



جامعة العربي التبسي - تبسة  
Université Larbi Tébessi - Tébessa

République Algérienne Démocratique et Populaire  
Ministère de l'enseignement supérieur et de la  
recherche scientifique

Université Larbi Tébessi - Tébessa

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie  
Département : Mathématiques et Informatique



كلية العلوم الدقيقة وعلوم الطبيعة والحياة  
FACULTÉ DES SCIENCES EXACTES  
ET DES SCIENCES DE LA NATURE ET DE LA VIE

Mémoire de fin d'études  
élaboré pour l'obtention du diplôme de *MASTER*  
Domaine : Mathématiques et Informatique  
Filière : Informatique  
Option : Systèmes multimédia

Thème

## **RECONSTRUCTION 3D SANS TRIANGULATION**

Présenté Par :  
TOUAHRIA Anouar

Devant le jury :

Mr R. Menassel	MCB	Université Larbi Tébessa	Président
Mr Y. Souli	MAA	Université Larbi Tébessa	Examineur
Mr D. Rouabhia	MAA	Université Larbi Tébessa	Encadreur

# Remerciements

*Je tiens tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui ma a donné la force et la patience d'accomplir ce Modeste travail J'exprime mes profonds remerciements à mon encadreur, Mr D.rouabhia, pour toutes les informations qu'il m'a apporté, pour les conseils qu'il m'a donné, pour son suivi, sa disponibilité, sa sympathie, sa patience et son intérêt porté sur le travail que j'ai réalisé. Je tiens aussi à exprimer mes profonds remerciements au Membres de jury pour leurs efforts et leur l'examinassions et l'évaluation de mon travail.*

*J'adresse mes sincères remerciements à tous les professeurs, intervenants et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté à me rencontrer et répondre à mes questions durant mes recherches. Je remercie très spécialement mes très chers parents qui ont toujours été là pour moi, « Vous avez tout sacrifié pour vos enfants n'épargnant ni santé ni efforts. Vous m'avez donné un magnifique modèle de labeur et de persévérance »*

*Je remercie également mes frères Dia el Haak et Chames El Dieen et ma sœur Rouaida pour leur encouragement.*

*Enfin, je remercie tous mes Ami(e)s que j'aime tant, pour leur sincère amitié et confiance, et à qui je dois ma reconnaissance et mon attachement. À tous ces intervenants, je présente mes remerciements, mon respect.*

# Dédicace

*Dédicaces Je dédie ce travail,*

*A ma chère mère,*

*A mon cher père*

*Qui n'ont jamais cessé, de formuler des prières à mon égard, de me soutenir et de m'épauler pour que je puisse atteindre mes objectifs.*

*A mes chers frères*

*A ma sœur*

*A mes amis*

*Et ceux qui ont partagé avec moi tous les moments d'émotions lors de la réalisation de ce travail, il m'ont chaleureusement supporté et encouragé tout au long de mon parcours.*

*A ma famille et mes proches.*

*Et à ceux qui m'aide et me supporté dans les moments difficiles.*

*A tous mes amis qui m'ont toujours encouragé, et à qui je souhaite plus de succès.*

*A tous ceux que j'aime et ce qui m'aime.*

# **Table des matières**

<b>TABLE DES MATIERES.....</b>	<b>I</b>
<b>LISTE DES TABLEAUX .....</b>	<b>II</b>
<b>LISTE DES FIGURES .....</b>	<b>III</b>
<b>RESUME.....</b>	<b>IV</b>
<b>ABSTRACT.....</b>	<b>V</b>
<b>الملخص.....</b>	<b>VI</b>
<b>INTRODUCTION GENERALE.....</b>	<b>1</b>

## **CHAPITRE01 : RCONSTRUCTION 3D**

<b>Introduction .....</b>	<b>3</b>
<b>1.Définition de la reconstruction 3d : .....</b>	<b>3</b>
<b>2.Systèmes de coordonnées : .....</b>	<b>3</b>
<b>2.1 Coordonnées homogènes.....</b>	<b>4</b>
<b>3.Transformations usuelles :.....</b>	<b>4</b>
<b>3.1 Translation :.....</b>	<b>5</b>
<b>3.2 Changement d'échelle .....</b>	<b>5</b>
<b>3.3 Rotation.....</b>	<b>5</b>
<b>4.Modélisation de la caméra .....</b>	<b>6</b>
<b>4.1Transformation scène-Caméra .....</b>	<b>7</b>
<b>5.Formation de l'image .....</b>	<b>7</b>
<b>6.Relations géométriques entre les images.....</b>	<b>8</b>
<b>6.1. Géométrie épipolaire .....</b>	<b>8</b>
<b>7.Méthodes de modélisation et rendu à base d'images : IBMR .....</b>	<b>9</b>
<b>8.Quelques approches d'IBMR.....</b>	<b>11</b>
<b>8.1. Forme à partir X (Shape from X) .....</b>	<b>11</b>
<b>8.1.1 Forme à partir d'ombres (Shape from shading) .....</b>	<b>11</b>
<b>8.1.2 Forme à partir silhouette (Shape from silhouette).....</b>	<b>12</b>
<b>8.1.3 Forme à partir de textures (Shape from texture) .....</b>	<b>12</b>
<b>8.1.4 Forme à partir du mouvement (Shape from motion) .....</b>	<b>13</b>
<b>8.1.5 Forme à partir Focus/ defocus (Shape from focus/defocus) .....</b>	<b>13</b>
<b>8.2 Les imposteurs .....</b>	<b>14</b>
<b>9.Classifications multicritères des méthodes de reconstruction 3D.....</b>	<b>15</b>
<b>9.1 Niveau d'automaticité.....</b>	<b>16</b>
<b>9.1.1 Reconstruction automatique .....</b>	<b>16</b>
<b>9.1.2 Reconstruction semi-automatique.....</b>	<b>16</b>

