

Plusieurs simulations ont été faites à l'aide du logiciel APILUX sur la pièce test. Le modèle CAO de la pièce a été utilisé, ainsi que les caractéristiques liées à sa texture. Nous nous sommes intéressés aux distributions d'éclairement et de luminance obtenues sur la pièce et dans son environnement, à savoir le socle sur lequel est posée la pièce et le fond devant lequel elle est située. L'évolution de ces distributions a été étudiée lorsque la source lumineuse ponctuelle est déplacée selon l'axe Y (voir figure 3.20).

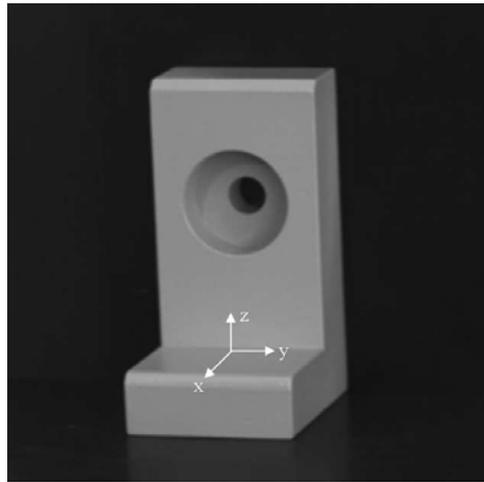


FIG. 3.20 – Image de synthèse utilisée pour étudier les distributions de paramètres lors du déplacement de la source lumineuse par rapport au repère choisi

L'étude des distributions d'éclairement et de luminance simulées sur les zones violettes de l'objet (voir figure 3.21) ne permet pas de faire apparaître les formes réelles de la pièce.

Cependant, les effets de la forme de l'objet sur les diagrammes de distribution peuvent être observés. Il est effectivement possible de détecter l'effet induit par la lumière réfléchie par des surfaces adjacentes (voir figures 3.22, 3.23) : la distribution d'éclairement fait apparaître un niveau d'éclairement supérieur sur le côté gauche des diagrammes correspondant au début du plan vertical de la pièce. De plus, l'atténuation observée sur le côté droit des diagrammes (figure 3.23) peut aisément être reliée à la forme de l'objet qui comporte une zone de transition arrondie entre le plan vertical et le plan horizontal du bas de l'objet.

Cependant, la seule quantification de la réflexion lumineuse et son atténuation dans certaines zones ne permet pas de reconstruire l'objet 3D, ni de relier ces informations à des paramètres de la source lumineuse (position et/ou intensité). Les effets mesurés démontrent seulement que la propagation de la lumière modélisée à l'aide d'une approche physique conduit à en observer les effets dans les zones analysées. Cette observation peut

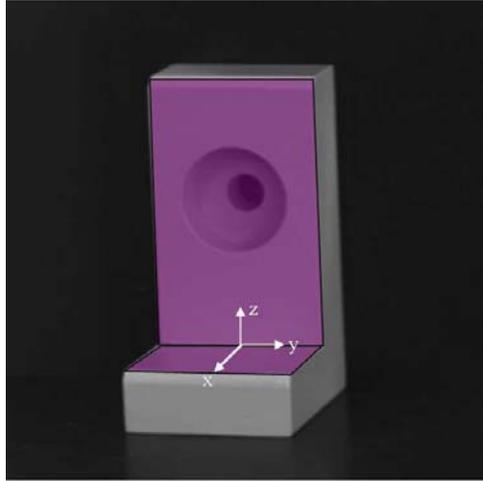


FIG. 3.21 – Zones test violettes d'un objet 3D typique utilisées dans nos simulations des effets lumineux

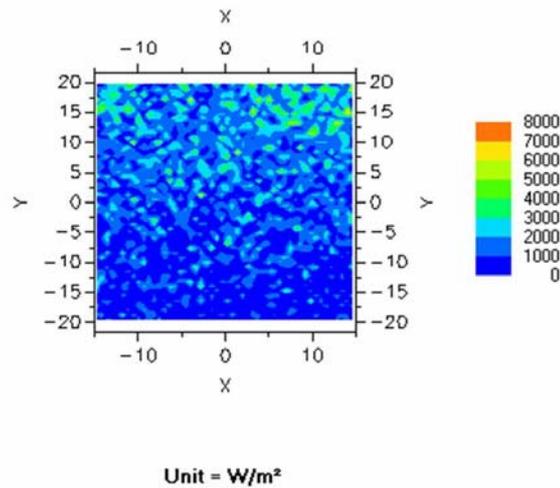


FIG. 3.22 – Distribution d'éclairage modélisée pour la face violette horizontale de l'objet réel (figure 3.21) avec une source positionnée à $Y=90mm$

3.3 Modélisation de l'illumination

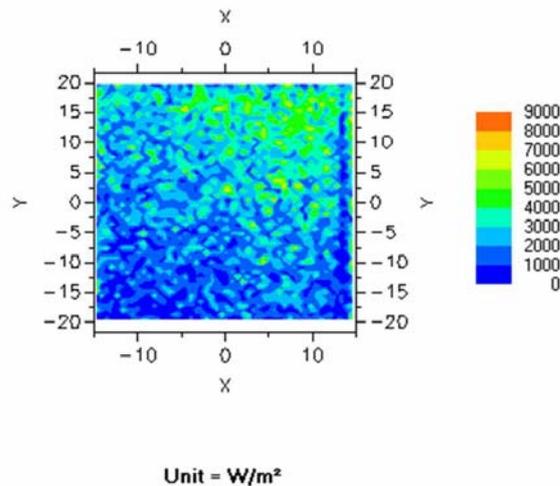


FIG. 3.23 – Distribution d'éclairement modélisée pour la face violette horizontale de l'objet réel (figure 3.21) avec une source positionnée à $Y=30\text{mm}$

être exploitée, lors de l'exécution de l'application en ligne, pour effectuer un traitement d'images local permettant, dans l'image, de réduire par exemple les surintensités dues au système d'illumination.

Par contre, l'étude des distributions d'éclairement et de luminance sur les zones verticale et horizontale du fond de l'image (figure 3.24) permet d'obtenir des informations sur la position de la source lumineuse [KFH05].

En effet, étudions l'évolution de la distribution de l'éclairement lorsque la source lumineuse ponctuelle est déplacée selon l'axe Y (voir figures 3.25, 3.26, 3.27, 3.28).

Lorsque la source lumineuse est positionnée en $y = 0$, cela signifie qu'elle est située exactement face à la pièce. Le fond et le support sur lequel est posée la pièce reçoivent donc des distributions d'éclairement symétriques ainsi que cela apparaît sur les images de gauche des figures (figures 3.25, 3.26).

Lorsque la source est déplacée dans le sens des y positifs, la forme de la pièce provoque des effets d'ombre et diminue l'éclairement du côté inverse du déplacement de la source. Étant donnée la forme de la pièce, à l'aide des déformations observées des distributions de l'éclairement, il est possible d'évaluer la position de la source lumineuse. En mesurant le déplacement de la projection du point A (resp. B ; voir figure 3.24) selon la position de la source lumineuse dans la figure 3.27 (resp. figure 3.28), on peut tracer la droite passant par A et son projeté (resp. B et son projeté) comme montré exagérément sur la figure 3.29. La recherche du milieu de la distance minimale entre ces deux droites permet d'estimer la position de la source lumineuse.

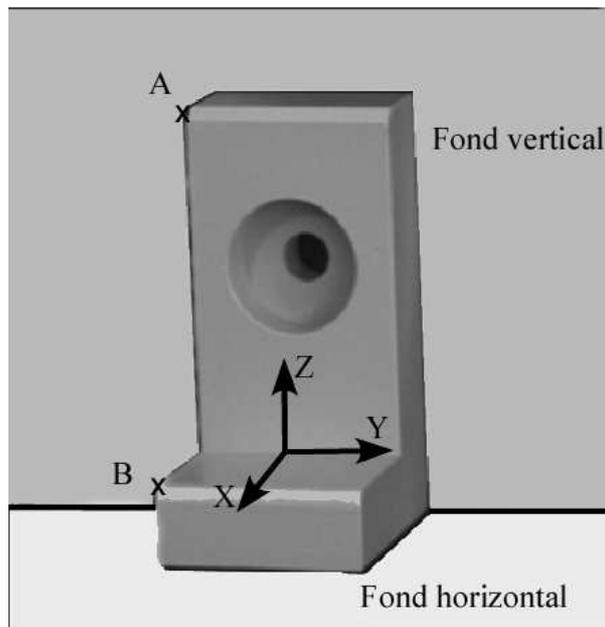


FIG. 3.24 – Fonds d'image vertical et horizontal pour lesquels l'analyse de l'éclairément et de la luminance est effectuée

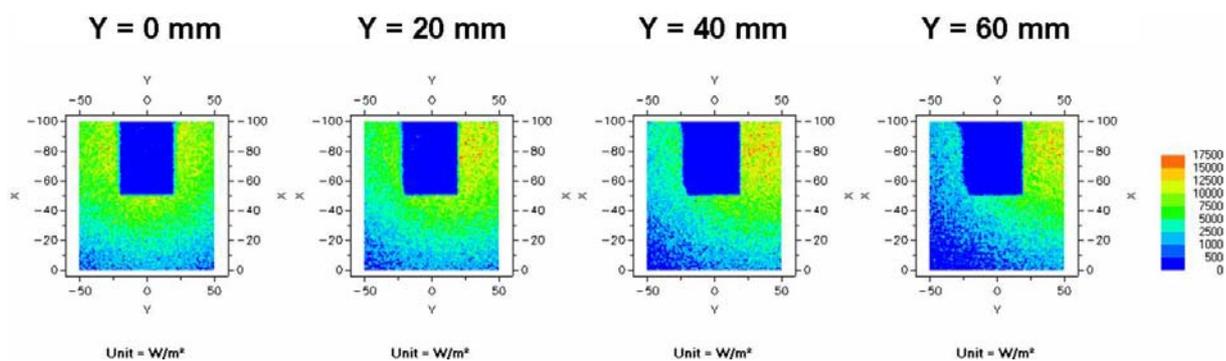


FIG. 3.25 – Distribution de l'éclairément sur la surface horizontale du fond de l'image selon différentes positions de la source le long de l'axe Y

3.3 Modélisation de l'illumination

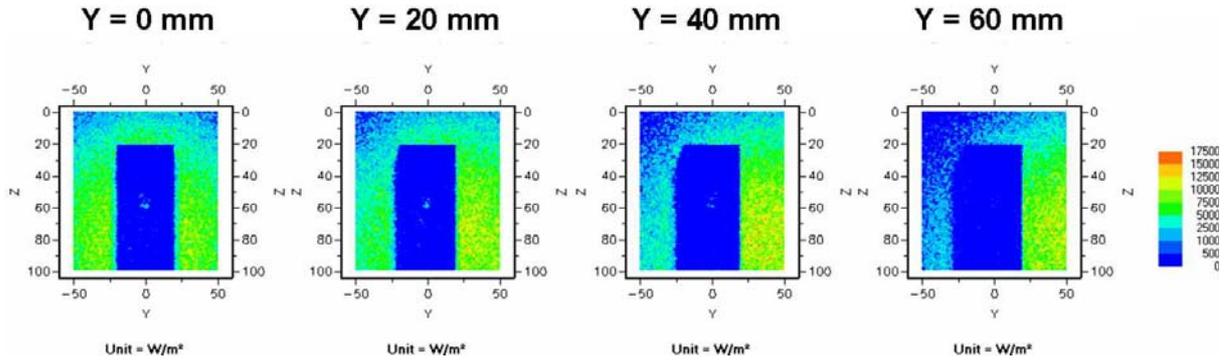


FIG. 3.26 – Distribution de l'éclairage sur le fond vertical de l'image selon différentes positions de la source le long de l'axe **Y**

Prenons, par exemple, les images des figures 3.25 et 3.26 correspondant à $Y=60\text{mm}$. La source réelle a été positionnée à $X=150\text{mm}$, $Y=60\text{mm}$ et $Z=150\text{mm}$ dans le repère de la figure 3.24. L'estimation de la position de la source conduit à des coordonnées relativement proches des coordonnées réelles : $X=161\text{mm}$, $Y=75\text{mm}$ et $Z=173\text{mm}$.

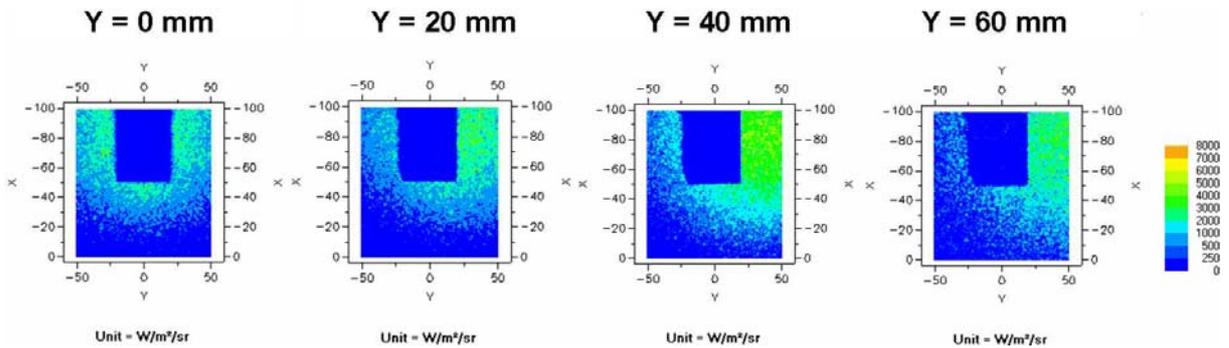


FIG. 3.27 – Distribution de la luminance sur la surface horizontale du fond de l'image pour différentes positions de la source le long de l'axe **Y**

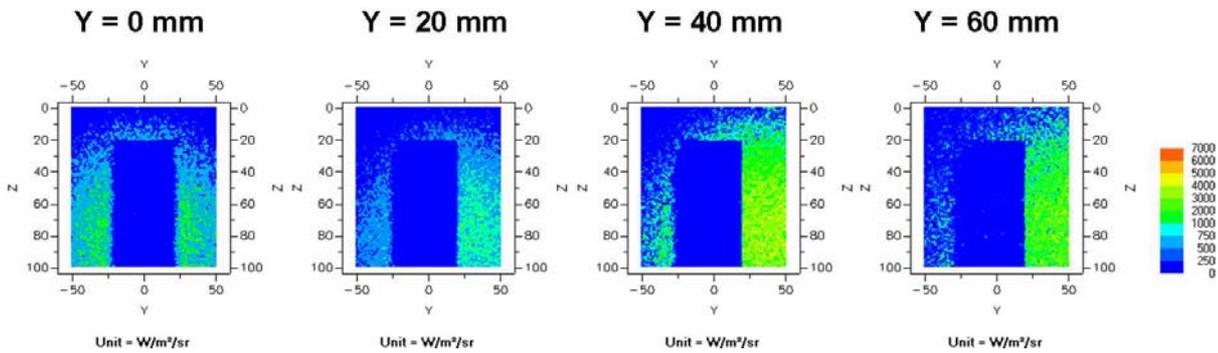


FIG. 3.28 – Distribution de la luminance sur la surface verticale du fond de l'image pour différentes positions de la source le long de l'axe Y

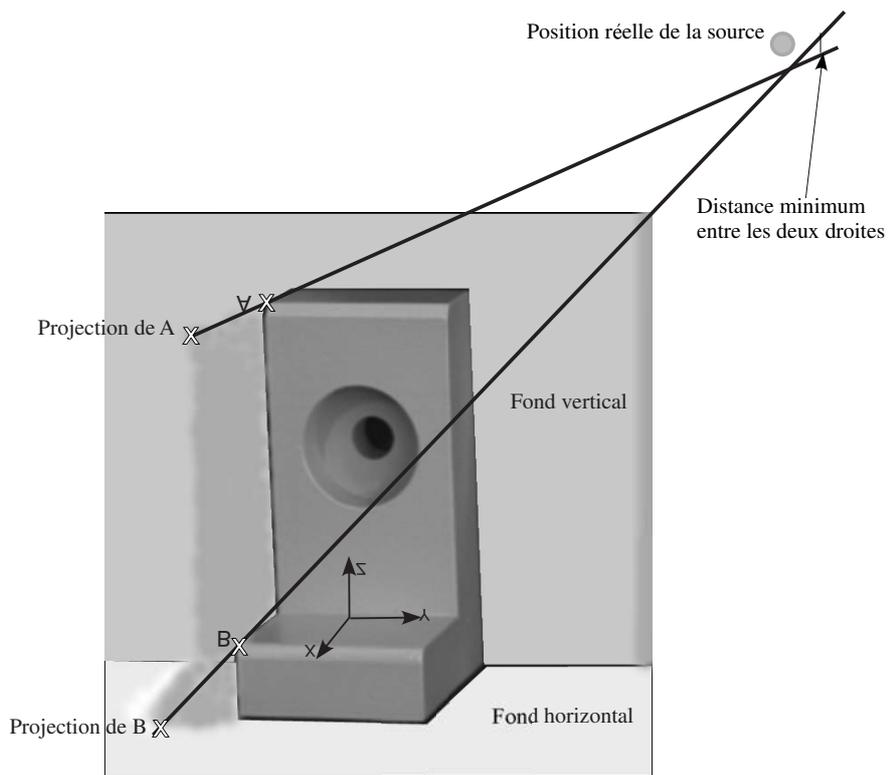


FIG. 3.29 – La position de la source lumineuse peut être obtenue en calculant l'intersection des deux droites passant respectivement par A et la projection de A, et B et la projection de B

3.3.5 Conclusion de la modélisation de l'illumination

L'étude des distributions de différents paramètres peut ainsi être menée à la fois pour la pièce et pour son environnement. Les résultats simulés avec la pièce en « L » permettent de définir des traitements d'images locaux. En utilisant de façon simultanée les distributions sur le fond et le socle, il devient également possible, pendant l'exécution de l'application en ligne, de repositionner la source lumineuse en fonction d'effets spécifiques à faire apparaître ou disparaître dans les images réelles. Ces premières études sur la pertinence des paramètres liés aux phénomènes d'illumination n'en sont qu'à leurs débuts. En effet, seules les sources ponctuelles ont été abordées. L'éclairage et la luminance sont les seuls paramètres dont les distributions ont été étudiées. Or, dans les systèmes d'inspection, les sources ponctuelles ne sont pas parmi les plus utilisées. D'autre part, il est vraisemblable que d'autres paramètres, notamment des paramètres directement calculés à partir des images réelles, puissent aussi être pertinents dans la maîtrise de la position et du réglage d'une source lumineuse. Le problème crucial est enfin de pouvoir tenir compte de plusieurs sources lumineuses. La maîtrise des paramètres liés à chaque source semble, pour le moment encore, un problème relativement complexe à résoudre.

3.4 Conclusion

Dans ce chapitre dédié à la modélisation CAO et à la modélisation de l'illumination, nous avons exposé l'intérêt d'un modèle CAO capable de fournir la géométrie idéale de l'objet à étudier. En utilisant l'outil CATIA V5, il est possible de déterminer les primitives communes (segments de droite et segments curvilignes) dans des images obtenues avec un système d'acquisition stéréoscopique. En effet, en réalisant une modélisation complète de la scène, des images conceptuelles ont été générées et ont conduit à l'obtention d'images filaires. Seules sont conservées les primitives de type segment de droite et segments curvilignes visibles du point de vue du capteur stéréoscopique. Ces primitives 3D peuvent alors être projetées sur le plan image 2D et être comparées aux primitives contenues dans les images réelles.

La modélisation de phénomènes d'illumination va contribuer à l'adaptation de traitements d'images locaux et à l'amélioration des conditions de prises de vue. De premiers résultats encourageants conduisent à de nombreuses pistes à explorer pour parvenir à maîtriser au mieux les paramètres d'illumination. À terme, tous ces mécanismes d'évaluation de l'information extraite d'une image seront intégrés dans un processus capable de choisir la technique de modélisation la plus adéquate et les sources de lumière les plus appropriées, compte tenu de l'application et du type des objets à évaluer. Pour être opérationnel, notre système de mesure dimensionnelle permettant la reconstruction tridimensionnelle par stéréovision de pièces manufacturées doit encore recourir à une planification. L'outil de planification utilisé repose sur les SGT (Situation Graph Tree) ou arbres de graphes de situations. Le développement de cet outil fait l'objet du chapitre suivant.

Chapitre 4

Automatisation et mise en place d'un système d'inspection

4.1 Introduction

Dans le cadre de l'inspection dimensionnelle, un système de contrôle qualité assurant une inspection métrologique automatique de pièces manufacturées est développé dans notre laboratoire. Ce système est composé de deux modules : le premier utilise les graphes de situations qui permettent l'automatisation et la gestion des différentes procédures de la reconstruction 3D. Le deuxième, appelé HTN, ou Hierarchical Task Plan en anglais, se charge du contrôle de la séquence d'acquisitions ainsi que de la reconstruction 3D complète de l'objet et de son évaluation dimensionnelle. Le schéma de la figure (4.1) illustre notre système de mesure par vision, en vue d'un contrôle qualité.

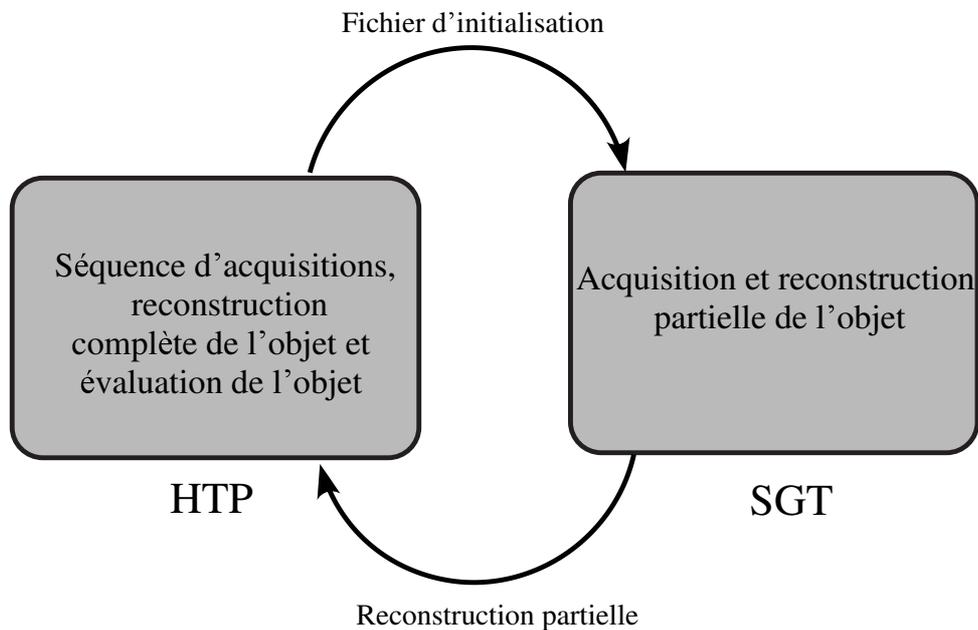


FIG. 4.1 – Modules composant le système de planification et de contrôle d'une mesure dimensionnelle par vision

L'articulation de l'ensemble des outils présentés et développés dans les chapitres précédents va se faire autour des graphes de situations, en développant un graphe de situations (SGT, Situation Graph Tree, en anglais) spécifique à l'automatisation de la reconstruction 3D partielle par vision stéréoscopique. Une caractéristique de ces graphes intéressante à exploiter est qu'ils sont capables de s'adapter dynamiquement aux conditions réelles de prise de vue (effets d'illumination, position de la tête stéréoscopique) et ainsi de modifier automatiquement les conditions effectives d'acquisition et de traitement de données si nécessaire.

De ce fait, nous avons développé une approche permettant une replanification dynamique adaptée au type de problème rencontré en utilisant le système de contrôle. L'intérêt réside dans la possibilité de réagir à des actions qui ne sont pas exécutées comme prévu par le planificateur. Sur cette base, nous avons développé des règles de décision permettant au système de planification d'adapter les traitements en cours d'exécution. Ces règles

se basent en premier lieu sur l'ajustement de l'éclairage de la scène et, dans un second temps, sur le déplacement de la tête de mesure. Les modules développés et intégrés dans la chaîne de traitement sont capables de fournir une reconstruction partielle élaborée des objets à évaluer. Cette reconstruction doit ensuite être comparée avec le modèle CAO correspondant, pour décider de la qualité de l'objet.

Nous avons, jusqu'ici, décrit brièvement dans la partie introductive notre système de planification dédié à la reconstruction des pièces industrielles. À présent, nous allons analyser en détail le comportement du système ainsi que les modules développés constituant ce dernier.

Ce chapitre est organisé comme suit : nous introduisons en premier lieu quelques généralités sur l'inspection automatique. Dans un second temps, nous étudions les graphes de situations tout en présentant leur contexte classique d'utilisation ainsi que leur adaptation à notre application de contrôle qualité. Nous présentons ensuite l'analyse d'une reconstruction partielle contrôlée par les SGT en présentant la chaîne de traitement dans sa globalité. La replanification dynamique des tâches de traitement est exposée en présentant les règles de décision permettant d'adapter notre SGT aux résultats effectifs des traitements ainsi qu'aux conditions réelles d'acquisition. Enfin, une description du dispositif expérimental est donnée, avec des résultats expérimentaux se rapportant à ce travail.

4.2 Inspection automatique

Parmi les applications possibles de la vision par ordinateur, la majorité des applications porte sur l'inspection de pièces industrielles. Plus précisément, dans notre cas, disposant du modèle CAO de l'objet à inspecter, ainsi que d'un ensemble de points 3D décrivant la pièce à évaluer, il est possible par mise en correspondance des deux jeux de données de vérifier la conformité de la pièce usinée et cela en tenant compte des tolérances. Les différentes pièces qui peuvent être inspectées sont classées généralement en deux catégories en fonction de la complexité de leurs formes [Mor96]. La première catégorie concerne les objets de forme simple comprenant des surfaces planes ou cylindriques. La deuxième catégorie, quant à elle, comprend les objets de forme complexe comportant des surfaces gauches et peu de surfaces planes. Ceux-ci sont en général difficilement mesurables par les techniques traditionnelles. La validité de la procédure d'inspection automatique passe par une connaissance précise des erreurs induites par l'incertitude de mesure du capteur utilisé d'une part et, d'autre part, par la mise en correspondance entre le modèle et les points 3D décrivant la pièce à inspecter. Ces erreurs doivent être impérativement négligeables ou dans les limites de tolérance pour les dimensions contrôlées pendant l'opération d'inspection.

Par définition, l'opération d'inspection dimensionnelle consiste à déterminer la conformité d'un produit usiné, tout en vérifiant s'il dévie d'une liste de spécifications données. Ainsi, l'inspection implique la mesure de certaines caractéristiques telles que les dimensions géométriques, la finition des surfaces, etc. L'inspection non visuelle réalisée, par exemple, par l'utilisation de capteurs télémétriques ou de systèmes de mesures 3D de type scanner est très lente et demande une program-

4.3 Graphes de situations

mation détaillée pour chaque objet traité. Son avantage est d'aboutir à des informations 3D de grande précision. L'inspection visuelle permet de contrôler les pièces industrielles sans les manipuler. Un avantage peut aussi résider dans le fait que la tâche d'inspection ne nécessite pas l'arrêt de la chaîne de fabrication. Cette inspection peut être qualitative (accepter ou rejeter une pièce) ou bien quantitative (classer un type de défauts constaté). Le contrôle de qualité est souvent effectuée par des opérateurs humains qui peuvent s'aider de dispositifs de mesure. Mais ces opérations restent des tâches répétitives, difficiles et fastidieuses. De ce fait, l'opérateur humain ne peut pas toujours suivre la cadence de production et l'inspection ne porte souvent que sur des échantillons de la production et ne représente par conséquent qu'une valeur statistique.

L'inspection visuelle automatisée comble en partie les lacunes d'un évaluateur humain, dont l'aptitude à résister à la lassitude due au caractère répétitif de l'inspection et l'efficacité ne sont pas optimales. De plus, certains environnements sont hostiles à l'intervention humaine et, par ailleurs, certains défauts ne sont pas détectables par l'œil humain. L'automatisation de l'inspection industrielle, à l'aide de la vision par ordinateur, par exemple, est ainsi apparue comme un moyen d'obtenir un contrôle de qualité fiable de la production. Un autre avantage de l'inspection visuelle automatisée est la faisabilité de l'inspection en temps réel, dans certains cas.

L'utilisation de données CAO est d'un grand intérêt pour l'inspection. Le modèle CAO utilisé pour une tâche d'inspection est considéré comme étant une description idéale de l'objet. En effet, les données CAO représentent une description analytique de la forme de l'objet incluant souvent une paramétrisation explicite de la forme des surfaces. À partir d'une modélisation explicite des spécifications de tolérance associées aux éléments géométriques de pièces quasi-polyédriques, Zhou [Zho00] a implémenté une approche pour la comparaison entre images réelles et représentations conceptuelles. Cette approche offre la possibilité d'évaluer quantitativement des pièces manufacturées avec précision et de vérifier si des tolérances prédéfinies sont respectées. Elle repose sur une comparaison entre listes de points de contour extraites de l'image réelle et les zones d'intérêt correspondantes modélisant les tolérances. En effet, l'exploitation de l'approche développée par Zhou, permet de remonter à des données quantitatives caractérisant la pièce à évaluer.

Après avoir abordé l'apport de l'utilisation des modèles CAO, nous allons introduire un bref état de l'art sur les graphes de situations, que nous avons utilisés pour l'automatisation de notre système d'inspection.

4.3 Graphes de situations

Notre système de mesure se charge de la gestion de l'ensemble des différentes tâches de traitement, à savoir l'acquisition, l'évaluation et les mesures dimensionnelles. Dans le cadre de cette thèse, nous nous sommes ainsi intéressés à l'utilisation des SGT (Situation Graph Trees) pour planifier et automatiser les différentes tâches de traitement nécessaires à une reconstruction tridimensionnelle partielle par stéréovision.

4.3.1 Éléments de base constituant un SGT

Un SGT [AN02, AON02] décrit le comportement d'un agent en termes de situations dans lesquelles ce dernier peut se trouver. Un schéma de situation (« Situation Scheme ») est constitué de deux éléments de base : un ensemble de prédicats d'état (« State Scheme ») décrivant à l'aide de prédicats logiques l'état (binaire ou flou) de l'agent, ainsi que celui de son environnement, et un schéma d'actions (« Action Scheme ») décrivant l'ensemble des actions que l'agent doit exécuter. Si l'ensemble des prédicats d'état est satisfait, l'agent instancie le schéma de situation correspondant et toutes ses actions sont exécutées (voir figure (4.2)). On peut relier les schémas de situation par des liens de prédiction (« Prediction

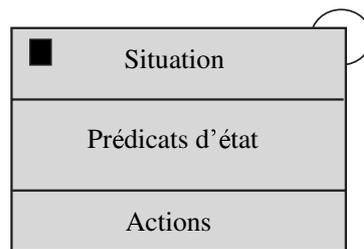


FIG. 4.2 – Élément de base d'un SGT : schéma de situation

Edges ») orientés [HN00]. Un schéma de situation peut posséder un lien d'auto-prédiction (« Self Prediction »). Un agent peut donc rester dans une même situation pendant plusieurs pas temporels. On appelle graphe de situations (« Situation Graph »), un ensemble de schémas de situation connectés [Nag99] (voir figure (4.3)). Certains schémas de situation peuvent être marqués comme étant des situations de début (« Start Situation ») constituant des points d'entrée du graphe ou bien comme étant des situations de fin (« End Situation ») constituant des points de sortie du graphe.

Une séquence de situations est ensuite définie par le chemin reliant une situation de début à une autre de fin. Nous pouvons ainsi obtenir des séquences de situations possibles. Mais, pour ces séquences, seul un schéma de situation est activé à un instant donné. Afin d'affiner un schéma de situation, il est possible de le connecter à un ou plusieurs graphes de situations par le biais d'un lien de spécialisation (« Specialization Edges ») [Nag99]. L'ensemble résultant constitue un arbre de graphes de situations (SGT). Il est possible dans un graphe de situations d'avoir des spécialisations multiples.

4.3.2 Contexte du développement et d'utilisation classique des SGT

La conception des SGT a été faite à l'Institut d'Algorithmique et des Systèmes Cognitifs à l'Université de Karlsruhe en Allemagne [AN03]. L'objectif initial était de construire automatiquement une description textuelle du contenu d'une séquence vidéo ; il s'agissait par exemple de décrire l'ensemble des actions effectuées par un véhicule (jouant le rôle d'agent) dans une scène urbaine. L'ensemble des situations dans lesquelles l'agent est susceptible de se trouver est ainsi décrit par plusieurs schémas de situation connectés pour former un graphe de situations.

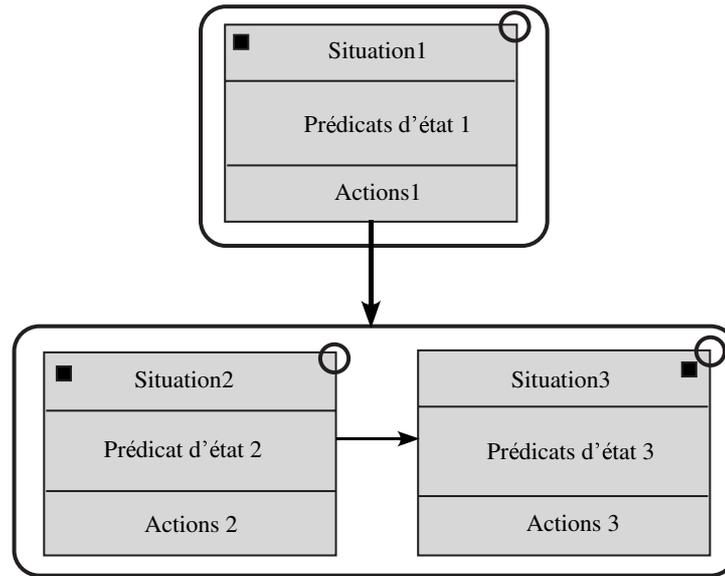


FIG. 4.3 – SGT (graphe de situations) constitué de schémas de situation

Les prédicats d'état des différents schémas de situation dépendent uniquement du contenu de l'image courante de la séquence vidéo et leur valeur est fournie par un processus de traitement d'images (p. ex., les coordonnées de l'agent sont (x_0, y_0)), et la seule action à exécuter consiste à afficher le nom et la description de la situation instanciée [Nag99].

4.3.3 SGT adapté à notre application : reconstruction tridimensionnelle

Les SGT ont été adaptés par notre équipe à notre application d'inspection dimensionnelle. Chaque situation représente une étape bien définie du processus d'acquisition et de reconstruction 3D. Les actions exécutées lors de l'instanciation d'une situation sont susceptibles de modifier les prédicats d'états d'autres situations. Le graphe conduit donc à plusieurs séquences de situations possibles.

Le SGT est composé de schémas de situation permettant de décrire toutes les situations que le capteur stéréoscopique est en mesure de rencontrer : déplacement, acquisition, traitement et reconstruction 3D partielle.

Un niveau d'abstraction différent peut être atteint en spécialisant une situation par un lien de spécialisation.

Le parcours du graphe consiste à déterminer la séquence de situations qu'un agent (dans notre application, il s'agit des têtes de mesure) est amené à instancier. Les schémas de situation peuvent être successivement instanciés, si l'ensemble des prédicats d'état du schéma de situation instancié est vérifié et si ce schéma de situation est connecté au schéma de situation actif à l'étape précédente. Il faut aussi que le schéma de situation précédemment actif ait exécuté avec succès l'ensemble de ses actions.

Le parcours d'un SGT repose donc sur l'évaluation des prédicats d'état. La valeur de ces prédicats peut être modifiée par certaines actions d'autres schémas de situation (p. ex., par

le biais de figures de mérite obtenues à la suite d'un traitement d'images), mais elle peut également être dépendante de l'état d'un capteur. Il peut ainsi être nécessaire de modifier la valeur de certains paramètres de traitement, si les figures de mérite prédéfinies ne sont pas réalisées (p. ex., à la suite d'un échec de la segmentation d'une image). Cette correction, automatique, peut également amener le processus à ajuster les paramètres physiques de l'application (p. ex., en modifiant les conditions d'illumination ou bien en ajustant la position de la tête de mesure). Notre SGT est ainsi capable d'adapter dynamiquement son comportement en fonction des conditions réelles d'acquisition et des résultats effectifs des différentes procédures de traitement déjà exécutées ou en cours d'exécution.

Nous avons abordé l'utilisation classique des SGT et nous venons de voir que ces derniers peuvent être adaptés à une application d'inspection dimensionnelle et de contrôle qualité, en permettant une planification d'une séquence de reconstruction 3D de pièces industrielles à inspecter. À présent, nous allons présenter l'utilisation des graphes de situations pour définir une planification et l'automatisation d'une procédure de reconstruction tridimensionnelle par stéréovision.

4.4 Modélisation du comportement du capteur par les SGT

Dans le cadre de notre contribution au développement du système de planification, nous avons utilisé les SGT comme outil et notre tâche a été de développer un graphe adapté à une reconstruction partielle par stéréovision. En effet, les SGT ont été d'abord développés, dans notre équipe, pour une reconstruction partielle par lumière structurée [KLH05].

Notre SGT exécute un ensemble d'opérations permettant d'obtenir une reconstruction partielle de l'objet et correspondant au point de vue effectif du capteur. Cette description partielle de la pièce, sous la forme de primitives 3D quantifiées, doit être évaluée en faisant appel aux primitives prévues idéales obtenues à partir du modèle CAO.

La modélisation du comportement de la tête de mesure consiste à prévoir *a priori* toutes les séquences de situations possibles pour les traitements. Ces parcours peuvent être schématisés par un graphe comportant des situations appropriées à chaque étape de traitement. Le graphe permet donc de déterminer la séquence de situations que la tête de mesure est amenée à instancier effectivement.

Notre SGT a été aussi développé pour synthétiser toutes les configurations possibles que le capteur stéréoscopique peut rencontrer. On modélise, pour ce faire, les étapes permettant une reconstruction de l'objet par des schémas de situation spécifiques. Les situations contiennent un ensemble de prédicats d'état et un ensemble d'actions. Les conditions d'activation d'une situation sont spécifiées par les prédicats d'état, tandis que les actions représentent les traitements prêts à être exécutés si la situation est activée.

L'exécution des traitements est spécifiée par les actions des schémas de situation. Toute action exécutée peut modifier les prédicats d'états d'autres situations et, par conséquent, le parcours du graphe dépend des résultats des situations déjà instanciées. L'enchaînement

4.5 Reconstruction 3D guidée par les SGT : tête de mesure stéréoscopique

des différentes procédures de traitement pour aboutir à une reconstruction 3D partielle est donné dans la figure (4.4).

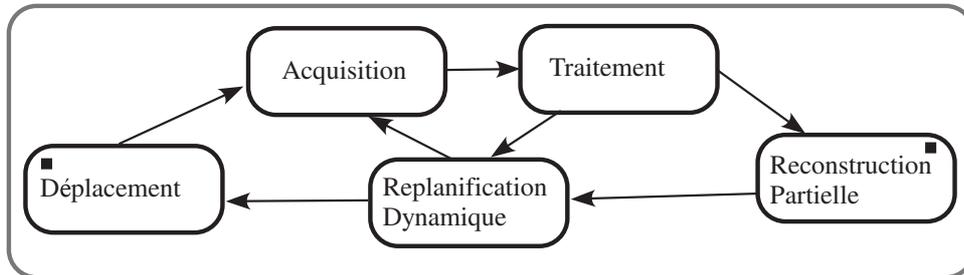


FIG. 4.4 – Parcours du SGT développé pour une reconstruction partielle

Comme nous l'avons vu dans la section sur l'état de l'art des SGT, ceux-ci décrivent le comportement d'un agent en termes de situations dans lesquelles ce dernier peut se trouver. Dans notre application, l'agent est représenté par la tête de mesure (capteur). Un SGT est donc capable de décrire la connaissance comportementale de notre système de vision. Chaque étape du processus représenté figure (4.4) est modélisée par une situation. Pour une tâche de reconstruction partielle de l'objet, des procédures d'acquisition et de traitement de données sont nécessaires. Pour coordonner les différentes tâches de traitement, nous avons été amenés à définir un graphe de situations bien précis.