9.1.3 Reconstruction manuelle .....	17
9.2 Niveau de complexité de la scène.....	17
9.2.1 Grandes scènes.....	17
9.2.2 Scènes moyenne .....	18
9.2.3 Petites scènes .....	18
9.3 Mouvement de la scène.....	19
9.3.1 Scènes statiques .....	19
9.3.2 Scènes dynamiques .....	19
9.4 Types d'approches utilisées .....	19
9.4.1 Approche ascendante.....	19
10.4.2 Approche descendante .....	19
9.4.3 Approche hybride.....	20
9.5 Passive ou active .....	20
9.5.1 Les méthodes actives .....	20
9.5.2 Les méthodes passives .....	21
9.6 Multiplicité des objets .....	22
9.6.1 Reconstruction mono-objet.....	22
9.6.2 Reconstruction multi-objet .....	22
9.7 Degré de précision .....	23
9.7.1 L'apparence visuelle.....	23
9.7.2 Exigences de description .....	23
9.7.3 Exactitude du modèle géométrique .....	24
9.7.4 Plusieurs aspects à la fois.....	24
9.8 Le degré de généralité de la modélisation 3D .....	24
9.8.1 Reconstruction projective.....	24
9.8.2 Reconstruction euclidienne .....	25
9.8.3 Reconstruction affine.....	25
9.8.4 Reconstruction métrique .....	26
Conclusion.....	26

## CHAPITRE02 : STEREOSCOPIE

Introduction .....	28
1. Définition de la stéréoscopie .....	28
2. La stéréovision .....	28
2.1 Calibration .....	29
2.1.2 Paramètres extrinsèques .....	30

2.1.2.1 Matrice essentielle .....	31
2.1.2.2 Matrice fondamentale.....	31
2.1.2.2.1 Calibrage classique (hors ligne).....	32
2.1.2.2.2 Auto-calibration (en ligne).....	32
2.2 La mise en correspondance.....	32
2.2.1 Primitives stéréoscopiques.....	33
2.2.2 Notion de la disparité .....	33
2.2.2.1 Fonction de disparité.....	33
2.2.2.2 Vecteur de disparité .....	34
2.2.2.2 Carte de disparité.....	34
2.2.3 Problème lié à la mise en correspondance.....	35
2.2.4 Les contraintes de l'appariement .....	37
2.2.4.1 Contraintes géométriques .....	37
Contrainte épipolaire .....	37
Contrainte d'unicité .....	39
Contrainte de continuité de disparité.....	40
Contrainte d'ordre .....	40
2.2.4.2 Contraintes figurales .....	41
Disparité locale constante.....	41
Continuité figurale .....	41
2.2.5 Techniques de mise en correspondance .....	41
2.2.5.1 Méthodes d'appariement denses.....	42
2.2.5.2 Méthodes d'appariement éparses .....	42
2.2.5.3 Méthodes d'appariement locales.....	42
2.3.1 La triangulation.....	44
Conclusion.....	45

### CHAPITRE03 : RECONSTRUCTION 3D SANS TRIANGULATION

Introduction .....	46
1.Système d'acquisition de l'image .....	46
2.Matériels et langage de programmation.....	47
2.1.Programmation avec MATLAB .....	47
3.Démarche proposée du processus de reconstruction 3D.....	47
3.1.Acquisition des images.....	48
3.2.Isolation d'objet de son arrière-plan.....	49
3.3 Application de la stéréoscopie.....	51

<b>3.4 L'estimation de carte de disparité .....</b>	<b>54</b>
<b>3.5 Inférence de la troisième dimension à partir de la carte de disparité.....</b>	<b>55</b>
<b>3.6 Positionnement des points dans l'espace : .....</b>	<b>56</b>
<b>Conclusion.....</b>	<b>58</b>
<b>Conclusion general.....</b>	<b>60</b>
<b>Références.....</b>	<b>61</b>