4.5 Reconstruction 3D guidée par les SGT : tête de mesure stéréoscopique

Pour valider notre approche, nous avons conçu un graphe de situations synthétisant toutes les configurations que la tête de mesure stéréoscopique peut rencontrer lorsqu'elle élabore une reconstruction partielle de l'objet. Toutes les étapes du processus d'acquisition et de reconstruction 3D sont modélisées par des schémas de situation.

Le graphe de contrôle HTP permet la conception d'une séquence d'acquisitions en vue de la reconstruction complète d'un objet. Cette conception détermine à la fois le nombre et l'ordre des acquisitions à réaliser, la tête de mesure la plus adaptée au type de primitive recherchée, les paramètres de l'éclairage de la scène ainsi que les paramètres des différents procédures de traitement d'images. Le graphe de contrôle se charge aussi de la gestion de la reconstruction 3D ; il permet ainsi le recalage successif des reconstructions partielles réalisées par les SGT pour obtenir une reconstruction complète de l'objet.

L'ensemble des opérations effectuées pour réaliser une reconstruction partielle de l'objet guidée par les SGT est regroupé dans deux classes, à savoir les traitements hors ligne et les traitements en ligne. Les traitements hors ligne sont effectués par le graphe de contrôle et les traitements en ligne sont les traitements réalisés par le SGT en cours d'exécution (voir figure 4.1).

4.5.1 Traitements effectués hors ligne

Une phase de modélisation hors ligne s'avère être nécessaire pour prévoir et définir toutes les possibilités de traitement que la tête de mesure peut avoir à effectuer lors de l'instanciation en ligne du graphe de situations. Ce travail permet de construire un fichier contenant les différents paramètres estimés *a priori* et nécessaires à la reconstruction de la pièce manufacturée. Ses paramètres vont être fournis au SGT pour lancer les différents traitements liés à la reconstruction. Ces paramètres sont les suivants :

Paramètres d'étalonnage du capteur

C'est un module que nous avons développé et qui exploite la méthode d'étalonnage de Tsai [Tsa86]. Il permet de modéliser les deux caméras composant notre tête de mesure stéréoscopique. Cet étalonnage est effectué hors ligne et permet d'obtenir les paramètres intrinsèques et extrinsèques des caméras formant la tête de mesure stéréoscopique. La figure (4.5) illustre un exemple de fichier contenant les paramètres d'étalonnage d'une caméra.

```

/*-----*/
/*
/*           Calibration Results
/*-----*/
/* Number of calibration points : 99
/*-----*/
/* Rotation Matrix :
/* r1 = -0.023165 r2 = -0.959744 r3 = -0.279921 */
/* r4 = -0.970523 r5 = -0.045599 r6 = +0.236656 */
/* r7 = -0.239893 r8 = +0.277152 r9 = -0.930397 */
/*-----*/
/* Translation Matrix (mm) :
/* Tx = -1.435 Ty = +4.901 Tz = +463.772 */
/*-----*/
/* Focal length (mm) : 37.388
/*-----*/
/* Radial lens distortion : +0.002517
/*-----*/
/* CCD chip, image acquisition :
/* Sx = 0.011000 Sy = 0.011000 */
/* Cx = 256.000 Cy = 256.000
/*-----*/
/* Minimal error distance : 0.112 pixels 0.016 mm */
/* Maximal error distance : 3.235 pixels 0.468 mm */
/* Average error distance : 0.992 pixels 0.141 mm */
/*-----*/
/* p1 = -0.302
/* p2 = -6.881
/* p3 = -2.565
/* p4 = +245.485
/* p5 = -7.245
/* p6 = -0.181
/* p7 = -1.221
/*-----*/

```

FIG. 4.5 – Exemple de fichier contenant les paramètres d'étalonnage estimés par la méthode de Tsai

Primitives prédites CATIA

Les primitives prédites vont être utilisées comme information *a priori* pour évaluer la segmentation des images 2D. Ces primitives 2D prédites résultent de la conception

4.5 Reconstruction 3D guidée par les SGT : tête de mesure stéréoscopique

du modèle de la pièce avec le logiciel CATIA et sont obtenues en introduisant les paramètres d'étalonnage (voir chapitre 3, section 3.2.4). Le modèle CAO est aussi utilisé pour sélectionner les primitives communes aux deux images stéréoscopiques.

Position de la tête de mesure

Chaque position particulière de la tête de mesure permet de réaliser une reconstruction partielle correspondant à ce champ de vision. Pour une reconstruction complète de la pièce, nous pouvons réaliser plusieurs déplacements de la tête de mesure autour de l'objet. Pour chaque position, l'estimation des paramètres extrinsèques des caméras est indispensable pour effectuer la reconstruction 3D à partir des images acquises.

Par ailleurs, un module spécifique a été développé dans l'équipe pour réaliser le recalage entre les reconstructions partielles obtenues pour chaque parcours du SGT, pour obtenir une reconstruction complète de l'objet.

Paramètres d'illumination

Un réglage des conditions d'éclairage préalablement aux acquisitions des images s'avère nécessaire. Ce réglage porte sur le choix du type d'éclairage (source ponctuelle, rectangulaire, surfacique, etc.), ainsi que sur le réglage de l'intensité lumineuse. En définissant les modifications à apporter au système d'illumination, nous pouvons améliorer la qualité des images acquises. Pour cela, il est indispensable d'être capable de modéliser les effets lumineux (p. ex., des surintensités) observés lorsque les paramètres de la source lumineuse (position, intensité, etc.) ne sont pas optimisés, compte tenu de l'objet particulier à reconstruire.

Règles de décisions pour replanifier les traitements en cours d'exécution

Ces règles ont été développées pour corriger un traitement incohérent. Elles interviennent lors de l'appel de la situation « replanification dynamique ». Ces règles définies *a priori* sont abordées en détail dans la partie portant sur la replanification dynamique.

4.5.2 Traitements effectués en ligne

Cette phase de traitement implique tous les programmes opérationnels que nous avons développés. Chacun de ces programmes est liée à une action particulière d'un schéma de situation du SGT (voir figure (4.4)). Ces procédures sont :

- **Contrôle du contraste de l'image.** C'est un programme permettant de valider le contraste de l'image acquise. Ce traitement est indispensable pour vérifier la qualité de l'image acquise, en termes d'intensité des sources d'éclairage. Une image sombre ou saturée n'autorise souvent pas une bonne extraction des indices visuels.
- **Segmentation et classification des contours des images acquises.** C'est la procédure nécessaire à l'extraction des primitives des images (contours). Ces primitives vont être classées selon leur forme géométrique pour faciliter le travail de la procédure d'appariement entre contours extraits du couple d'images stéréoscopiques.

- **Élimination des primitives non communes dans le couple d'images stéréoscopiques.** Les images d'un couple stéréoscopique ne contiennent pas forcément les mêmes contours. De ce fait, les contours non perçus simultanément dans les deux images n'ont pas de correspondant. Par conséquent, ils doivent être supprimés avant d'exécuter l'algorithme d'appariement. Cette procédure de traitement a été présentée en détails dans le deuxième et le troisième chapitres du présent mémoire.
- **Appariement de contours.** C'est l'algorithme permettant d'apparier les contours des images gauche et droite acquises par les deux caméras. L'approche a été développée dans le chapitre 2 de ce mémoire.
- **Calcul des primitives 3D.** C'est la dernière étape de notre chaîne de traitement. Le programme approprié permet d'aboutir à une information 3D liée aux primitives appariées auparavant.

4.6 Automatisation d'une reconstruction 3D par stéréovision et replanification dynamique utilisant les SGT

Les différentes situations formant notre SGT sont spécialisées, mise à part la première situation appelée « déplacement », qui constitue la situation de début. Les autres situations exécutent un enchaînement de procédures permettant d'obtenir des primitives 3D de la pièce à inspecter. Elles sont appelées respectivement : acquisition, traitement, reconstruction partielle et *replanification dynamique*. Dans la figure 4.6 nous avons schématisé notre graphe de situations.

Dans ce qui suit, nous présentons la description de chaque situation, ainsi que sa spécialisation. Les situations de début et de fin d'un graphe sont mentionnées sur les figures par un carré situé respectivement à gauche et à droite de la situation concernée.

1. **Déplacement.** Cette situation représente l'état initial du système et permet de fixer la position de la tête de mesure. La reconstruction partielle sera effectuée pour ce point de vue. Dans notre cas, cette situation n'est pas spécialisée.
2. **Acquisition.** Cette situation permet au système de faire l'acquisition des images par l'intermédiaire de la tête stéréoscopique. La position du capteur stéréoscopique a été définie préalablement et est contrôlée par la situation « déplacement ». Nous avons spécialisé la situation « acquisition » à l'aide d'un sous-graphe comportant trois situations : *Test_image*, *Ajustement_intensité* et *Validation_contraste*.
 - **Test_image.** Cette situation permet de vérifier la qualité des images acquises avant de lancer l'exécution des différents traitements de la reconstruction. Cette étape va aider le processus à éliminer au mieux tous les artefacts dus aux conditions d'acquisition. Cette situation se charge de lancer une procédure sous la forme d'un fichier exécutable ayant comme entrée une image et fournit en sortie un paramètre permettant de décider de la qualité de l'image (image sombre, saturée ou équilibrée). Dans le cas où la situation *Test_image* donne en sortie la valeur de paramètre liée à une image sombre ou saturée, un ajustement de l'intensité est déterminé pour corriger la qualité de l'image. Dans le cas d'une image évaluée comme équilibrée, le contraste est validé et on peut instancier la situation

4.6 Automatisation d'une reconstruction 3D par stéréovision et replanification dynamique utilisant les SGT

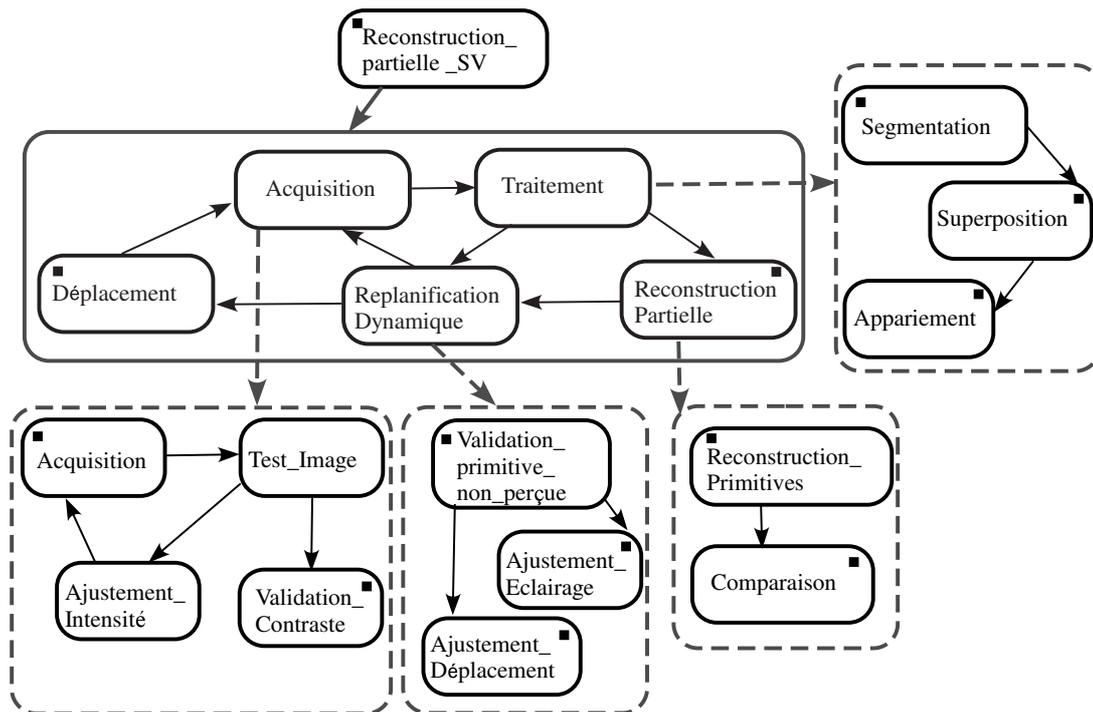


FIG. 4.6 – Analyse d'une reconstruction partielle et replanification dynamique par les SGT . La figure indique aussi les spécialisations

suivante. Ce traitement est global et se fait sur toute l'image.

- **Ajustement intensité.** Dans cette situation, la valeur du prédicat d'état est fournie par la situation Test_image. En fonction de ce paramètre (sombre/saturé), on augmente/diminue l'intensité de l'éclairage. Une fois l'intensité ajustée, l'acquisition des images est refaite. La situation suivante va nous permettre de valider ou non les nouvelles images acquises.
 - **Validation contraste.** C'est une situation de fin pour le graphe de spécialisation de l'acquisition. Cette situation permet de quitter cette spécialisation pour continuer le parcours générique du graphe source.
3. **Traitement.** Cette situation contrôle les différents traitements de la méthode d'appariement développée dans le chapitre 2. Elle récupère en entrée le couple d'images stéréo-scopiques et donne comme résultat les contours des deux images appariés correctement. Le traitement est spécialisé par trois situations : *Segmentation*, *Superposition* et *Appariement*. Les actions des différentes situations sont disponibles sous la forme de fichiers exécutables. Nous allons maintenant expliciter le traitement correspondant à chaque situation séparément (voir figure (4.7)).
- **Segmentation.** C'est une situation de début qui reçoit en entrée les images acquises. Celles-ci valident les prédicats et permettent de lancer la procédure de segmentation qui ne conserve que les contours valides du couple d'images stéréo-scopiques. Ces contours vont être divisés et classés selon leur forme géométrique. En sortie, nous obtenons des images résultats contenant des segments de contour (voir chapitre 2).

- **Superposition.** Cette situation reçoit comme paramètres pour les prédicats d'état les contours extraits et classés par la situation « segmentation ». Un deuxième groupe de paramètres pour les prédicats d'état est donné par les images « fil de fer » générées hors ligne. Dans cette situation, une superposition des contours réels avec les contours conceptuels « fil de fer » est faite pour accentuer ceux qui vont être appariés et reconstruits. Cette situation exploite l'information *a priori* fournie par les images « fil de fer » construites avec les mêmes conditions de prise de vue (points de vue des caméras et réglage de l'éclairage). Cette situation peut être une situation de fin, si on aboutit à une mauvaise superposition (p. ex., des primitives réelles sont non perçues). Dans ce cas, le SGT fait appel à une replanification dynamique permettant de corriger les conditions d'acquisition des images réelles. Dans le cas contraire, c'est-à-dire lorsque nous pouvons identifier, à partir des primitives prédites, toutes les primitives réelles correspondantes, celles-ci vont être utilisées par la prochaine situation.

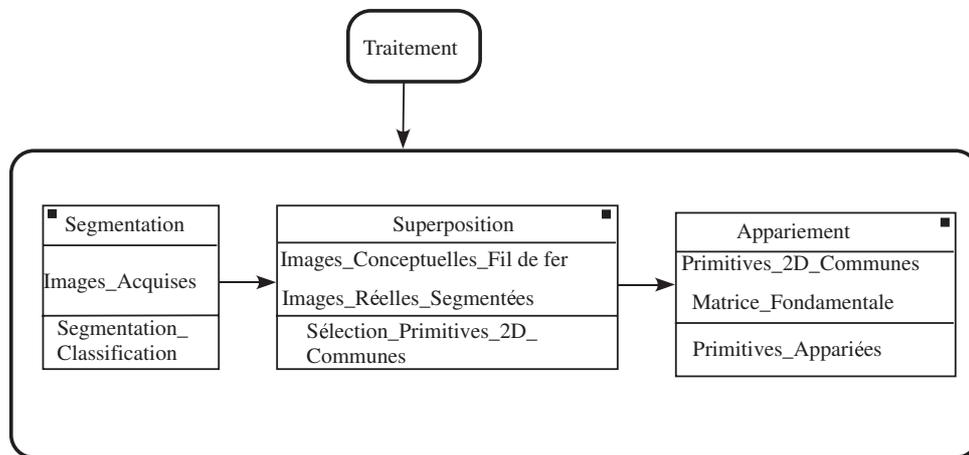


FIG. 4.7 – Spécialisation de la situation traitement

- **Appariement.** C'est une situation de fin pour la spécialisation du traitement, qui reçoit en entrée les primitives 2D communes aux deux images réelles. En utilisant la matrice fondamentale obtenue hors ligne comme paramètre d'un prédicat d'état, nous pouvons lancer la procédure d'appariement entre les primitives des deux images. Cette situation comprend tous les programmes développés pour la méthode d'appariement exposée dans le deuxième chapitre du présent mémoire. Une fois la situation « appariement » instanciée après vérification des prédicats d'état à partir des primitives 2D communes et de la matrice fondamentale estimée *a priori*, le schéma lance les actions en exécutant le programme approprié qui génère en sortie des paires appariées de primitives de type contour.
4. **Reconstruction partielle.** Nous avons spécialisé cette situation par deux situations : *Reconstruction_Primitives* et *Comparaison*. La situation « reconstruction partielle » reçoit en entrée les primitives 2D appariées et fournit en sortie des primitives 3D reconstruites. Cette situation est une situation de fin, parce que l'on conserve la dernière reconstruction partielle obtenue au bout d'un maximum de n replanifications dynamiques (voir la situation 5 ; voir figure (4.8)).

4.6 Automatisation d'une reconstruction 3D par stéréovision et replanification dynamique utilisant les SGT

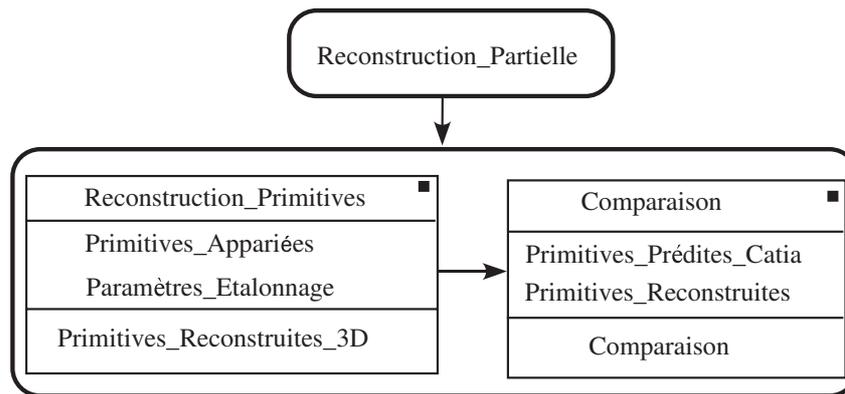


FIG. 4.8 – Spécialisation de la situation reconstruction partielle

- **Reconstruction_primitives.** C'est la situation représentant le schéma de début de cette spécialisation. Elle évalue ses prédicats d'état à partir des paramètres d'étalonnage estimés hors ligne et des primitives 2D appariées résultant de la situation « traitement ». Une fois que cette situation a été instanciée, nous obtenons en sortie des primitives contours reconstruites dans l'espace 3D.
 - **Comparaison.** Cette situation va permettre d'évaluer la reconstruction partielle de l'objet selon le point de vue utilisé. La comparaison repose sur une théorie exploitant des tolérances géométriques permettant la construction du modèle de tolérances 3D de la pièce à inspecter [Zho00]. Ce module de comparaison entre primitives 3D reconstruites et primitives prédites incluant des zones de tolérance fait partie des travaux développés auparavant dans notre laboratoire. L'instanciation de cette situation va permettre d'exécuter le module de comparaison. Si les primitives 3D reconstruites sont incluses dans les zones de tolérance du modèle 3D, la reconstruction est dite réussie. Dans le cas contraire, une replanification dynamique est appelée pour déplacer la tête de mesure. Si, avec ce déplacement, on aboutit encore à des primitives reconstruites non incluses dans les zones de tolérance, alors la pièce est considérée comme défectueuse.
5. **Replanification dynamique des conditions d'acquisition.** En cours d'exécution, le SGT adapte dynamiquement les paramètres physiques de l'application si nécessaire. Selon les conditions réelles d'acquisition et les résultats effectifs des procédures de traitement en cours, une correction peut s'avérer nécessaire afin d'améliorer les conditions d'éclairage ou de modifier la position de la tête de mesure. La replanification dynamique est appelée si les traitements globaux (p. ex., l'ajustement d'un paramètre de segmentation, l'ajustement de l'intensité d'éclairage) utilisés pour la correction d'une anomalie n'améliorent pas le résultat. L'appel à la situation « replanification dynamique » favorise les traitements locaux qui permettront de traiter les zones mal analysées de l'objet dans l'image, parce que les conditions d'acquisition (position de la tête de mesure et/ou des sources lumineuses) n'étaient pas optimales pour les analyser. La situation « replanification dynamique » prend en compte le résultat du dernier traitement et évalue un ensemble de règles de décision définies *a priori*. Ces règles vont permettre au graphe de choisir le ou les paramètres physiques qui vont être ajustés. Pour des raisons de simplicité, nous avons développé ces

règles de décision en favorisant une action sur les paramètres de l'éclairage, avant de modifier ceux du capteur. La replanification dynamique est une approche innovante dans ce système de planification de la reconstruction 3D dédié au contrôle qualité de pièces polyédriques. Cette situation est appelée pour donner au système d'inspection un comportement dynamique (voir figure 4.10). Par un appel à cette situation, nous pouvons intervenir au niveau des différentes situations constituant notre graphe afin d'améliorer la qualité de la reconstruction partielle. L'ajustement se fait par l'appel d'un ensemble de règles appelées règles de décision. À titre d'exemple, dans notre implémentation actuelle, la replanification dynamique est spécialisée par trois situations : *Règles_de_décision*, *Ajustement_éclairage* et *Ajustement_déplacement* (voir figure (4.9)).

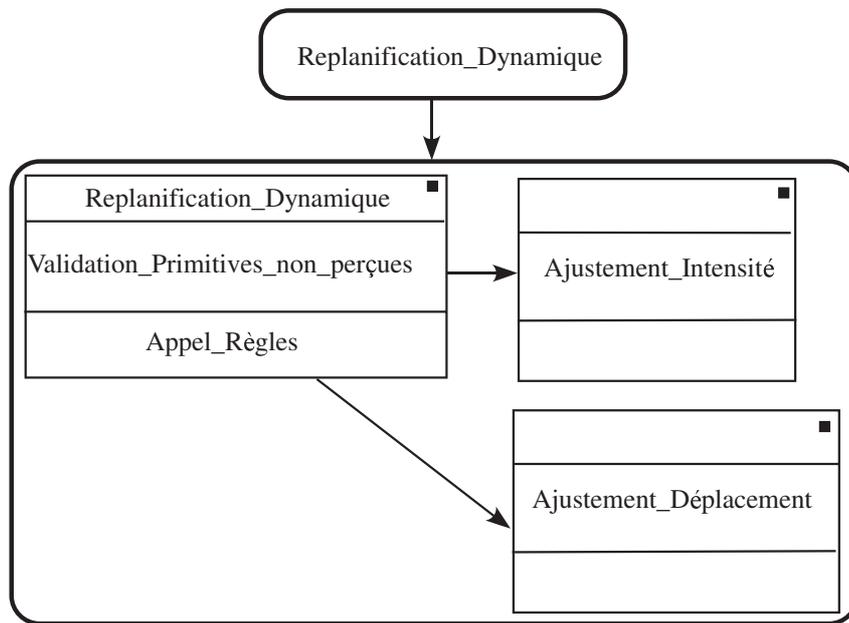


FIG. 4.9 – Spécialisation de la replanification dynamique

- **Règles de décision.** Nous avons développé pour la replanification dynamique trois règles de décision. Ces règles, du type *Si/Alors*, vont guider grandement la correction des traitements. Ces trois règles se basent par ordre de priorité sur l'ajustement de l'intensité de l'éclairage, sur la position des sources lumineuses et enfin sur le déplacement de la tête de mesure stéréoscopique. Ici, nous nous reposons sur des traitements locaux autour des primitives non perçues pour corriger les conditions de prises de vue. Soit I_{max} , I_e et I_{min} les paramètres identifiant respectivement une image saturée, équilibrée ou sombre. Nos règles sont les suivantes :

Règles 1 : Si une primitive n'est pas vue et si $I \ll I_e$ « image sombre » ou $I \gg I_e$ « image saturée », alors augmenter (diminuer) l'intensité.

Règle 2 : Si une primitive n'est pas vue et si $I_{min} \leq I \leq I_{max}$ « image équilibrée »,

4.7 Description du dispositif expérimental

alors changer la position de la source lumineuse de *pos1* à *pos2*.

Règle 3 : Si une primitive n'est pas vue et si $I_{min} \leq I \leq I_{max}$ « image équilibrée », et si la source est en position *pos2*, alors faire un déplacement de la tête de mesure (capteur stéréoscopique).

- **Ajustement_éclairage.** Cette situation permet de valider l'ajustement de l'éclairage.
- **Ajustement_déplacement.** Cette situation permet d'instancier la situation de déplacement si les différents réglages de l'éclairage ne donnent pas de résultat correct. Dans ce cas, nous faisons un déplacement local de la tête de mesure.

La figure 4.10 illustre l'appel de la replanification dynamique par le SGT pour corriger l'éclairage de la pièce à analyser.

4.7 Description du dispositif expérimental

Le dispositif expérimental qui va servir à l'ensemble des expériences qui suivent est constitué d'une tête de mesure stéréoscopique, composée de deux caméras CCD du type Sony XC – 55, avec une focale fixe de 12 mm. Le dispositif mécanique permettant le déplacement de l'objet est composé d'une table micrométrique comportant quatre degrés de liberté. Trois de ces degrés sont des translations selon les axes X, Y, Z de la table avec une course maximale de 100mm et un pas de 10 micromètres. Le dernier est une rotation de la table, effectuée avec un pas de 0.01° autour de l'axe vertical du dispositif. Nous avons utilisé ce système pour contrôler les acquisitions des images réelles des pièces à évaluer. Le même système est utilisé pour étalonner les caméras en plaçant une mire d'étalonnage sur la table. Plusieurs sources lumineuses sont placées sur le banc expérimental supportant la table micrométrique, autour de l'objet (pièce L) (voir figure 4.11).

4.8 Correction de l'éclairage

Lorsqu'on veut ajuster l'intensité d'une source lumineuse, la question qui se pose est la suivante : comment peut-on « mesurer » l'intensité de l'éclairage pour pouvoir l'évaluer puis l'ajuster ? Nous avons vu dans le chapitre sur l'illumination qu'il n'y a pas de mesure d'intensité pouvant être utilisée directement. Les différents modèles d'éclairage qui existent sont souvent déduits des résultats de diverses expériences. Plusieurs expériences ont ainsi été réalisées sur notre banc d'essais. Nous avons placé sur la table micrométrique une source lumineuse ponctuelle et, par réglage de l'intensité de cette source (*Stabilite*TM 40 DCD de StockerYale) de 0 à 10 unités, des images de la scène contenant la pièce à étudier ont été acquises.

L'algorithme utilisé pour l'évaluation de la qualité des images acquises est simple. La séparation de l'objet du fond de l'image se fait par la méthode classique des k-moyennes (*K-means*)[DHS01]. Puis, on étudie l'histogramme de l'intensité de l'éclairage de l'objet.

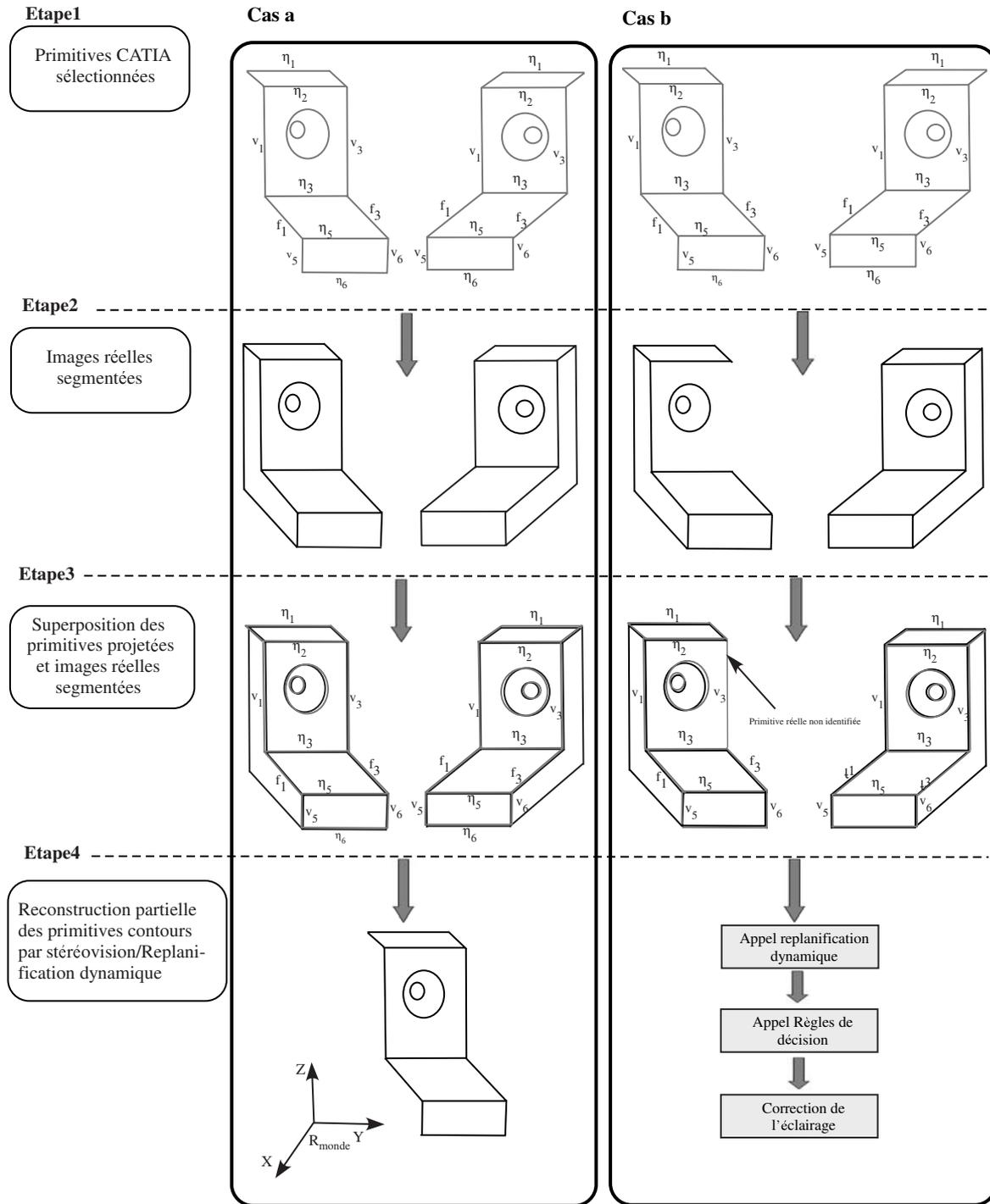


FIG. 4.10 – Appel de la replanification dynamique et identification des primitives réelles par comparaison aux primitives prédites (Cas a : pas de replanification. Cas b : avec replanification)

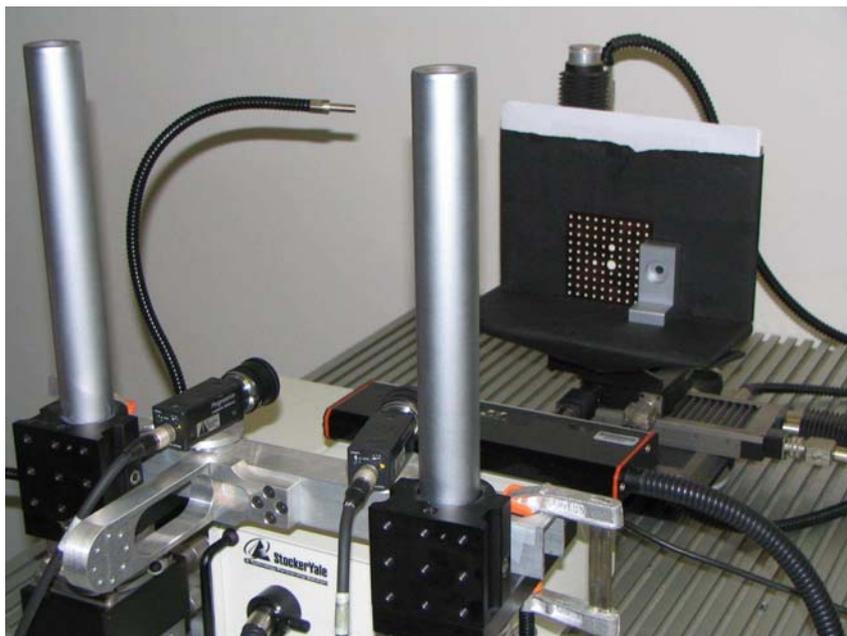


FIG. 4.11 – Dispositif expérimental

Cet histogramme est caractéristique pour une image saturée. Nous observons, en effet, la présence d'un pic dès qu'une zone de saturation apparaît dans l'image (voir figure 4.13). Pour décider si une image est sombre ou équilibrée, nous avons préféré utiliser les histogrammes cumulés pour évaluer leur niveau d'illumination. D'un côté une image équilibrée possède un histogramme presque uniforme, ce qui donne idéalement un histogramme cumulé sous la forme d'une droite de pente moyenne P_1 (voir figure 4.12). D'un autre côté, l'histogramme d'une image sombre présente un pic aux alentours des faibles intensités, ce qui conduit à un histogramme cumulé avec une pente P_2 importante aux alentours des intensité faibles.

En pratique, nous avons calculé les pentes P_1 et P_2 entre 20 % et 80 % du maximum de l'histogramme cumulé comme indiqué par la figure 4.12.

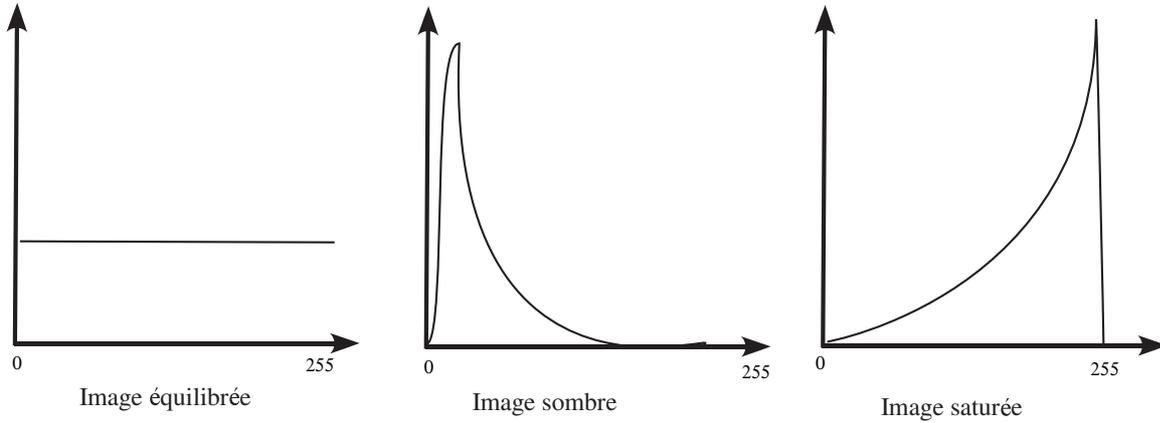
En nous basant sur cette mesure, nous avons pu mesurer l'intensité de l'éclairage d'images acquises équilibrées, sombres et saturées. La saturation est détectée par le pic présent dans l'histogramme aux valeurs fortes d'intensité (voir figure 4.13). Une image équilibrée ou trop sombre est aisément détectée par l'étude de l'histogramme cumulé (voir figures 4.14, 4.15).

4.9 Implémentation

Les premières expériences ont été menées avec le dispositif expérimental de la figure 4.11. Les informations liées à la modélisation portent sur les paramètres d'étalonnage du capteur, les primitives prédites de la pièce, les valeurs spécifiques des paramètres du traitement d'image dédié à la stéréovision, etc.

Nous allons présenter dans cette section l'implémentation de la méthode, ainsi que quelques résultats obtenus.

Histogrammes idéaux



Histogrammes cumulés idéaux

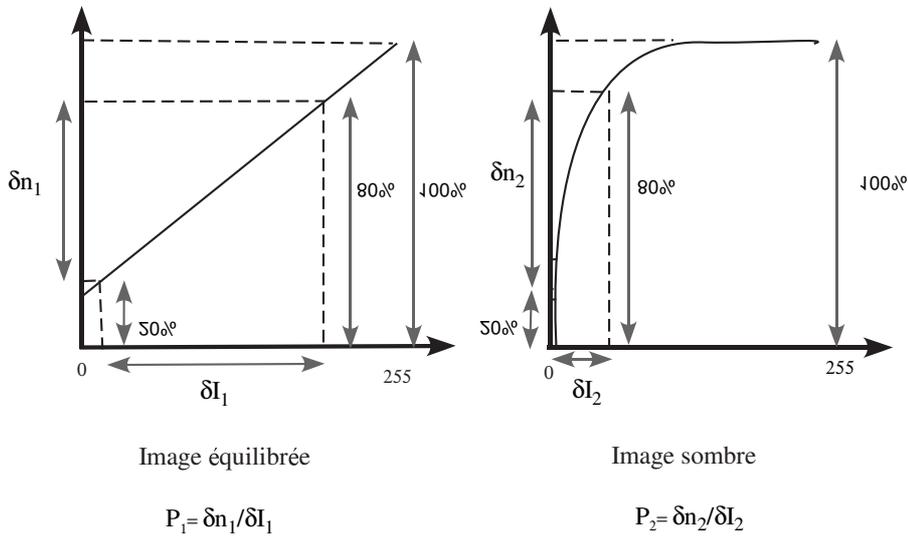


FIG. 4.12 – Schéma illustrant l'allure des histogrammes idéaux pour une image équilibrée, sombre et saturée

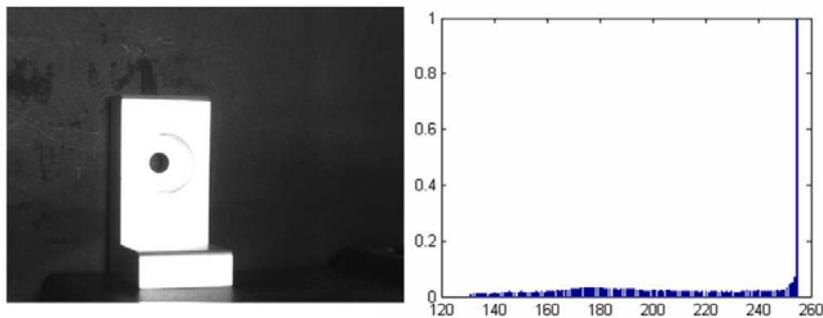


FIG. 4.13 – Exemple de résultat de la mesure de l'éclairage dans une image saturée

4.9 Implémentation

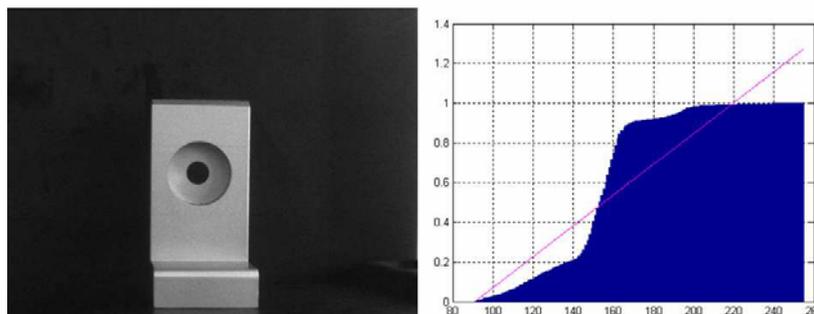


FIG. 4.14 – Exemple de résultat de la mesure de l'éclairage dans une image équilibrée

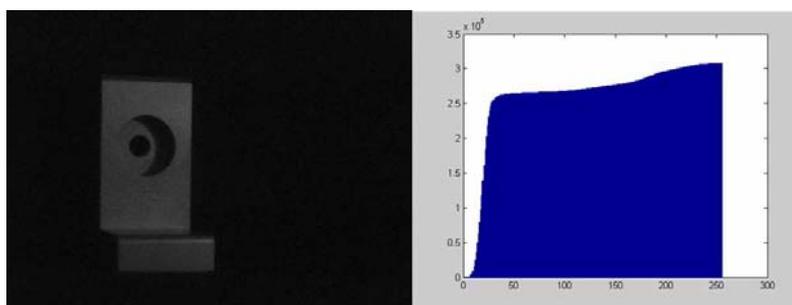


FIG. 4.15 – Exemple de résultat de la mesure de l'éclairage dans une image sombre

4.9.1 Génération du SGT dédié à la reconstruction 3D partielle par stéréovision

Nous avons utilisé SGTEditor qui est un éditeur graphique adapté à la conception des graphes de situations [Nag99]. Nous avons ainsi généré notre graphe de situations dédié à la reconstruction tridimensionnelle de pièces polyédriques, voir figure (4.16).