# Liste des figures

## CHAPITRE 01 : RECONSTRUCTION 3D

Figure [I.1] : Système de coordonnées.....	4
Figure [I.2]: Translation .....	5
Figure[I.3]: Les trois transformations élémentaires du modèle sténopé, et les repères associés .....	6
Figure [I.4] : Fonction de projection de l'image .....	7
Figure[I.5]: Géométrie épipolaire .....	9
Figure[I.6] : Modélisation basée images. (Florian Moreno - Avril 2009 – Portal of Architectural Image-Base Modeling.....	10
Figure[I.7] : Rendu basée image. (Fujita and, Masahiro, Takashi, Kanai - hard ware assiste relief texture mapping).....	10
Figure[I.8] : Résultat des hape from shading. A gauche l'image source, au centre l'image en niveau de gris et le résultat dans l'image de droite.....	11
Figure[I.9]: problème de Shape from shading.....	12
Figure[I.10]: Shape from silhouette partir 3 vues différentes.....	12
Figure[I.11]: Quelques exemples de textures.....	13
Figure[I.12]: (a) Texture extraite d'une robe, (b) surface reconstruite de la texture. . . . .	13
Figure[I.13]: Forme à partir de focus/defocus .....	14
Figure[I.14]: De gauche en droite. Modèle original, ensemble optimal de rectangles, plaquage de texture, Panneau d'affichage rendu.....	15
Figure[I.15]: Exemple d'une voiture reconstruite automatiquement .....	16
Figure[I.16]: Reconstruction semi-automatique: Exemple de détection des vertices et arrêtes pour obtenir le modèle final.....	17
Figure[I.17]: Exemple de reconstruction 3D d'une ville.....	17
Figure[I.18]: Modèle 3D d'un bâtiment service publiques.....	18
Figure[I.19]: Modèle de reconstruction 3D d'une petite scène.....	18
Figure [I.20] : Reconstruction ascendante.....	19
Figure [I.21] : Reconstruction descendante .....	20
Figure [I.22] : Exemple de méthode active de reconstruction 3D, Télémètre Laser.....	21
Figure [I.23] : Exemple de méthode active de reconstruction 3D, Lumière structurée illuminant un objet.....	21
Figure[I.24]: Simple exemple du système des téléovision.....	22
Figure[I.25]: Exemple de reconstruction mono-objet.....	22
Figure[I.26]: Regroupement des éléments reconstruits.....	23

Figure[I.27]: Restitution de l'apparence visuelle d'un édifice.)	23
Figure [I.28] : Exemples de reconstructions descriptives	24
Figure [I.29] : exemple scanne l'objet	24
Figure [I.30] : Résultat d'une reconstruction projective. (Kassem Al Ismaeil - Structure from Motion & Camera Self-Calibration, 2011)	25
Figure[I.31]: Une reconstruction euclidienne à droite	25
Figure[I.32]: Une reconstruction affine à droite	26
Figure[I.33]: Résultat d'une reconstruction métrique. (Kassem Al Ismaeil – Structure from Motion & Camera Self-Calibration, 2011)	26

## CHAPITRE 02: STEREOSCOPIE

Figure[II.1]: Principe de la vision stéréoscopique	28
Figure[II.2]: Exemple d'objet avec fantôme	29
Figure[II.3]: (a) Géométrie épipolaire quelconque. (b) Géométrie épipolaire rectifiée	29
Figure[II.4]: Modèle géométrique de prise de vues	30
Figure[II.5]: Vecteur de disparité	34
Figure[II.6]: a) Structure cubique b) Carte de disparité: Plus proche est plus sombre c) Carte de disparité: Plus le plan focal est sombre	34
Figure[II.7]: Les zones occultées doivent avoir une texture transparente. La profondeur du nouveau coin doit être calculée	35
Figure[II.8]: Zone de discontinuité de profondeur – Les deux pixels d'un point rouge se correspondent mais leurs voisinages différents car ils sont proches d'une discontinuité de profondeur	36
Figure[II.9]: Zone homogène – Manque d'information qui caractérise le pixel	36
Figure[II.10]: Zone avec textures répétitives – Pour un pixel il existe plusieurs candidats dans l'autre image. (D. Scharstein and R. Szeliski - Middlebury Stéréo Vision dataset)	36
Figure[II.11]: l'appariement stéréoscopie	37
Figure[II.12]: Géométrie épipolaire	38
Figure[II.13]: Exemple de la rectification épipolaire. (G. Galles - Pixel matching for binocular stereovision by propagation of feature points matches and region-based randomized voting scheme)	38
Figure[II.14]: Le point P visualisée par deux caméras $F$ et $F'$	39
Figure[II.15]: deux caméras visualisent deux points	40
Figure[II.16]: Contrainte d'ordre	40
Figure[II.17]: Contrainte de continuité figurale	41

Figure[II.18]: Classification possible des méthodes d'appariement.....	42
Figure[II.19]: Mise en correspondance locale.....	43
Figure[II.20]: Système de projection des scène–Représentation de triangle formé par le système.....	44