Ce graphe permet de mettre en œuvre la modélisation comportementale de la tête de mesure stéréovision. Il contient toutes les situations abordées dans la section 4.6 de ce chapitre et appropriées aux traitements. Cet éditeur nous a permis aussi de générer les spécialisations de chaque situation. Cette description graphique peut être aussi donnée sous la forme d'une version textuelle. En ce qui concerne l'implantation logicielle, nous avons utilisé les logiciels fournis avec SGTEditor.

4.9.2 Système d'acquisition et d'évaluation

À terme, le système développé à l'aide du dispositif expérimental sera implémenté dans un robot de mesure d'ores et déjà disponible dans notre laboratoire et composé d'une structure circulaire sur laquelle plusieurs bras articulés, comportant les têtes de mesure, peuvent se déplacer. Des sources lumineuses seront disposées autour de l'objet. La position et l'orientation des têtes de mesure (stéréovision et lumière structurée) seront exprimées dans un repère de référence du système d'acquisition. Ces deux têtes de mesure sont :

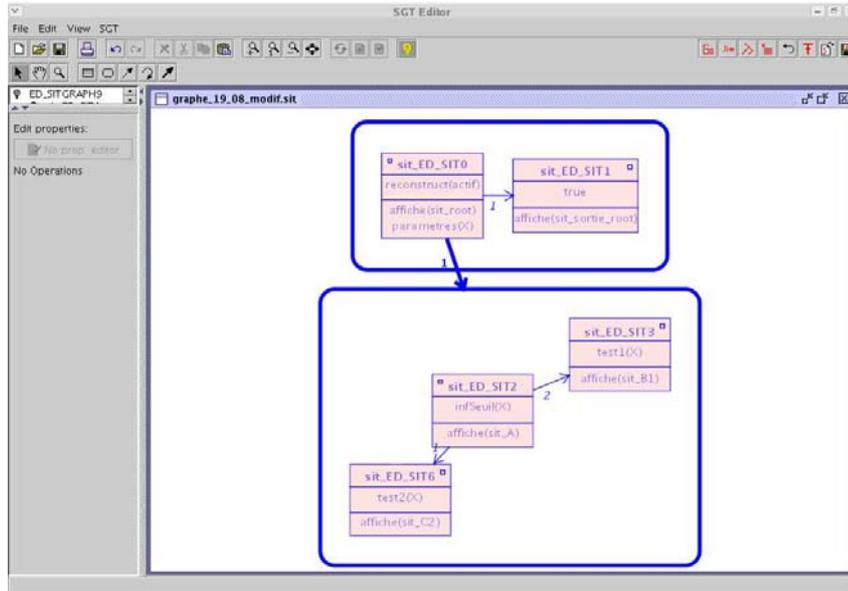


FIG. 4.16 – Exemple d'un Graphe de situations réalisé avec SGTeditor

Tête de mesure stéréovision : elle est composée de deux caméras CCD standards et permet d'effectuer des acquisitions de paires d'images stéréoscopiques.

Tête de mesure lumière structurée : elle est composée d'une caméra CCD standard couplée avec un projecteur de lumière. On aboutit à une information 3D par l'étude de la déformation de la projection du motif lumineux régulier projeté (p. ex., des franges verticales alternativement noires et blanches) sur la pièce à inspecter.

Pour le moment, ces têtes de mesure sont utilisées indépendamment l'une de l'autre et uniquement dans les dispositifs expérimentaux. Le robot de mesure sera amené à les gérer conjointement.

La validation de la replanification par les SGT est présentée dans la section suivante. Cette validation a été faite en utilisant la tête de mesure stéréoscopique dans le dispositif expérimental.

4.10 Validation de la replanification par SGT

L'automatisation de la reconstruction 3D par vision stéréoscopique a conduit au développement d'un SGT spécifique. Ce dernier est chargé de la gestion des opérations d'acquisition et de reconstruction partielle planifiées, tout en permettant une replanification dynamique adaptée au type de problèmes rencontrés pendant les traitements. Un graphe de situations a ainsi été développé afin de valider notre approche. Ce dernier est composé de cinq situations de base : *Déplacement*, *Acquisition*, *Traitement*, *Reconstruction 3D*, *Replanification dynamique*.

Les premiers résultats obtenus portent sur une reconstruction partielle avec une correction dynamique de l'intensité de l'éclairage.

La segmentation obtenue en ligne est comparée aux images « fil de fer » décrivant les primitives prédites. Dans le cas où nous n'arrivons pas à détecter une primitive, par comparaison avec l'information *a priori*, on fait appel à une replanification dynamique afin de modifier les paramètres de la segmentation. S'il est toujours impossible de détecter une primitive prédite, on procède aussi à une replanification dynamique qui va permettre d'agir sur les paramètres physiques de l'éclairage. Une nouvelle acquisition des images est réalisée avec de nouveaux paramètres pour l'éclairage. Les nouvelles images segmentées sont ensuite comparées une nouvelle fois avec les primitives prédites. Si la comparaison est positive, on réalise la reconstruction partielle de l'objet.

Notre graphe de situations permet ainsi de générer deux types de séquence ; le premier correspond au cas où le système arrive à identifier toutes les primitives prédites (voir figure 4.10. Cas a), ce qui amène directement à une reconstruction 3D des contours de la pièce à inspecter. La séquence est donc (voir figure 4.17) :

Reconstruction_partielle_SV > Déplacement > Acquisition > Test_Image > Ajustement_Intensité > Acquisition > Test_Image > Validation_Contraste > traitement > Segmentation > Superposition > Appariement > reconstruction_3D_partielle.

Le deuxième type, quant à lui, fait appel à une replanification dynamique, parce que le système n'arrive pas à identifier toutes les primitives prédites (voir figure 4.10. Cas b). Cette replanification permet de modifier les paramètres de l'éclairage pour ensuite tenter d'obtenir toutes les primitives de type contours de la pièce. Cela conduit à une séquence de traitement du type (voir figure 4.18) :

Reconstruction_partielle_SV > Déplacement > Acquisition > Test_Image > Validation_Contraste > traitement > Segmentation > Superposition > Replanification_Dyn-amique > Validation_Primitive_non_perçue > Ajustement_Éclairage > Traitement > Segmentation > Superposition > Appariement > reconstruction_3D_partielle.

Le parcours du SGT en cas de correction des paramètres d'illumination s'effectue d'une façon tout à fait correcte et conduit aux résultats montrés dans la figure 4.19.

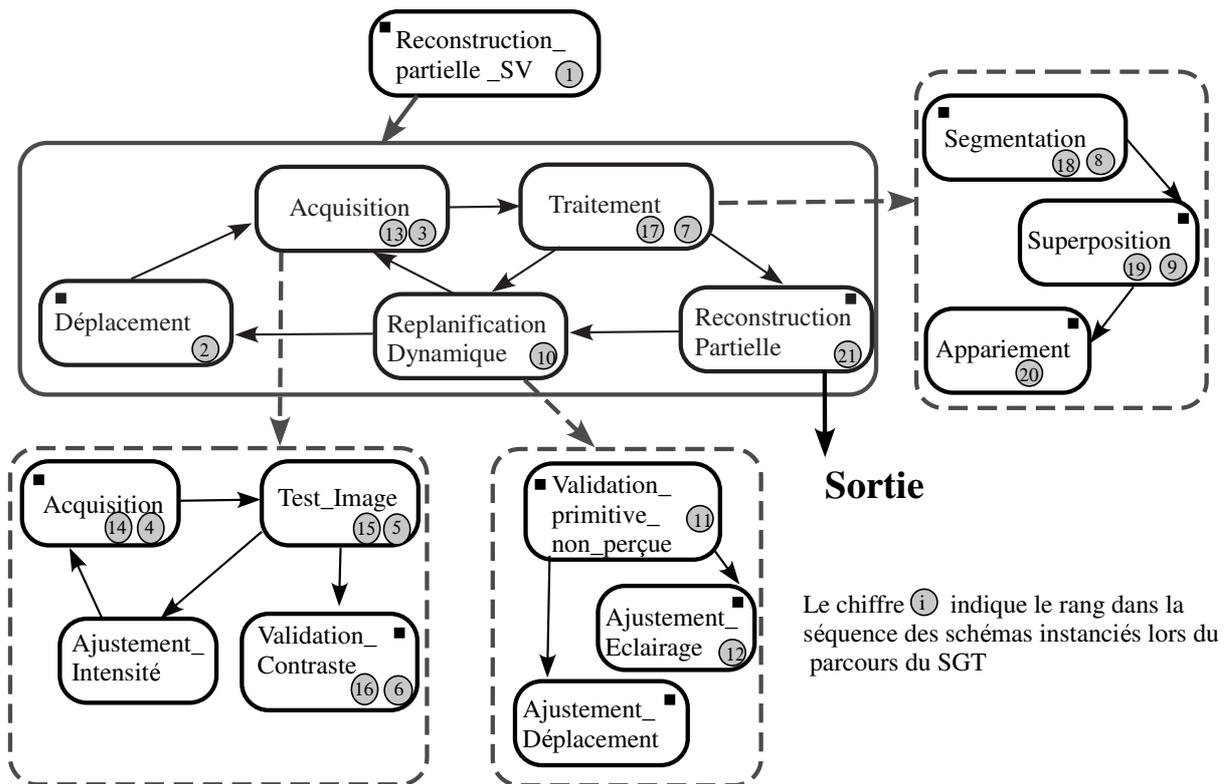


FIG. 4.18 – Séquence de schémas de situation pour une reconstruction partielle par stéréovision avec replanification dynamique

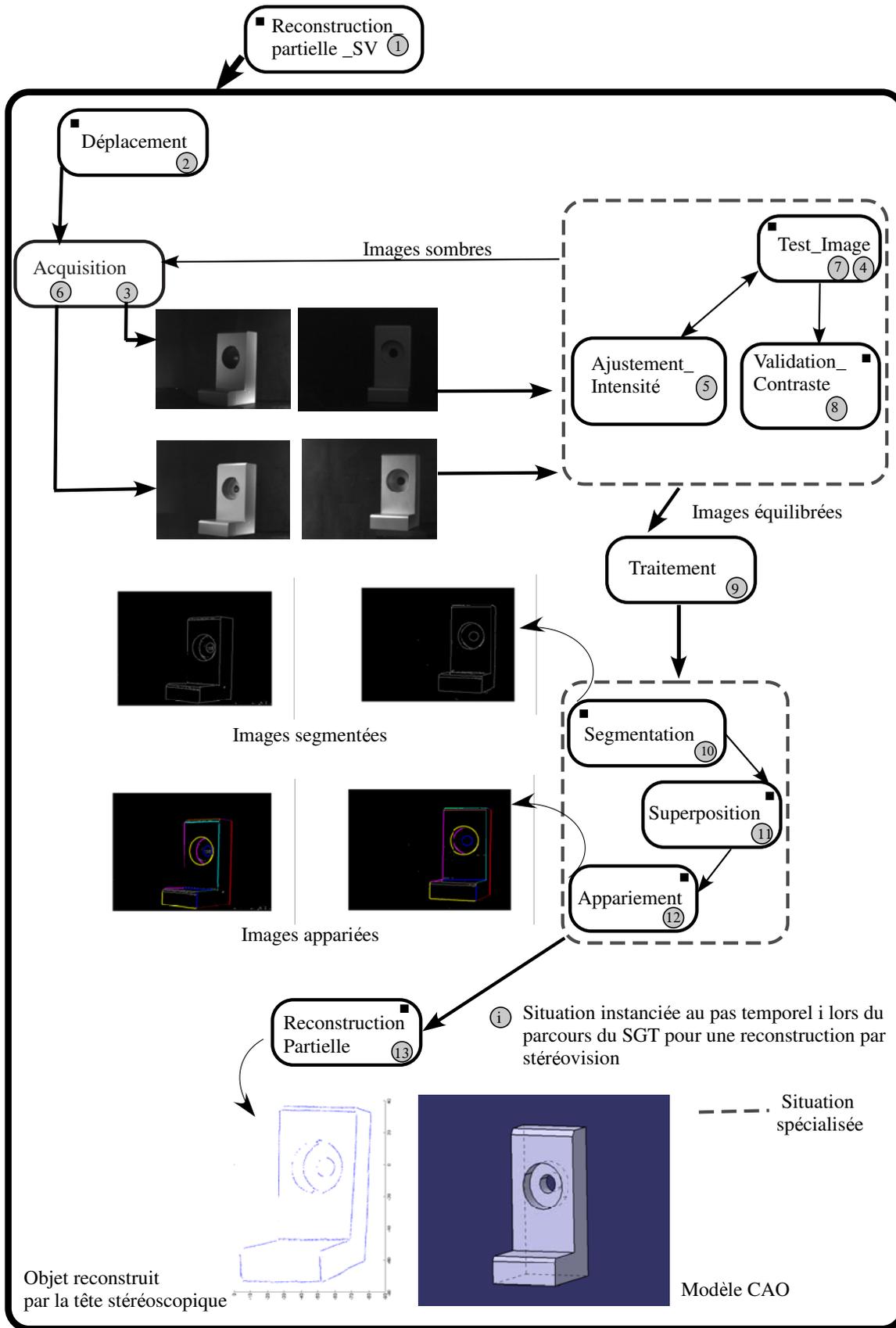


FIG. 4.19 – Diagramme simplifié du parcours du SGT en cas de correction des paramètres d'illumination

4.11 Conclusion

L'approche décrite porte sur le développement d'un système de planification permettant de gérer l'inspection et l'évaluation dimensionnelle d'une pièce manufacturée. La connaissance *a priori* mise en œuvre porte sur l'ensemble des composants de l'application, à savoir le modèle de la tête de mesure, des sources lumineuses et de l'objet à analyser.

Dans le cadre de cette thèse, notre contribution a porté sur une analyse complète du processus de reconstruction tridimensionnelle par stéréovision de pièces polyédriques, en utilisant les SGT. Nous avons ainsi été amenés à étudier l'automatisation du processus de reconstruction, tout en respectant les règles de fonctionnement d'un SGT. Nous avons ensuite utilisé nos développements présentés dans les deuxième et troisième chapitres de ce mémoire et concernant, respectivement, une méthode d'appariement permettant d'obtenir une reconstruction 3D partielle de la pièce étudiée, et la conception d'un modèle CAO CATIA, approprié aux pièces à évaluer et permettant la génération d'images de synthèse respectant le point de vue réel du capteur.

Notre SGT est capable de corriger les traitements en cours d'exécution en cas de résultat non satisfaisant, que ce soit au niveau du traitement ou bien au niveau de la reconstruction partielle. Ce graphe rassemble un ensemble de situations. Chacune d'elle concerne une étape de la reconstruction tridimensionnelle. Ces situations sont au nombre de 5 : déplacement, acquisition, traitement, reconstruction partielle et planification dynamique.

Perspectives Parmi les perspectives, nous pouvons citer la réalisation, en cours de développement, d'une reconstruction complète de l'objet effectuée en parcourant plusieurs fois le SGT. À chaque parcours, nous devons fournir au SGT les paramètres nécessaires pour réaliser une reconstruction tridimensionnelle partielle selon le point de vue spécifié. Nous réalisons ainsi, par exemple, quatre acquisitions correspondant à quatre points de vue (vue de face, vue de gauche, vue de droite et vue de derrière) (voir figure 4.20)).

La chaîne de traitement permettant une reconstruction complète de l'objet, ainsi que son évaluation, combine des connaissances *a priori* liées au modèle de l'objet, au système d'acquisition (tête de mesure stéréoscopique), aux traitements, ainsi qu'à une étape de recalage des reconstructions partielles. Toutes les acquisitions en vue de la reconstruction complète de l'objet et les traitements des images sont réalisés par le SGT. La gestion de la reconstruction 3D par stéréovision est réalisée par le graphe de contrôle HTP (voir figure 4.1) qui fournit les valeurs des différents jeux de paramètres estimés *a priori* au SGT. La reconstruction complète, et la comparaison des données réelles mesurées et conceptuelles issues du modèle de référence CATIA qui suit, permettra l'évaluation dimensionnelle de l'objet.

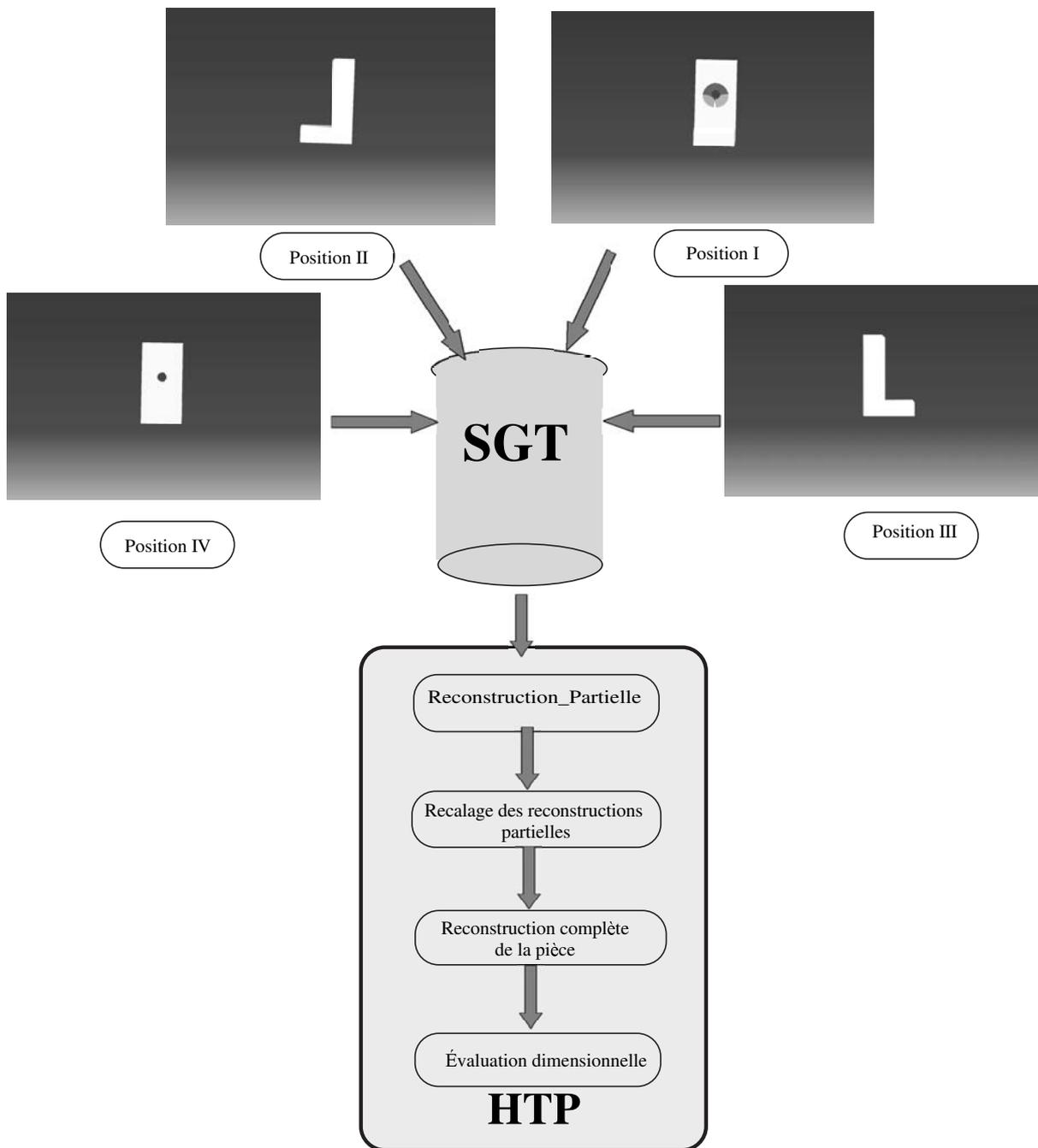


FIG. 4.20 – Simulation à l'aide de CATIA d'une séquence de 4 acquisitions conduisant à la reconstruction complète de la pièce « L »

Conclusion générale et perspectives

Notre projet, dans sa globalité, vise au développement d'une plate-forme de mesure en vue d'un contrôle qualité de pièces industrielles par vision. L'idée principale est d'utiliser un bras manipulateur avec plusieurs degrés de liberté afin d'être en mesure de déplacer le capteur autour de la pièce à inspecter. Les têtes de mesure permettant l'acquisition des images sont placées sur l'extrémité du bras du robot avec pour objectif de réaliser une mesure dimensionnelle de la pièce manufacturée. L'approche adoptée consiste à reconstruire en 3 dimensions le contenu des images de l'objet par stéréovision. Cette reconstruction est ensuite comparée avec le modèle de l'objet issu d'un outil CAO, l'automatisation du système d'inspection étant réalisée par l'utilisation des graphes de situations (SGT).