### CHAPITRE 03 : RECONSTRUCTION 3D SANS TRIANGULATION

Figure [III.1]: Une image utilisée (gauche ;Objet 1) .....	48
Figure [III.2]: Une image utilisée (gauche ;Objet2) .....	49
Figure [III.3]: résultat d'Isolation d'objet 1 .....	50
Figure [III.4]: résultat d'Isolation d'objet2.....	50
Figure [III.5]: résultats de la stéréoscopie de l'objet 1 .....	53
Figure [III.6]: résultats de la stéréoscopie de l'objet 2 .....	53
Figure [III.7]: Cartes de disparité résultantes des objets utilisés .....	54
Figure [III.8]: Cartes de disparité résultantes des objet utilisés.....	55
Figure [III.8]: les résultats de reconstruction 3d sans triangulation de l'objet 1 .....	57
Figure [III.9]: les résultats de reconstruction 3d sans triangulation de l'objet 2 .....	58

# Liste des Tableaux

## **CHAPITRE01 : RECONSTRUCTION 3D**

Tableau [I.1] : Quelques techniques de rendu basées sur l'image..... 10

Tableau [I.2] : Critères de classification des méthodes de reconstruction 3D ..... 16

## **CHAPITRE02 : STEREOSCOPIE**

## **CHAPITRE03 : RECONSTRUCTION 3D SANS TRIANGULATION**

[III.1] : Tableau comparatif de temps de calcul de reconstruction 3d de deux objets..... 58

# RESUME

La reconstruction 3D est un domaine très large et plusieurs méthodes sont proposées. Parmi les méthodes les plus connues nous pouvons citer la stéréoscopie et la triangulation, qui regroupées ensemble permettent de reconstruire un objet à partir d'une paire d'images.

Dans ce mémoire, nous essayons d'utiliser la carte de disparité seule pour inférer les profondeurs perdues lors du processus de formation de l'image. Nous avons détecté que l'aspect visuelle de la carte de disparité donne une idée de la forme de l'objet à reconstruire, la couleur rouge foncé indique que le point en question est proche de la caméra tandis que la couleur bleu foncé indique l'éloignement de la caméra. C'est à partir de ce constat que l'idée est venue à notre esprit, et la question « pourquoi ne pas utiliser ces informations pour reconstruire un objet ? » c'est imposé.

Nous proposons alors de reconstruire partiellement un objet directement à partir de la carte de disparités, sans avoir besoin à utiliser le processus de triangulation gourmand en termes de temps de calcul et d'espace mémoire. Nous avons essayé d'informatiser cette idée, et nous avons effectivement obtenu une reconstruction 3D comme imaginée. Les résultats obtenus sont pas si satisfaisants vue la courte durée consacrée à réaliser notre travail, et il est évident qu'ils peuvent être amélioré par rectification et approfondissement des recherche.

L'essentiel ici, est que nous avons prouvé qu'il est vraiment possible de réaliser une reconstruction 3D sans faire appel à un algorithme complexe tel-que la triangulation, et les résultats obtenus forment bien une preuve constatable pour cette hypothèse. Nous espérons que notre proposition soit considérée comme repère de départ pour d'autres personnes voulant approfondir cette idée par l'élaboration de processus visant à améliorer la qualité visuelle des résultats obtenus

**Mots clés :** Reconstruction 3D, stéréoscopie, triangulation, carte de disparité, l'aspect visuelle.



# ABSTRACT

3D reconstruction is a very large area and several methods are proposed. Among the best-known methods are stereoscopy and triangulation, which are grouped together to reconstruct an object from a pair of images.

In this thesis, we try to use the disparity map alone to infer the depths lost during the image formation process. We detected that the disparity map gives a visual approximation of the real object. For example, the darkest red color says that the considered point is near to the camera. It is from this observation that the idea comes on mind and the question "why not to use this information to reconstruct an object?" has been imposed.

We propose then to partially reconstruct an object directly from the disparity map, without having to use the greedy triangulation process (in terms of computing time and memory space). We tried to represent this idea using a computer, and we actually got a partial 3D reconstruction as imagined. The results are not as interesting as known algorithms due to short the duration of our studies, but it is obvious that they can be improved and deepened.

The bottom line here is that we have proved that it is really possible to perform a 3D reconstruction without resorting to a complex algorithm such as triangulation, and that the results are well established for this hypothesis. The ideas can be deepening by rigorous research in this area.

**Key words:** 3D reconstruction, stereocopy, triangulation, disparity map, visual approximation.

المُلخَص

بعد إعادة البناء ثلاثي الأبعاد مجالاً واسعاً جداً وُقِّتِرح عدة طرق. من بين أفضل الطرق المعروفة ، يمكننا أن نذكر الصورة المجسمة والتثليث ، والتي جمعت معاً لتتيح إعادة بناء كائن من زوج من الصور. في هذه الرسالة ، نحاول استخدام خريطة التباين وحدها لاستنتاج الأعماق المفقودة أثناء عملية تكوين الصورة. لقد اكتشفنا أن الجانب المرئي لخريطة التباين يعطي فكرة عن شكل الكائن المراد إعادة بناؤه ، يشير اللون الأحمر الداكن إلى أن النقطة المعنية قريبة من الكاميرا بينما يشير اللون الأزرق الغامق إلى بعيدا عن الكاميرا. من هذه الملاحظة ، جاءت الفكرة إلى أذهاننا ، والسؤال "لماذا لا تستخدم هذه المعلومات لإعادة بناء كائن؟ إنه مفروض.

ثم نقترح إعادة بناء كائن جزئياً مباشرة من خريطة التباينات ، دون الحاجة إلى استخدام عملية التثليث البحت من حيث وقت الحساب ومساحة الذاكرة. حاولنا حوسبة هذه الفكرة ، وحصلنا على إعادة بناء ثلاثية الأبعاد كما يتصور. النتائج التي تم الحصول عليها ليست مرضية للغاية بالنظر إلى الوقت القصير المكرس لتحقيق عملنا ، ومن الواضح أنه يمكن تحسينها عن طريق تصحيح وتعميق البحث.

خلاصة القول هنا هي أننا أثبتنا أنه من الممكن حقاً إجراء إعادة بناء ثلاثية الأبعاد دون اللجوء إلى خوارزمية معقدة مثل التثليث ، والنتائج التي تم الحصول عليها تشكل دليلاً قاطعاً على هذه الفرضية. نأمل أن يتم اعتبار اقتراحنا كنقطة انطلاق للأشخاص الآخرين الذين يرغبون في تعميق هذه الفكرة من خلال تطوير عمليات لتحسين الجودة المرئية للنتائج.

الكلمات المفتاحية: إعادة البناء ثلاثي الأبعاد، صور مجسمة، التثليث، خريطة التباين.

# **Introduction générale**

## Introduction générale

La création de modèles 3D photoréalistes d'une scène à partir de plusieurs photographies est un problème fondamental en vision par ordinateur et modélisation à base d'images. Pour la plupart des algorithmes de vision par ordinateur, l'accent est mis sur la reconstruction automatique de scènes avec peu ou pas d'intervention de l'utilisateur.

En fait, le processus de reconstruction 3D est bien complexe, et fait intervenir différents algorithmes dont chacun constitue en lui-même un champ de recherche à part entière. La diversité des méthodes et stratégies proposées dans ce contexte fait l'objet d'une large littérature, et les chercheurs visent tous à aboutir à une reconstruction 3D de qualité acceptable en moins de temps possible.

Dans ce mémoire, **nous proposons un système pour la création automatique de modèles 3D partiels en minimisant le nombre d'algorithmes nécessaire**. L'entrée dans notre système est un couple d'images calibrées d'une scène prises à partir de points de vue connus ; et le résultat sera un modèle tridimensionnel de l'objet à reconstruire.

Comme l'indique l'intitulé de notre travail « Reconstruction 3D sans triangulation », nous essayons de proposer une méthode de reconstruction 3D tout en éliminant l'une des phases les plus gourmandes et plus compliquées des processus de reconstruction 3D. Nous asseyons ainsi ; d'extraire les informations de l'objet à reconstruire, et ce, en se contentant juste des informations issues de l'algorithme de stéréoscopie et stockées dans les cartes de disparités.

L'idée est intervenue par le constat que la carte de disparité obtenue après l'application d'un algorithme de stéréoscopie sur une paire d'images, est colorée de façon à indiquer une forme approximative de l'objet à reconstruire malgré le fait qu'elle soit une matrice 2D. Les couleurs présentes dans la carte de disparités donne l'illusion visuelle d'une forme de l'objet, par exemple la couleur rouge la plus foncée indique que le point en question est proche de la caméra ; tandis que la couleur bleu la plus foncée indique que le point appartient à l'arrière-plan de l'objet.

Le principe de la méthode que nous proposons est de traduire cette illusion en informations exploitable par un ordinateur. Ainsi, nous utilisons directement les cartes de disparités pour inférer les profondeurs perdues de l'objet à reconstruire, ce qui veut dire ; la possibilité de repositionnement de chaque point dans l'espace.

A cet effet, nous organisons notre mémoire de la façon suivante :

- Le chapitre 1 est consacré à la reconstruction 3D et ces différentes approches, et dont le but principal est de faire introduire le lecteur à ce domaine large et en pleine expansion.
- Dans le chapitre 2, nous présentons les notions et techniques liées à la stéréoscopie qui fait partie essentielle de notre proposition.
- Le chapitre 3, est réalisé pour justifier l'essai de contourner la phase de triangulation dont nous présentons alors les aspects fondamentaux et le degré de complexité.

- Notre méthode est expliquée dans le chapitre 4, où nous présentons aussi quelques résultats obtenus par son application sur des images réelles.

Finalement, nous concluons notre travail par une conclusion générale et un ensemble de ces futures perspectives.