Dans une première contribution à cette thèse, nous avons introduit une technique originale d'appariement de contours. La performance de cet algorithme d'appariement dépend fortement de l'estimation de la matrice fondamentale et de la qualité de la segmentation. Développée spécifiquement pour des pièces manufacturées quasi-polyédriques, en partant d'une classification de contours en formes géométriques simples, la procédure d'appariement évalue les correspondances établies de la gauche vers la droite et vice-versa, en utilisant des matrices d'appariement. L'approche repose aussi sur une sélection des contours à apparier, obtenue en comparant les images acquises avec des images conceptuelles générées à l'aide d'un outil CAO.

Dans une deuxième contribution, ces images conceptuelles générées à l'aide d'un modèle CAO ont permis d'évaluer certains paramètres d'illumination utiles à une éventuelle replanification des traitements liés à la reconstruction (p. ex., réglage ou déplacement de la source lumineuse). La position de la source lumineuse a ainsi effectivement pu être estimée en étudiant l'évolution de la luminance dans les images acquises.

Dans une troisième contribution, nous avons analysé la procédure de reconstruction par stéréovision de pièces manufacturées en utilisant les graphes de situations. L'évaluation des images s'effectue en exploitant les mêmes outils que ceux utilisés pour la simulation et permet d'ajuster les traitements (replanification dynamique). À partir d'une analyse multi-images, nous avons ainsi déterminé des descriptions scéniques en faisant intensément appel à une connaissance *a priori*. Ces descriptions ont été planifiées et sont obtenues automatiquement par l'utilisation d'un outil de planification faisant appel aux graphes de situations.

L'approche développée permet une reconstruction tridimensionnelle partielle de l'objet, avec une replanification dynamique fonction des résultats obtenus pendant les traitements. Pour ce faire, nous avons développé des règles de décision adaptées aux problèmes rencontrés sur notre système de mesure et de reconstruction. Lors de l'application de ces règles, nous avons mis la priorité sur le réglage de l'éclairage. Ce choix est justifié par le fait que tout changement de position de la tête de mesure nécessite un nouvel étalonnage des caméras formant la tête stéréoscopique.

Nos règles sont simples mais efficaces et ne représentent qu'une première liste de consignes qui devra être complétée. Lors du développement de ces règles, nous nous sommes ainsi basés sur le contrôle de la qualité des images réelles acquises. En effet,

la qualité de la prise de vue a une influence considérable sur la localisation des contours, ce qui influence par la suite l'exactitude des mesures dimensionnelles et, cela, quelles que soient les performances des traitements effectués par notre système. Il est donc intéressant de mettre en œuvre des procédures d'ajustement permettant d'obtenir des images réelles qui reflètent au mieux le contenu de la scène, en contrôlant les paramètres d'acquisition (éclairage, caméra).

Notre système de mesure par vision de pièces manufacturées a été élaboré en faisant appel à une connaissance *a priori* provenant d'un modèle CAO. Ce modèle a été conçu à l'aide du logiciel CATIA V5 (Dassault Systems). Il est composé de descriptions géométriques et topologiques de la pièce. On peut ainsi déterminer *a priori*, pour une phase ultérieure de mesure à partir d'images de la pièce, les informations nécessaires pour évaluer des caractéristiques plus ou moins élémentaires de la pièce. Cette phase de traitement hors ligne permet d'initialiser les opérations de mesure en ligne.

Le système de mesure est basé sur la comparaison d'images réelles de la pièce, acquises à l'aide de caméras CCD, avec des images conceptuelles générées à partir des données géométriques de la pièce telle qu'elle doit être usinée. Les conditions de prise de vue, telles que la position de l'objet par rapport au capteur ainsi que le placement des sources lumineuses, sont connues afin de simuler la scène à analyser pour générer les images de synthèse. La comparaison des images réelles et synthétiques permet ainsi un contrôle quantitatif de la pièce. Pour effectuer la comparaison avec les images réelles, celles-ci doivent être soigneusement traitées, et conduire à une extraction soignée et précise des contours.

En conclusion, l'approche proposée dans cette thèse permet d'effectuer une mise en correspondance efficace de primitives extraites d'images stéréoscopiques de pièces manufacturées, pour assurer une reconstruction 3D précise. Notre contribution a notamment permis l'automatisation de la procédure de reconstruction en utilisant les graphes de situations.

Le développement d'un dispositif d'acquisition contrôlée d'images permet d'envisager, à moyen terme, la maîtrise complète des conditions d'acquisition, autorisant ainsi une extraction contrôlée d'indices images et leur évaluation quantitative de manière robuste et fiable.

Les travaux en cours dans notre laboratoire visent à développer une méthodologie combinant une approche orientée surface avec la méthode orientée contour présentée, en vue d'une reconstruction 3D plus complète. Pour ce faire, nous relierons les données de type contour 3D avec des données de type surface, afin d'obtenir une description plus riche de l'objet. Nous envisageons de combiner les deux têtes de mesure, stéréoscopique et lumière structurée, dans le robot de mesure. Le système de planification, validé sur la version prototype du robot de mesure, est en cours d'amélioration, afin d'assurer le contrôle des deux têtes de mesure simultanément. Chacune des deux têtes de mesure est actuellement utilisée indépendamment pour générer, soit des données 3D spécifiques aux contours (Stéréovision), soit des mesures liées aux surfaces (Lumière structurée). Une reconstruction 3D efficace de pièces manufacturées peut être obtenue en combinant ces deux approches,

et permettra d'obtenir une description bien plus riche de l'objet.

Les principales applications considérées portant sur une analyse dimensionnelle des pièces par comparaison avec un modèle conceptuel. Un robot d'inspection capable de couvrir le volume d'une demi-sphère de 2 mètres de rayon est en cours de développement. Ce robot sera intégré à la plate-forme d'analyse contrôlée d'images. L'ensemble du travail présenté dans cette thèse s'intégrera dans cette plate-forme. La figure (4.21) illustre la démarche envisagée.

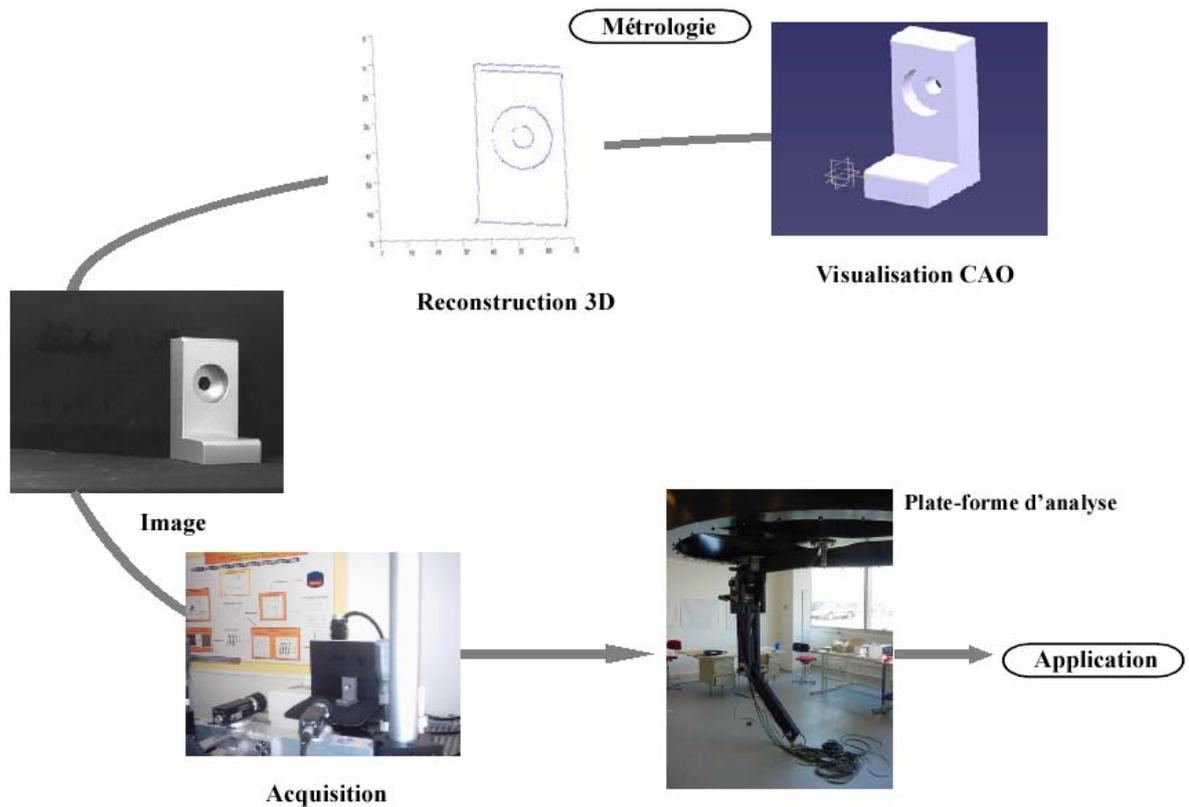


FIG. 4.21 – Plate-forme d'analyse contrôlée d'images pour la modélisation 3D et la robotique

Bibliographie

- [AF87] N. Ayache and B. Faverjon. Efficient registration of stereo images by matching graph descriptions of edge segments. *International Journal of Computer Vision*, 1(2), pages 107–132, 1987.
- [AG92] P. Aschwanden and W. Guggenbuhl. Experimental results from a comparative study on correlation-type registration algorithms. In *Robust Computer Vision : Quality of Vision Algorithms*, W. Förstner and S. Ruwiedel, Eds. Karlsruhe, Germany : Wichmann, pages 268–282, 1992.
- [Alo88] A. Aloimonos. Visual shape computation. In *Proceedings of the IEEE*, volume 76, pages 899–916, 1988.
- [AMDG02] D. Aluze, F. Merienne, C. Dumont, and P. Gorria. Vision system for defect imaging, detection, and characterization on a specular surface of a 3D object. *Image and Vision Computing*, 20(8), pages 569–580, Juin 2002.
- [AN02] M. Arens and H. H. Nagel. Representation of behavioural knowledge for planning and plan-recognition in a cognitive vision systems. In M. Jarke, J. Koehler, and G. Lakemeyer, editors, *Proceedings of 25th German Conference on Artificial Intelligence, LNAI 2479, Springer-Verlag, Berlin*, pages 268–282, 2002.
- [AN03] M. Arens and H. H. Nagel. Behavioural knowledge representation for the understanding and creation of video sequences. In A. Gunther, R. Kruse, and B. Neumann, editors, *Proceedings of 26th German Conference on Artificial Intelligence, LNAI 2821, Springer-Verlag, Berlin*, pages 149–163, 2003.
- [AON02] M. Arens, A. Ottlik, and H. H. Nagel. Natural language text for a cognitive vision system. In F. Van Harmelen, editor, *Proceedings of 15th European Conference on Artificial Intelligence (ECAI), Springer-Verlag, Lyon, France*, pages 455–459, 2002.
- [APL] Web site :<http://www.oplusplus.com/>.
- [APS01] X. Armangué, J. Pagès, and J. Salvi. Comparative survey on fundamental matrix estimation. *International Journal of Computer Vision*, 1, pages 227–232, 2001.
- [ATLB04] B. Albouy, S. Treuillet, Y. Lucas, and D. Birov. Fundamental matrix estimation revisited through a global 3D reconstruction framework. In *Proceeding of Advanced Concepts for Intelligent Vision Systems, Bruxelles, Belgique*, pages 185–192, 2004.
- [AWB87] A. Aloimonos, I. Weiss, and A. Bandopadhy. Active vision. *International Journal of Computer Vision*, 1(4), pages 333–356, 1987.

- [Aya89] N. Ayache. *Vision stéréoscopique et perception multisensorielle. Application à la robotique mobile*. Interedition, Paris, 1989.
- [Baj86] R. Bajcsy. Active perception. In *Proceedings of the IEEE*, volume 76, pages 996–1005, 1986.
- [Bal91] D. H. Ballard. Animate vision. *Artificial Intelligence*, 48(1), pages 57–86, 1991.
- [BB81] H. H. Baker and T. O. Binford. Depth from edge- and intensity-based stereo. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence, Vancouver, Canada*, pages 631–636, 1981.
- [BB82] D. H. Ballard and C. M. Brown. *Computer Vision*. Prentice-Hall, Englewood Cliffs, New Jersey, 1982.
- [BH87] B. Bhanu and C. Ho. CAD based 3D object representation for robot vision. *Computer*, 20(8), pages 19–35, 1987.
- [BHH85] B. G. Batchelor, D. A. Hill, and D.C. Hodgson. *Automated Visual Inspection*. IFS Ltds., Bedford, UK, 1985.
- [BMB94] P. Brand, R. Mohr, and P. Bobet. Distorsions optiques : correction dans un modèle projectif. In *Proceedings Reconnaissance des Formes et Intelligence Artificielle, Paris, France*, pages 87–98, 1994.
- [BMZ92] P. Beardsley, D. Murray, and A. Zisserman. Camera calibration using multiple images. In *Proceedings of the Second European Conference on Computer Vision, Genoa*, pages 312–320, 1992.
- [BN96] D. N. Bhat and S. K. Nayar. Ordinal measure for visual correspondence. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, San Francisco, California, USA*, pages 351–357, 1996.
- [Bén02] M. Bénallal. *Système de calibration de caméra, localisation de formes polyédriques par vision monoculaire*. Thèse de doctorat, École des Mines de Paris, Novembre 2002.
- [Bou93] S. Boukir. *Reconstruction 3D d'un environnement statique par vision active*. Thèse de doctorat, Université de Rennes 1, Octobre 1993.
- [Bou94] B. Boufama. *Reconstruction tridimensionnelle en vision par ordinateur : cas des caméras non étalonnées*. Thèse de doctorat, Institut Nationale Polytechnique de Grenoble, Novembre 1994.
- [BS63] P. Beckmann and A. Spizzichino. The scattering of electromagnetic waves from rough surfaces. 1963.
- [BT80] S. T. Barnard and W. B. Thompson. Disparity analysis of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2(4), pages 333–340, 1980.
- [Can86] J. Canny. A computational approach to edge detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8(6), pages 679–698, 1986.
- [Cata] Web site : <http://cfao-catiav5.grenoble.cnrs.fr/>.
- [Catb] Web site : <http://www.synchrotron-soleil.fr/>.

BIBLIOGRAPHIE

- [CBBJ96] F. Chaumette, S. Boukir, P. Bouthemy, and D. Juvin. Structure from controlled motion. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18(5), pages 492–504, 1996.
- [CBvdHG04] W. Chojnacki, M. J. Brooks, A. van den Hengel, and D. Gawley. A new approach to constrained parameter estimation applicable to some computer vision problems. *Image and Vision Computing*, 22(2), pages 85–91, 2004.
- [Chr98] S. Christy. *Localisation et modélisation tridimensionnelles par approximations successives du modèle perspectif de caméra*. Thèse de doctorat, Institut National Polytechnique de Grenoble, Septembre 1998.
- [CM92] S. D. Cochran and G. Medioni. 3D surface description from binocular stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10), pages 981–994, 1992.
- [CMB⁺95] J. P. Coquerez, A. Montenvert, Ph. Bolon, J. M. Chassery, and C. Graffigne. *Analyse d'images : filtrage et segmentation*. Masson, Paris, 1995.
- [COR] Web site :<http://www.graphics.cornell.edu>.
- [Cou94] B. Couapel. *Stéréovision par ordinateur*. Thèse de doctorat, Université de Rennes I, Mars 1994.
- [Cou97] C. Coulot. *Etude de l'éclairage de surfaces métalliques pour la vision artificielle : application au contrôle dimensionnel*. Thèse de doctorat, Université de Bourgogne, Juin 1997.
- [CR89] F. Chaumette and P. Rives. Modélisation et calibration d'une caméra. In *Proceedings of Reconnaissance des Formes et Intelligence Artificielle, Paris, France*, volume 1, pages 527–536, 1989.
- [Dau94] C. Daul. *Construction et utilisation de listes de primitives en vue d'une analyse dimensionnelle de pièces à géométrie simple*. Thèse de doctorat, Université Louis Pasteur de Strasbourg, 1994.
- [Dev97] F. Devernay. *Vision stéréoscopique et propriétés différentielles des surfaces*. Thèse de doctorat, école Polytechnique, Février 1997.
- [DF96] F. Devernay and O. Faugeras. From projective to euclidean reconstruction. In *Proceedings of International Conference on Computer Vision and Pattern Recognition, San Francisco, USA*, pages 264–269, Juin 1996.
- [DHS01] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. John Willy and Sons, New York, 2001.
- [DR81] L.S. David and A. Rosenfeld. Cooperating processes for low-level vision : A survey. *Artificial Intelligence*, 17(1-3), pages 245–263, Août 1981.
- [DR03] G. Damiand and P. Resch. Split-and-merge algorithms defined on topological maps for 3D image segmentation. *Graphical Models*, 65(3), pages 149–167, 2003.
- [DRLR89] M. Dhome, M. Richetin, J. T. Lapreste, and G. Rives. Determination of the attitude of 3D objects from a single perspective view. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11(12), pages 1265–1278, 1989.

- [DZLF94] R. Deriche, Z. Zhang, Q. T. Luong, and O. Faugeras. Robust recovery of the epipolar geometry for an uncalibrated stereo rig. In *Proceedings of 3rd European Conference on Computer Vision, Stockholm, Suede*, volume 1, pages 567–576, Mai 1994.
- [Eva98] L. Evard. *Contribution à la reconstruction surfacique par organisation perceptuelle sous vision monoculaire en lumière structurée*. Thèse de doctorat, Université du Littoral, Boulogne-sur-Mer, 1998.
- [Fau92] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig. In *Proceedings of 2nd European Conference on Computer Vision, ECCV 92 Santa Margherita ligure, Italy*, volume 1, pages 536–578, Mai 1992.
- [Fau93] O. Faugeras. *Three dimensionnal computer vision : a geometric viewpoint*. MIT Press, Cambridge, Massachusetts, 1993.
- [FG93] A. Filbois and D. Gemerle. Modélisation d’objets polyédriques. In *Proceedings of ORASIS 1993, journées francophones des jeunes chercheurs en vision par ordinateur, Mulhouse*, pages 128–131, Octobre 1993.
- [FKH05a] A. Far, S. Kohler, and E. Hirsch. 3D reconstruction of manufactured parts using bi-directional stereovision-based contour matching and comparison of real and synthetic images. In *Proceedings of the 9th Conference on Machine Vision Applications, Tsukuba Science City, Japan*, pages 456–459, 2005.
- [FKH05b] A. Far, S. Kohler, and E. Hirsch. Reconstruction 3D de pièces industrielles à partir d’une segmentation et d’un appariement robuste d’images stéréoscopiques. In *Proceedings of ORASIS 2005, journées francophones des jeunes chercheurs on vision par ordinateur, Clermont-Ferrand*, Proceeding sur CD–Rom, article n°35, 2005.
- [FLM92] O. Faugeras, Q. T. Luong, and S. J. Maybank. Camera self-calibration : Theory and experiments. In *Proceedings of the Second European Conference on Computer Vision, ECCV 92, Santa Margherita Ligure*, volume 1, pages 321–334, Mai 1992.
- [FT86] O. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, USA*, pages 15–20, Juin 1986.
- [FTV97] A. Fusiello, E. Trucco, and A. Verri. Rectification with unconstrained stereo geometry. *Proceedings of the British Machine Vision Conference, United Kingdom*, pages 400–409, 1997.
- [Gev00] P. Geveaux. *Conception d’un système de detection de défauts sur surfaces texturées par traitement d’images : application au contrôle haute cadence de pièces en defilement*. Thèse de doctorat, Université de Bourgogne, Mai 2000.
- [Ghe92] R. Gherbi. *Mise en correspondance de chaînes de contour. Application à la stéréovision*. Thèse de doctorat, Université Paris XI, Orsay, Décembre 1992.