# **CHAPITRE 01 :**

# **RECONSTRUCTION 3D**



## Introduction

La vision par ordinateur est utilisée pour rendre une machine capable d'accomplir des tâches similaires à la vision humaine. Le processus de vision dépend de deux étapes importantes, la première étant l'acquisition des images, qui sont généralement capturées par des caméras et représentées par une matrice à deux dimensions, chaque élément de cette matrice étant appelé pixel. La deuxième étape est le traitement de ces images.

Dans notre travail, nous nous intéressons au traitement d'images afin de restaurer la troisième dimension perdue lors de la prise de vue. C'est une reconstruction en trois dimensions qui est un domaine de recherche très actif. Plusieurs moyens sont proposés pour atteindre cet objectif, chacun étant dédié à un type spécifique d'objets à reconstruire. Dans ce mémoire, nous présentons une tentative de reconstruction d'images en trois dimensions sans passer par la phase de triangulation.

Contrairement à la synthèse d'images traditionnelle, les techniques de modélisation et de présentation à base d'images (IBMR) sont utilisées pour reconstruire la forme 3D d'objets réel. Ces méthodes ont beaucoup attiré l'attention de la communauté graphique ces dernières années. Le principe général des méthodes IBMR consiste à utiliser les informations contenues dans les photographies pour la reconstruction 3D.

Dans ce chapitre, nous allons présenter les différentes techniques utilisées dans le domaine de la reconstruction 3D. Ce passage est essentiel pour comprendre la démarche que nous proposons. Nous présenterons dans la première partie plus de détails sur ces méthodes ; à savoir : le processus, les méthodes les plus célèbres de la littérature, etc. Dans la deuxième partie, nous présenterons une classification des méthodes de reconstruction tridimensionnelles

### 1. Définition de la reconstruction 3d :

En vision par ordinateur et en infographie, la reconstruction 3D consiste à capturer la forme et l'apparence d'objets réels. Ce processus peut être accompli par des méthodes actives ou passives. Si le modèle change de forme dans le temps, on parle de reconstruction non rigide ou spatio-temporelle [1].

### 2. Systèmes de coordonnées :

Un système de coordonnées est caractérisé par trois axes perpendiculaires. Les cas les plus courants dans les systèmes de coordonnées qui déterminent la position de la caméra et de l'objet, se base sur les emplacements de référence de la caméra, de l'objet et de l'image. Il existe trois types différents de références [2] :

- Les coordonnées de l'objet sur lequel la scène est projetée  $R_s (O_s, X, Y, Z)$ .
- Les coordonnées de la caméra  $R_c (O_c, x, y, z)$ .
- Les coordonnées de l'images  $R_i (O_i, u, v)$ .

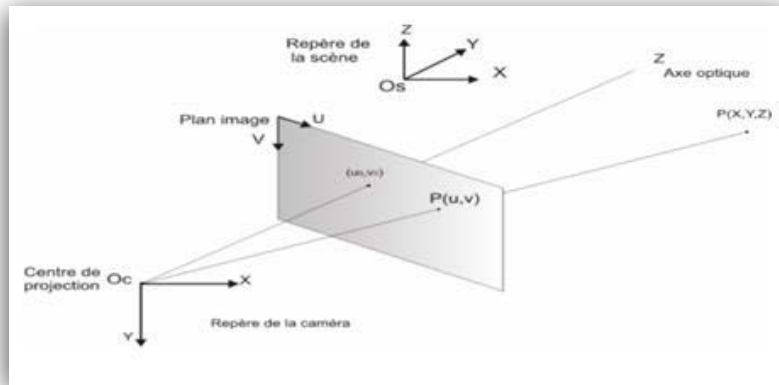


Figure [I.1] : Systèmes de coordonnées.

**2.1 Coordonnées homogènes**

Les coordonnées homogènes [3] ajoutent une coordonnée supplémentaire aux points, elles reposent sur une notation dans laquelle les vecteurs en N dimensions sont représentés par un vecteur en N+1 dimensions. Les systèmes de coordonnées homogènes permettent de définir une transformation entre deux référentiels avec une simple multiplication matricielle  $p=Tp'$ [5], On doit être capable de représenter les objets en 3D, décrire la pose de l'objet 3D, c'est-à-dire :

- sa position.
- son orientation.

En vision par ordinateur, on utilise souvent les coordonnées [4] homogènes qui ajoutent une coordonnée supplémentaire aux points[6].

Point :  $ph = [p_x , p_y , p_z , w]^t$ , avec  $w \neq 0$

Vecteur :  $v_h = [v_x , v_y , v_z , 0]^t$

$$\boxed{M} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \Rightarrow \tilde{M} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

coordonnées euclidiennes
coordonnées homogènes

Un point de l'espace projectif à n dimensions,  $\wp^n$ , est représenté par un vecteur de dimension n+1 de coordonnées  $x = [x_1 , \dots, x_{n+1}]^T$ , avec au moins une des coordonnées  $x_i$  non nulle.

Où  $x_i$  est appelée coordonnée homogène ou projective, X est appelé vecteur coordonnées.

**3.Transformations usuelles :**

Les transformations usuelles sur lesquelles on s'appuie pour déplacer un objet sont les suivantes [3]:

**3.1 Translation :**

Une translation 3D déplace un ensemble de points d'une certaine distance, dans une certaine direction. La matrice traduisant une translation dans l'espace est écrite sous la forme d'un vecteur T (dx, dy, dz):

$$p' = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & dx \\ 0 & 1 & 0 & dy \\ 0 & 0 & 1 & dz \\ 0 & 0 & 0 & 1 \end{bmatrix} = PT$$

$$T(d_x, d_y, d_z) = \begin{bmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

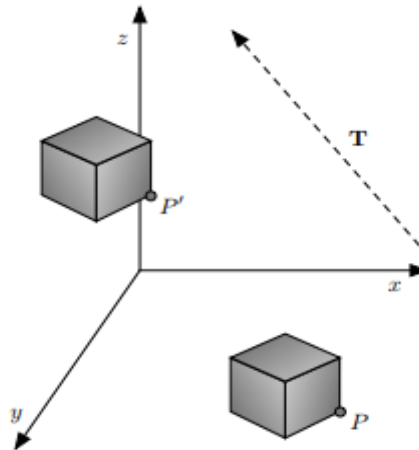


Figure [I.2] : Translation.

### 3.2 Changement d'échelle

Un changement d'échelle 3D modifie la taille d'un objet par rapport à l'origine, c'est une transformation linéaire qui agrandit ou réduit les objets selon un facteur d'échelle identique dans toutes les directions. Le résultat de la mise à l'échelle uniforme est similaire à l'original. Un facteur d'échelle de 1 est normalement autorisé, de sorte que les formes congruentes sont également classées comme similaires. Sa formule en transformation est dans la matrice suivante :

$$P' = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = SP$$

### 3.3 Rotation

En 3D, une rotation fait tourner un ensemble de points (ou d'objets) d'un angle  $\alpha$  autour d'un axe de rotation. Contrairement à la rotation dans le plan, l'axe autour duquel s'effectue la rotation 3D peut être une droite quelconque dans l'espace. Il sera alors nécessaire de décomposer la rotation en trois composantes : des rotations en x, y et z, ce qui donnera lieu à trois matrices de rotation différentes. Les formules en transformation sont données par les matrices suivante :

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P' = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = R_x P$$

#### 4. Modélisation de la caméra

Le modèle sténopé (« pinhole » en anglais) [3, 4] modélise une caméra par une projection perspective. Ce modèle transforme un point 3D de l'espace en un point-image et peut se décomposer en trois transformations élémentaires successives (Figure [I.3]) : la transformation entre le repère du monde et celui de la caméra, la transformation entre le repère caméra et le repère (plan rétinien) et la transformation entre le repère capteur et le repère image.

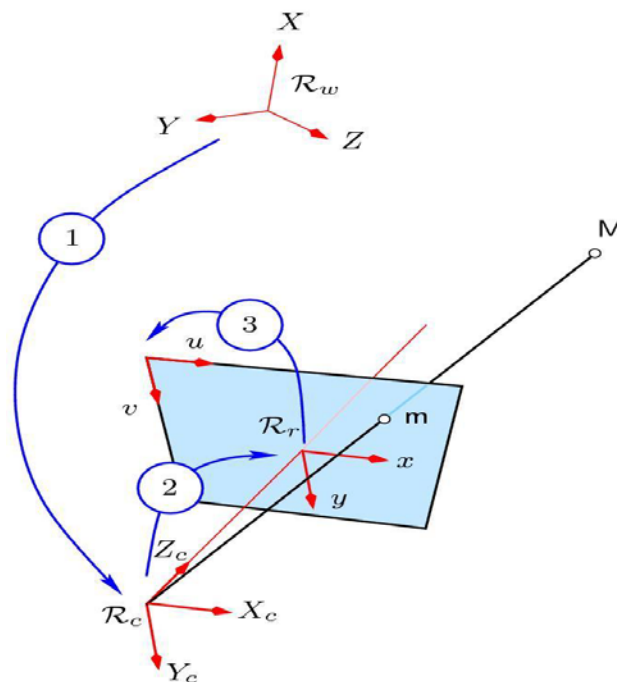


Figure [I.3] : Les trois transformations élémentaires du modèle sténopé, et les repères associés.

##### 4.1 Transformation scène-Caméra

Comme indiqué sur la (Figure [I.3]), représente une transformation entre le repère du monde  $R_w$  (choisi arbitrairement) et le repère caméra  $R_c$  (dont l'origine est situé au centre optique de la caméra). Cette transformation rigide peut se décomposer en une rotation  $[R]$  et

une translation  $[t]$ . Les paramètres de cette transformation sont appelés paramètres extrinsèques de la caméra [5].