BIBLIOGRAPHIE

- [GLZH02] P. Graebbling, A. Lallement, D. Y. Zhou, and E. Hirsch. Optical high precision 3D vision based quality control of manufactured parts using synthetic images and knowledge for image data evaluation and interpretation. *Applied Optics : Information processing*, 41(14), pages 2627–2643, 2002.
- [GM90] Z. Gigus and J. Malik. Computing the aspect graph for line drawings of polyhedral objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(2), pages 113–122, 1990.
- [Gri85] L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1), pages 17–34, 1985.
- [GS03] P. Gurdjos and P. Sturm. Methods and geometry for plane-based self-calibration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin*, volume 1, pages 491–496, Juin 2003.
- [GWCP05] O. Ghita, P. F. Whelan, T. Carew, and N. Padmapriya. Quality grading of painted slates using texture analysis. *Computers in Industry*, 56(8-9), pages 802–815, Décembre 2005.
- [GWES02] R. Gonzalez, R. E. Woods, S. L. Eddins, and L. Steven. *Digital Image Processing*. Prentice-Hall, 2nd edition, 2002.
- [HA89] W. Hoff and N. Ahuja. Surfaces from stereo : integrating feature matching, disparity estimation, and contour detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(2), pages 121–136, 1989.
- [HA94] X. Hu and N. Ahuja. Feature extraction and matching as signal detection. *International Journal of Pattern Recognition and Artificial Intelligence*, 8(6), pages 1343–1379, 1994.
- [Har97] R. Hartley. In defence of the eight-point algorithm. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(6), 1997.
- [HB89] B. K. P Horn and M. J. Brooks. *Shape from shading*. MIT Press, Cambridge, Massachusetts, 1989.
- [HGC92] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA*, pages 761–764, 1992.
- [HM93] R. Horaud and O. Monga. *Vision par ordinateur*. Hermès, Paris, 1993.
- [HN00] M. Haag and H. H. Nagel. Incremental recognition of traffic situations from video image sequences. *Image and Vision Computing*, 18(2), pages 137–153, 2000.
- [Hor00] R. Horaud. Vision 3D projective, affine et euclidienne. Technical report, INRIA, 2000.
- [HS89] R. Horaud and T. Skordas. Stereo correspondences through feature grouping and maximal cliques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11), pages 1168–1180, 1989.
- [HZ00] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, UK, 2000.

- [JNS04] F. Jurie, E. Noirfalise, and P. Sturm. Calibrage itératif de caméras à partir de scènes planes. In *Proceeding of 14ème Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle, Toulouse, France*, volume 1, pages 261–269, Janvier 2004.
- [JT95] J.M.Vezien and J. P. Tarel. A generic approach for planar patch stereo reconstruction. In *Proceedings of Scandinavian Conference on Image Analysis, Uppsala, Suède*, pages 1061–1070, 1995.
- [KA87] Y. C. Kim and J. K. Aggarwal. Positioning 3D objects using stereo images. *IEEE Journal on Robotics and Automation*, 3(4), pages 361–373, 1987.
- [KFH05] S. Kohler, A. Far, and E. Hirsch. Intelligent approaches for computer vision-based quality control. In *Proceedings of QCAV 2005, 7th International Conference on Quality Control by Artificial Vision, Nagoya, Japan*, pages 247–252, Mai 2005.
- [KH03] S. Kohler and E. Hirsch. Cognitive intelligent sensory system for vision-based quality control. In *Proceedings of QCAV 2005, 6th International Conference on Quality Control by Artificial Vision, SPIE*, pages 331–341, Mai 2003.
- [Kho91] M. K. Kholadi. *Représentation, modélisation et manipulation des connaissances spatiales en géomatique G.R.E.M.A.C.O.S.* Thèse de doctorat, INSA de Lyon, 1991.
- [KLH05] R. Khemmar, A. Lallement, and E. Hirsch. Design of an intelligent self-reasoning system for the automated vision based evaluation of manufactured parts. In *Proceedings of QCAV 2005, 7th International Conference on Quality Control by Artificial Vision, Nagoya, Japan*, pages 241–246, Mai 2005.
- [Kor88] A. F. Korn. Toward a symbolic representation of intensity changes in images. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 10(5), pages 610–625, 1988.
- [Kru97] W. M. Krueger. Characterization of generic light sources. In *Vision geometry VI, Proceedings SPIE*, volume 31, pages 300–306, 1997.
- [Lan97] Z. D. Lan. *Méthodes robustes en vision : application aux appariements visuels.* Thèse de doctorat, Institut National Polytechnique de Grenoble, 1997.
- [Len87] R. K. Lenz. Linsenfehlerkorrigierte eichung von halbleiterkamas mit standardobjektiven fur hochgenaue 3D messungen in echtzeit. *DAGM*, pages 212–215, 1987.
- [LF98] Q. T. Luong and O. D. Faugeras. The fundamental matrix : theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1), pages 43–75, 1998.
- [LH87] H. C. Longuet-Higgins. *A Computer Algorithm for Reconstructing a Scene from Two Projections.* Morgan Kaufmann Publishers Inc., 1987.
- [Lin99] D. Lingrand. Approximations de la projection perspective : modèle combiné et étude des singularités homographiques. Technical Report 3682, INRIA Sophia Antipolis, Avril 1999.

BIBLIOGRAPHIE

- [LKGC99] B. Lamalle, S. Kohler, P. Gorria, and C. Coulot. Dimensional control of metallic objects by artificial vision : contribution to lighting condition studies. *Journal of Optical Engineering*, 38(8), pages 1305–1311, 1999.
- [LTYM99] Y. Li, H. T. Tsui, S. Y. Yeng, and S. Ma. Shape from shading for multiple light sources. *Journal of the Optical Society of America (A : Optics and image science)*, 16(1), pages 36–52, 1999.
- [Luo92] Q. T. Luong. *Matrice fondamentale et calibration visuelle sur l'environnement : vers une plus grande autonomie des systèmes robotiques*. Thèse de doctorat, Université Paris-Sud, Orsay, Décembre 1992.
- [Lux85] A. Lux. *Algorithme et contrôle en vision par ordinateur*. Thèse de doctorat, INPG Grenoble, 1985.
- [Mar82] D. Marr. *Vision : a computational investigation into the human representation and processing of visual information*. W. H. Freeman, San Francisco, 1982.
- [Mar96] E. Marchand. *Stratégies de perception par vision active pour la reconstruction et l'exploration de scènes statiques*. Thèse de doctorat, Université de Rennes I, Juin 1996.
- [MBFM05] J. F. Menudet, J. M. Becker, T. Fournel, and C. Mennessier. Auto-calibrage plan d'une camera : une interpretation simple des contraintes géométriques. In *Proceedings of ORASIS 2005, journées francophones des jeunes chercheurs on vision par ordinateur, Clermont-Ferrand*, Proceeding sur CD–Rom, article n°12, 2005.
- [MF81] J. E. W. Mayhew and J. P. Frisbay. Psychophysical and computational studies towards a theory of human stereopsis. *Artificial Intelligence*, 17, pages 349–385, 1981.
- [Mor96] V. Moron. *Mise en correspondance de données 3D avec un modèle CAO : Application à l'inspection automatique*. Thèse de doctorat, INSA de Lyon, 1996.
- [Nag99] H. H. Nagel. Natural language description of image sequences as a form of knowledge representation. In W. Burgard, T. Christaller, and A. B. Cremers, editors, *Proceedings of 23rd Annual German Conference on Artificial Intelligence (KI-99), LNCS 1701, Springer-Verlag, Berlin*, pages 45–60, Septembre 1999.
- [NIK91] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflexion : physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7), pages 611–634, 1991.
- [Ois98] L. Oisel. *Reconstruction 3D de scènes complexes à partir de séquences vidéo non calibrées : estimation et maillage d'un champ de disparité*. Thèse de doctorat, Université de Rennes I, 1998.
- [OK85] Y. Ohta and T. Kanade. Stereo by intra and inter scanline search using dynamic programming. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 7(2), pages 139–154, 1985.

- [Onk89] L. Onken. *Entwurf eines Verfahrens für eine Schaltung mit Spezialprozessoren zur schnellen Verkettung von Konturpunkten und Beschreibung von Konturpunkten*, Diplomarbeit. Thèse de doctorat, Fraunhofer-Institut Für Informations- und Datenverarbeitung (IITB), Karlsruhe, Novembre 1989.
- [PAGM90] B. Peroche, J. Argence, D. Ghazanfarpour, and D. Michelucci. *La synthèse d'Images*. Hermès, Paris, 1990.
- [Rab05] C. Rabaud. *Une nouvelle approche de mise en correspondance stéréoscopique par méthodes possibilistes*. Thèse de doctorat, Université de Montpellier II, Juillet 2005.
- [RAD] Web site :<http://www.radsite.lbl.gov/radiance/>.
- [Rou99] D. Roussel. *Reconstruction de courbes et de surfaces 3D en stéréo-acquisition*. Thèse de doctorat, Université Paris XI, Janvier 1999.
- [SAB02] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(1), pages 1617–1635, 2002.
- [Som97] L. Sommellier. *Mise en correspondance d'images stéréoscopiques utilisant un modèle topologique*. Thèse de doctorat, Université Claude Bernard Lyon I, Septembre 1997.
- [Stu97] P. F. Sturm. *Vision 3D non calibrée : contribution à la reconstruction projective et études des mouvements critiques pour auto-calibrage*. Thèse de doctorat, Institut National de Polytechnique de Grenoble, Décembre 1997.
- [Tar96] J. P. Tarel. *Estimation géométrique et appariement en modélisation automatique*. Thèse de doctorat, Université Paris IX Dauphine, Juin 1996.
- [TG95] G. H. Tarbox and S. N. Gottshlich. An integrated volumetric inspection system. *Computer Vision and Image Understanding*, 61(3), pages 430–444, 1995.
- [Tor02] P. H. S. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. *International Journal of Computer Vision*, 50(1), pages 35–61, 2002.
- [Tos87] G. Toscani. *Système de calibration optique et perception du mouvement en vision artificielle*. Thèse de doctorat, Université d'Orsay, Décembre 1987.
- [TS67] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9), pages 1105–1114, 1967.
- [Tsa86] R. Y. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, Miami Beach, USA*, pages 364–374, 1986.
- [Tsa87] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4), pages 323–344, 1987.

-
- [Tsa91] R. Y. Tsai. Synopsis of recent progress on camera calibration for 3D machine vision. In *Autonomous Mobile Robots : Perception, Mapping, and Navigation (Vol. 1)*, pages 191–200. IEEE Computer Society Press, 1991.
- [VG91] J. M. Vezien and A. Gagalowicz. Reconstruction 3D en régions d’une paire d’images stéréoscopiques. In *Proceedings of Reconnaissance des Formes et Intelligence Artificielle, Lyon, France*, pages 649–660, 1991.
- [Via92] M. Viala. *Contribution à la reconstruction de scènes constituées d’objets cylindriques et polyédriques à partir d’une séquence d’images acquises par une caméra en mouvement*. Thèse de doctorat, Université d’Orsay, Novembre 1992.
- [WCR92] J. Weng, P. Cohen, and N. Rebiho. Motion and structure estimation from stereo image sequences. *IEEE Transactions on Robotics and Automation*, 8(3), pages 362–382, 1992.
- [Wil94] R. Willson. *Modeling and Calibration of Automated Zoom Lenses*. Thèse de doctorat, Carnegie Mellon University, 1994.
- [Xie95] M. Xie. Cooperative strategy for matching multi-level edge primitives. *Image and Vision Computing*, 13(2), pages 89–99, 1995.
- [XJ03] A. Xavier and S. Joaquim. Overall view regarding fundamental matrix estimation. *Image and Vision Computing*, 21(2), pages 205–220, 2003.
- [ZDFL95] Z. Zhang, R. Deriche, O. Faugeras, and Q. T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1), pages 87–119, 1995.
- [Zha89] J. Zhao. *Extraction d’information tridimensionnelle par stéréovision*. Thèse de doctorat, Université Paul Sabatier, Toulouse, 1989.
- [Zha96] Z. Zhang. Determining the epipolar geometry and its uncertainty : A review. *International Journal of Computer Vision*, 27(2), pages 161–198, 1996.
- [Zha99] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proceedings of International Conference on Computer Vision, Corfu, Greece*, pages 666–673, Septembre 1999.
- [Zha00] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), pages 1330–1334, 2000.
- [Zho00] D. Y. Zhou. *Application à la comparaison d’images réelles et conceptuelles à l’extraction contrôlée d’indices images et à la métrologie dimensionnelle*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, 2000.
- [ZLF93] Z. Zhang, Q. T. Luong, and O. Faugeras. Motion of an uncalibrated stereo rig : self calibration and metric reconstruction. Technical Report 2079, INRIA, 1993.
- [ZW94] R. Zabih and J. Woodll. Non-parametric local transforms for computing visual correspondance. In *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, volume II, pages 151–158, 1994.

Liste de publications

Conférences internationales avec actes et comité de lecture

Aicha Beya FAR, Sophie KOHLER and Ernest HIRSCH : « 3D reconstruction of manufactured parts using bi-directional stereovision-based contour matching and comparison of real and synthetic images ». *IAPR Conference on Machine Vision Applications, Tsukuba Science City, Japan* , pages 456-459, May 16-18, 2005.

Sophie KOHLER, **Aicha Beya FAR** and Ernest HIRSCH : « Intelligent approaches for computer vision-based quality control ». *QCAV 2005 : 7th International conference on Quality control by Artificial Vision in Nagoya, Japan*, pages 247-252, May 18-20, 2005.

Conférence nationale avec actes et comité de lecture

Aicha Beya FAR, Sophie KOHLER et Ernest HIRSCH : « Reconstruction 3D de pièces industrielles à partir d'une segmentation et d'un appariement robuste d'images stéréoscopiques ». *ORASIS 2005, journées francophones des jeunes chercheurs en vision par ordinateur, Fournol, France*, Proceedings sur CD-Rom, article N°35, 24-27 Mai 2005.

Annexes

Annexe A

Étalonnage d'une caméra

Cette annexe a été rédigée à partir de la thèse de Zhou [Zho00].

Introduction

La détermination des paramètres d'étalonnage d'une caméra est un moyen simple pour établir un lien entre les coordonnées d'un pixel dans l'image et les coordonnées du point correspondant dans l'espace tridimensionnel. Le soin avec lequel cette étape est réalisée est crucial pour l'étape de comparaison entre images réelles et représentations conceptuelles, en particulier en ce qui concerne la précision de la comparaison.

Les paramètres d'étalonnage forment deux sous-ensembles. Le premier est appelé le groupe des paramètres intrinsèques. Ceux-ci décrivent la géométrie interne de la caméra et certaines caractéristiques de l'optique associée. Le deuxième ensemble est constitué des paramètres extrinsèques qui permettent d'exprimer la position dans l'espace de la caméra par rapport à un repère de référence, par exemple le repère de la scène dans laquelle est placé l'objet à évaluer.

La méthode utilisée dans notre laboratoire repose sur une modélisation de la caméra CCD et de son optique associée proposée en particulier par Tsai dans sa suggestion d'une méthode d'étalonnage [Tsa86]. Cette méthode possède l'avantage d'être à la fois précise, rapide et relativement simple à implémenter. On illustre dans la figure 4.22 le modèle de Tsai avec les différents changements de repère à effectuer pour passer d'un point appartenant à l'espace 3D, ici un point de la mire d'étalonnage (centre d'un des cercles du plateau de la figure 4.22) au point correspondant dans l'image.

On note (x_w, y_w, z_w) les coordonnées d'un point P appartenant à la scène tridimensionnelle et (x, y, z) les coordonnées de ce point exprimées dans un repère lié à la caméra et dont l'origine est le centre optique O (l'axe z de ce repère est supposé confondu avec l'axe optique). La distance focale effective f est la distance entre le plan de formation de l'image et le centre optique. La droite passant par les points P et O coupe le plan image en un point P_1 et permet d'obtenir le point image correspondant au point P de la scène. Ce mécanisme correspond à une projection perspective reposant sur un modèle *sténopé* du capteur. Le point image est exprimé dans un repère (X_u, Y_u) contenu dans le plan image et correspond à une projection obtenue avec une optique sans défaut. On peut, par la suite, exprimer ce même point image dans un repère (X_d, Y_d) en tenant compte des éventuels défauts liés à l'optique (distorsion radiale, par exemple). Enfin, le point image est exprimé, après échantillonnage spatial, en unités pixels dans un repère (X_f, Y_f) . Les

paramètres intrinsèques (facteurs d'échelle horizontale et verticale de la matrice CCD, projection du centre optique O dans le plan image, longueur focale effective) doivent être étalonnés si l'on cherche à obtenir une relation quantitative entre les coordonnées (X_f, Y_f) d'un point exprimées dans le repère image et les coordonnées du même point données dans le système de coordonnées lié au capteur (x, y, z) . De même, il faut déterminer les paramètres extrinsèques pour être capable de lier le repère (x, y, z) au repère de référence (x_w, y_w, z_w) .

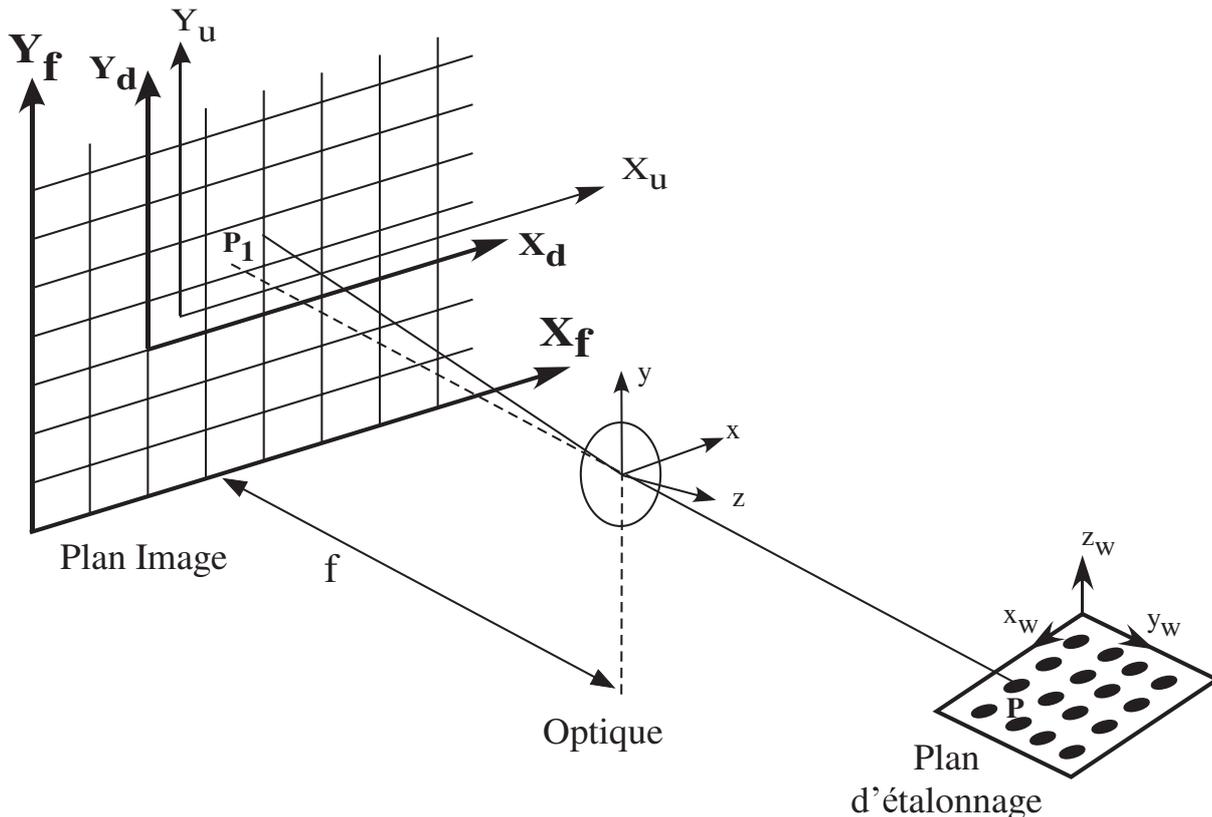


FIG. 4.22 – Modélisation du capteur et changements de repère liés à la méthode d'étalonnage proposée par Tsai [Tsa86]

Le processus de formation d'image peut ainsi être divisé en quatre étapes :

1. Passage du repère (x_w, y_w, z_w) de référence (lié à la scène) au repère caméra (x, y, z)

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} + T \quad (4.1)$$

La matrice de rotation R est classiquement définie par :

$$R = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{pmatrix} \quad (4.2)$$

De même, le vecteur de translation T est défini comme suit :

$$T = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (4.3)$$

Les paramètres à étalonner sont les éléments r_i et t_j des matrices R et T .

2. Passage du repère caméra (x, y, z) au repère image (X_u, Y_u) par application d'une transformation perspective donnée par :

$$\begin{cases} X_u = f \cdot \frac{x}{z} \\ Y_u = f \cdot \frac{y}{z} \end{cases} \quad (4.4)$$

Le paramètre à étalonner est la longueur focale effective f (distance entre le centre optique et le plan image (voir figure 4.22)).

3. Passage du repère (X_u, Y_u) au repère (X_d, Y_d) , en tenant compte des défauts de l'optique du capteur (ici une distorsion radiale due à l'objectif [Len87]) :

$$\begin{cases} X_d = \frac{2X_u}{1+\sqrt{1-4kR^2}} \\ Y_d = \frac{2Y_u}{1+\sqrt{1-4kR^2}} \end{cases} \quad (4.5)$$

avec :

$$R^2 = X_u^2 + Y_u^2 \quad (4.6)$$

Le coefficient de distorsion radiale k est le paramètre à étalonner. Pour les capteurs de bonne qualité associés à une bonne optique, k est habituellement négligeable. Dans ce cas, le repère (X_d, Y_d) , est confondu avec (X_u, Y_u) et il est souvent inutile de déterminer la valeur de ce coefficient.

4. Passage du repère (X_d, Y_d) au repère (X_f, Y_f)

Cette étape permet de rendre compte de la discrétisation (échantillonnage spatial) réalisé par le capteur CCD. Ce processus peut être représenté par les deux relations de l'équation 4.7.

$$\begin{cases} X_f = \frac{X_d}{S_x} + C_x \\ Y_f = \frac{Y_d}{S_y} + C_y \end{cases} \quad (4.7)$$

S_x et S_y correspondent respectivement au facteur d'échelle horizontale et au facteur d'échelle verticale. C_x et C_y sont les coordonnées de projection du centre optique O sur la matrice CCD de la caméra.

Dans le cas où le coefficient de distorsion k est nul, en regroupant les quatre transformations précédentes et en utilisant des coordonnées homogènes, on peut exprimer les coordonnées pixels d'un point image en fonction des coordonnées, dans le repère de référence, du point à l'origine du point image, à l'aide du système d'équations suivant :

$$\begin{pmatrix} kX_f \\ kY_f \\ k \end{pmatrix} = P \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (4.8)$$

avec :

$$P = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{pmatrix} \quad (4.9)$$

P est souvent appelé la matrice de projection du capteur. Ses coefficients sont donnés par les relations ci-dessous :

$$\begin{aligned} p_1 &= f \frac{r_1}{S_x} + C_x r_7 & p_2 &= f \frac{r_2}{S_x} + C_x r_8 & p_3 &= f \frac{r_3}{S_x} + C_x r_9 & p_4 &= f \frac{t_x}{S_x} + C_x t_z \\ p_5 &= f \frac{r_4}{S_y} + C_y r_7 & p_6 &= f \frac{r_5}{S_y} + C_y r_8 & p_7 &= f \frac{r_6}{S_y} + C_y r_9 & p_8 &= f \frac{t_y}{S_y} + C_y t_z \\ p_9 &= r_7 & p_{10} &= r_8 & p_{11} &= r_9 & p_{12} &= t_z \end{aligned} \quad (4.10)$$

Pour la détermination des coefficients p_j , on procède de la manière suivante. Pour chaque élément i d'un ensemble de points d'étalonnage, on établit une correspondance entre ses coordonnées (x_{wi}, y_{wi}, z_{wi}) dans le repère 3D de référence (dans notre cas le repère lié au plateau d'étalonnage montré dans la figure 4.22) et ses coordonnées (X_{fi}, Y_{fi}) dans l'image numérisée acquise par le capteur à étalonner. L'ensemble de ces correspondances, liées par l'équation matricielle 4.8, permet de déterminer les paramètres p_j de la matrice P en utilisant une procédure de minimisation d'une fonction erreur. Ensuite, en utilisant les relations 4.10, on peut en déduire les paramètres de la caméra. La fonction erreur a été définie de la façon suivante : on projette les coordonnées 3D des points s'étalonnage dans le système de coordonnées de l'image numérisée en utilisant les équations décrivant le modèle de la caméra et les valeurs des paramètres d'étalonnage. L'erreur, dans le plan image et exprimée en pixels, est alors définie par la distance euclidienne entre le point

image ainsi défini et les coordonnées mesurées dans l'image pour son correspondant. Le critère à minimiser est la somme des erreurs liées à l'ensemble des points d'étalonnage. L'erreur moyenne d'étalonnage est enfin définie par la moyenne de ces distances, calculées pour tous les points d'étalonnage en tenant compte des valeurs finales des paramètres.

Annexe B

Présentation de CATIA V5

Cette annexe a été rédigée à partir du document de Monsieur Jean Luc Giorgetta de la société SOLEIL [Catb]. Il a été mis à disposition via le site [Cata] du CNRS dans le cadre de la mutualisation des connaissances mis en place par cet organisme.

Introduction

CATIA est un puissant logiciel de CFAO (Conception et Fabrication Assistée par Ordinateur), conçu par Dassault Systems et commercialisé par IBM. Il est très utilisé en aéronautique et en automobile. La version 5 est disponible depuis fin 99. À chaque nouvelle mise à jour (release), de nouveaux produits sont proposés et des fonctionnalités supplémentaires sont ajoutées aux modules existants. La mise à jour 5 est opérationnelle depuis décembre 2000 (on parle de CATIA V5R5). À terme, tous les modules qui existaient dans la version 4 (V4) seront intégrés dans la version 5. Le logiciel fournit une large gamme de solutions intégrées pour couvrir tous les aspects de la conception et de la fabrication. Parmi les nombreuses fonctionnalités de base, on peut citer :

- la conception de pièces,
- l'assemblage,
- le rendu réaliste,
- le dessin interactif et génératif,
- les interfaces DXF/DWG, IGES.

CATIA permet ainsi de concevoir des pièces et des assemblages de pièces directement en 3 dimensions sans dessiner de plan. Lorsqu'on parle de la version 5, on utilise souvent la notion de maquette numérique. Ce terme désigne l'ensemble des données informatiques qui permettent de manipuler un objet aussi bien ou mieux qu'on peut le faire avec une maquette réelle ou un prototype. On peut ainsi tester sa résistance à diverses contraintes, vérifier qu'un sous-ensemble est montable ou démontable, s'assurer que la mobilité des composants les uns par rapport aux autres ne génère pas de collision, etc.

La maquette numérique permet de diminuer les coûts, les délais et d'augmenter la qualité car on évite de passer par une phase de prototypage ou de construction d'une maquette réelle. De plus, les modifications ultérieures sur les pièces sont beaucoup plus faciles à réaliser. Grâce au module de fabrication, on peut simuler l'usinage des pièces sur des machines à commande numérique et on peut générer automatiquement le fichier d'usinage qui est utilisé par la machine numérique réelle.

Généralités

CATIA V5 est un logiciel de CAO 3D volumique et surfacique de nouvelle génération. Il fait appel à des opérations élémentaires paramétriques pour générer les différents objets géométriques, contrairement aux logiciels de la génération précédente qui fonctionnaient strictement à partir d'opérations booléennes (CATIA V4, EUCLID 3).

CATIA V5 est organisé en modules fonctionnels nommés *Ateliers* permettant chacun de créer ou de modifier un type d'objet bien précis. L'architecture simplifiée de CATIA est résumée par le schéma de la figure 4.23.

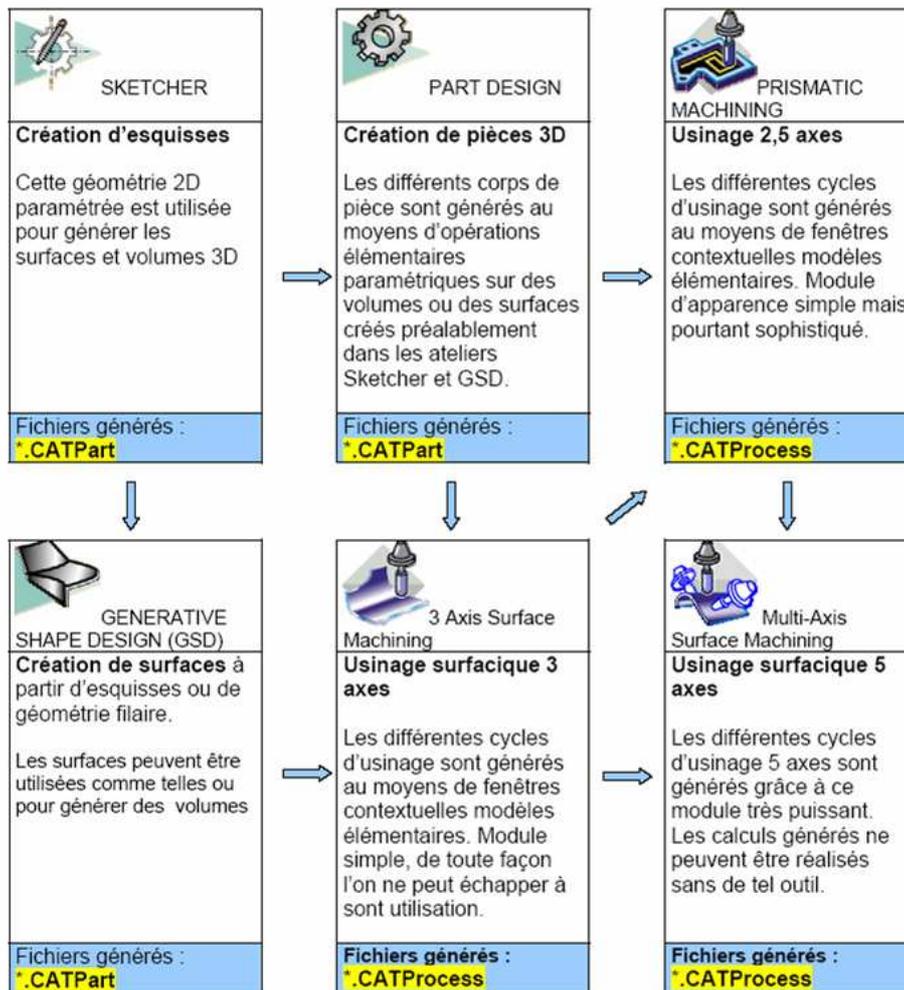


FIG. 4.23 – Architecture simplifiée de CATIA V5 d'après [Cata]

Interface graphique et fonctions de base

L'interface CATIA est présentée dans la figure 4.24. L'atelier *Esquisse* fonctionne selon la méthode *Sketcher* (dessin à main levée) : on crée une géométrie de forme approximative et l'on impose ensuite des contraintes dimensionnelles et géométriques aux différents

éléments pour obtenir un dessin rigoureux. Ces contraintes peuvent être modifiées ou supprimées, et l'on peut utiliser des équations pour lier les cotes entre elles ou à des paramètres extérieurs. La barre d'outils « Outils » permet de choisir les options de fonctionnement.

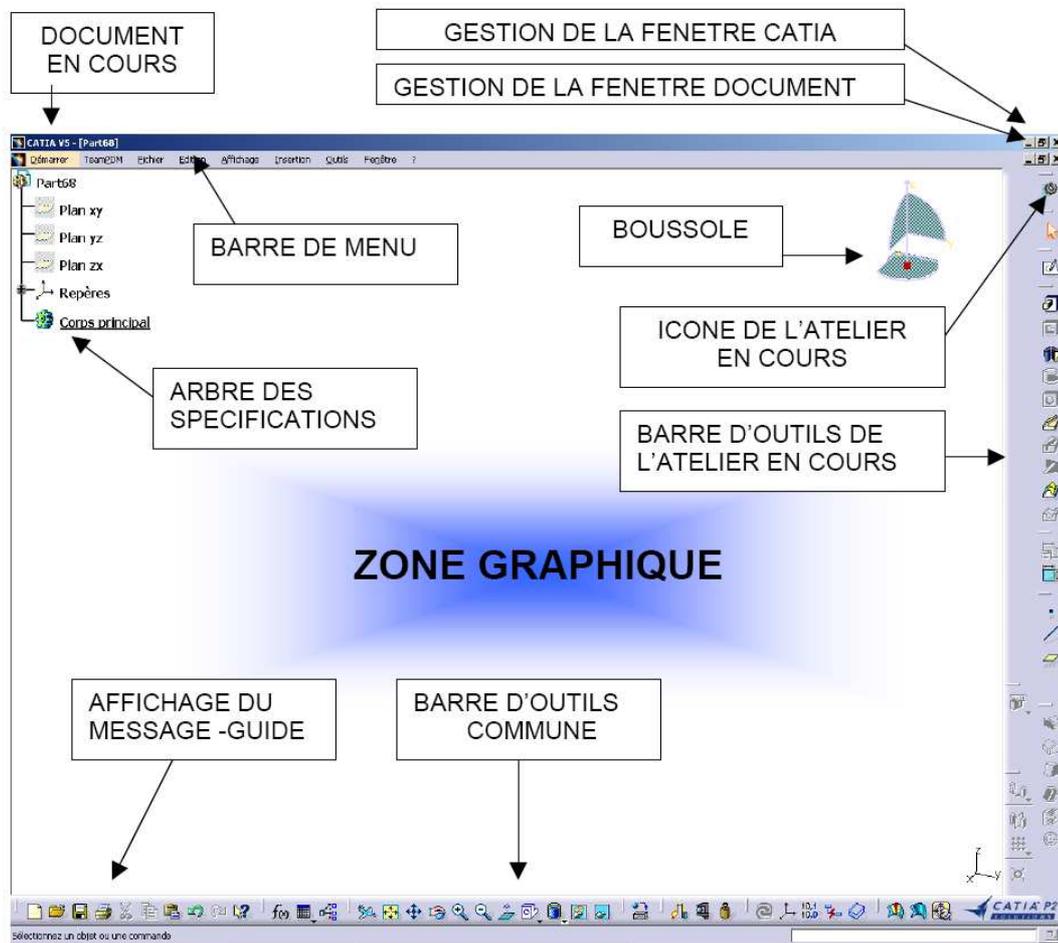


FIG. 4.24 – Interface CATIA V5 d'après [Cata]

Création des contours d'esquisses

Les esquisses constituent la base de départ servant à générer la géométrie 3D. Elles sont constituées d'éléments géométriques (lignes, cercles, courbes, etc.) placés sur un plan 2D. On peut créer autant d'esquisses que l'on veut, dans n'importe quel plan de l'espace, mais il est essentiel de suivre une méthodologie rigoureuse. En effet, le choix des références de départ (plans, directions, repères) est essentiel pour la mise en œuvre et la modification ultérieure des volumes. Il faut respecter les règles de base suivantes :

- Choisir le plan d'esquisse principal de manière à ce que :
 - il coïncide avec l'un des plans de base calé sur le repère absolu de la pièce.
 - il corresponde au plan d'interface de montage de la pièce dans l'assemblage.
- Choisir les directions d'extrusion de façon à définir une géométrie la plus complète possible à partir d'un nombre minimal d'esquisses.
- Pour les autres esquisses, utiliser les plans principaux ou créer d'autres plans si nécessaires.
- Utiliser une face plane de la pièce pour créer une esquisse. Une référence de type *corps surfacique* est alors créée.

Opérations sur la géométrie

On peut effectuer des opérations suivantes sur les tracés existants :

- Délimitation des intersections.
- Transformations géométriques.
- Congés et chanfreins.
- Extraction de géométrie à partir de données 3D.

Mise en place des contraintes

Il existe 2 types de contraintes :

- Les contraintes dimensionnelles.
- Les contraintes géométriques.

Les contraintes peuvent être définies entre les éléments d'une esquisse, entre les éléments et le repère d'esquisse ou entre certains éléments et des références extérieures (2D ou 3D). Lors du travail dans un contexte d'assemblage, il est possible de définir une contrainte d'esquisse par rapport à d'autres pièces : une référence externe est créée si l'option *garder le lien avec l'objet sélectionné* est active. Cette méthode est à éviter, car la gestion des liens dans l'assemblage est délicate à gérer.

Manipulation d'esquisses

Il est possible d'effectuer des manipulations globales sur des esquisses existantes. Pour modifier les éléments d'une esquisse, on « double-clique » sur l'esquisse ou sur son nom dans l'arbre. On bascule alors dans l'atelier esquisse pour éditer la géométrie. Pour changer de support d'esquisse, on effectue un clic droit sur le nom de l'esquisse dans l'arbre pour faire apparaître un menu contextuel, on désigne l'objet esquisse, puis on clique sur *nouveau support d'esquisse* et on désigne le nouveau plan support. Pour remplacer une esquisse par une autre, on clique sur *remplacer* dans le menu contextuel de l'esquisse à remplacer,

puis on désigne une nouvelle esquisse. L'opération générée prendra en compte la nouvelle esquisse.

Pour copier une esquisse il faut effectuer un clic droit sur le nom de l'esquisse dans l'arbre pour faire apparaître le menu contextuel, puis cliquer sur *copier*. Pour coller la copie, 2 méthodes sont possibles :

- Sélectionner par un clic droit dans l'arbre le corps de pièce ou l'opération devant recevoir la copie d'esquisse et cliquer sur *coller*. L'esquisse conserve son support et son orientation.
- Désigner un support (plan ou face) sur une pièce existante. L'esquisse est alors positionnée directement sur le support.

Éléments de référence

Utilisation des éléments de référence. Il existe 3 types d'éléments de référence, à savoir les points, les droites et les plans. Générés dans l'espace 3D indépendamment des esquisses, ils sont utilisés pour servir de base à d'autres éléments et apparaissent dans l'arbre comme *corps surfaciques*. Les principales utilisations sont de servir :

- de point d'origine à un repère local,
- de point à positionner un trou,
- de droite définissant une direction (axe d'un repère local),
- de plan utilisé comme support d'esquisse,
- de plan utilisé pour une opération de coupe.

Création de volumes

Dans CATIA V5, on peut générer des volumes de 2 manières :

- À partir du profil d'un plan quelconque placé dans une esquisse, le volume est obtenu en transformant le profil par une opération d'extrusion ou de révolution autour d'un axe. On parle de composants issus d'un contour. Même les volumes simples doivent être générés de cette manière. Il n'existe pas de fonction spéciale permettant de générer un cube, un pavé, un cylindre, etc. Il faudra créer une esquisse contenant un carré, un rectangle, un cercle et l'extruder.
- À partir d'une surface existante par remplissage ou ajout d'épaisseur.

La première méthode est la plus courante et couvre la majorité des volumes de forme simple. La méthode surfacique ne s'impose que pour concevoir des formes gauches ou des volumes à section évolutive. On peut effectuer des opérations sur des volumes existants, telles que chanfreins, congés, dépouilles, etc. Ces opérations sont nommées composants d'habillage. Normalement, il n'est pas nécessaire d'utiliser des opérations booléennes pour générer les volumes, car toutes les fonctions de base existent en version *ajout* ou *retrait* de matière. Les fonctions booléennes classiques existent. Elles sont utiles lorsque l'on travaille en utilisant différents corps de pièce et peuvent s'avérer utiles dans certains cas particuliers.