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = [R] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + t = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = [T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

T est une matrice  $4 \times 4$ .

$$t = \begin{bmatrix} t_X \\ t_Y \\ t_Z \end{bmatrix}; [R] = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

## 5. Formation de l'image

La formation de l'image correspond à une transformation de projection  $\pi$ , qui à partir du point de la scène  $P (X ; Y ; Z)$  fournit un point image  $p (u ; v)$ . L'étape de calibrage consiste de manière générale à estimer la transformation  $\pi$  telle que  $p = \text{Proj}\pi (P)$ . La littérature propose de nombreuses méthodes permettant d'approximer cette fonction. Elles possèdent en commun [7] :

- Une modélisation de la caméra.
- L'utilisation d'un objet de calibrage.
- L'approximation de la fonction de transfert.

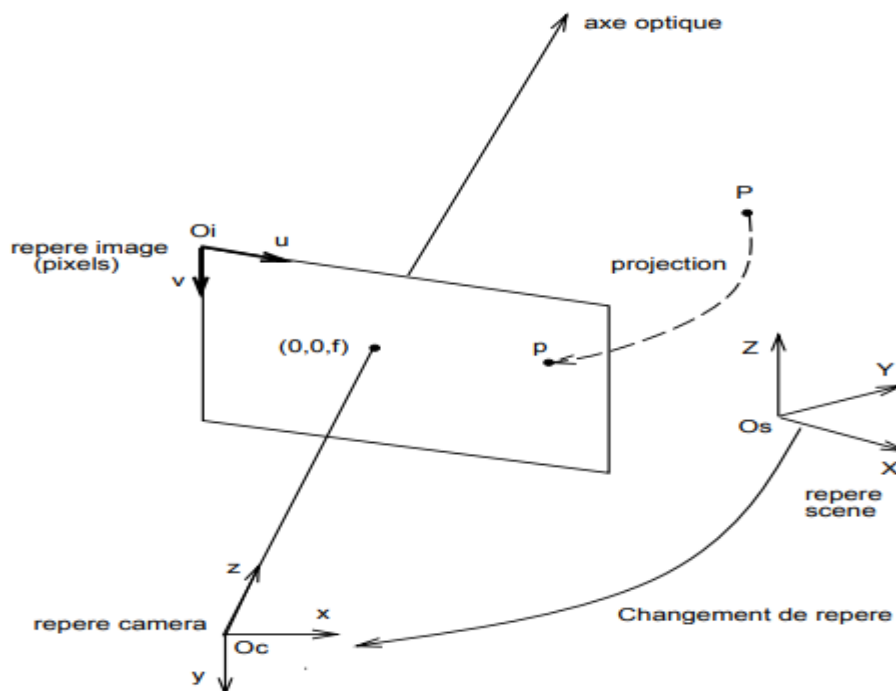


Figure [I.4] : Fonction de projection de l'image.

La formation de l'image passe par trois phases [8] :

- Un déplacement tridimensionnel : les points tridimensionnels exprimés dans un repère de la scène subissent un changement de repère pour passer au repère de la caméra. Ce

changement de repère comporte donc 6 paramètres : 3 pour la rotation et 3 pour la translation. Ces paramètres ne sont autres que la position et l'orientation de la caméra, ils sont appelés paramètres extrinsèques.

- Une projection 3D-2D : après le changement de repère de la phase précédente, les points tridimensionnels exprimés dans le repère de la caméra sont projetés sur le plan image. Les nouvelles coordonnées ainsi obtenues sont appelées coordonnées normalisées (elles sont toujours exprimées dans le repère caméra). Cette projection est une projection perspective dans le modèle de trou d'épingle.
- Changement de coordonnées : pour passer aux coordonnées pixels, les coordonnées normalisées subissent une transformation du plan. Cette dernière, comporte 5 paramètres appelés paramètres intrinsèques de la caméra. La formule de cette transformation peut être écrite par :

$$A = \begin{bmatrix} \alpha_u & -\alpha_u \cot \theta & u_0 \\ 0 & \alpha_v \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Tel que  $\alpha_u$  et  $\alpha_v$  sont deux facteurs d'échelles le long des axes plan image.  $u_0, v_0$  sont les coordonnées en pixel du centre d'image ; c'est à dire les coordonnées du point d'intersection de l'axe optique avec le plan image. Et  $\theta$  représente l'angle entre les deux axes du repère image, et comme cet angle est très proche de la valeur  $\pi$  la forme de la matrice de transformation devient :

$$A = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

## 6.Relations géométriques entre les images.

Les images prises par la caméra ne sont pas nécessairement sont relations. La relation de transformation qui se forme entre la scène dans l'image s'appelle homographie de plans, en plus les transformations scène-images ne sont pas inversibles [9].

### 6.1. Géométrie épipolaire

La géométrie épipolaire est un modèle mathématique de géométrie, qui décrit les relations géométriques de différentes photos du même objet, prises de différents points d'observation. Elle permet de décrire les dépendances entre les pixels en correspondance. Ceux formés par un seul point de l'objet observé sur chacune des images. La géométrie épipolaire concerne la géométrie projective intrinsèque entre deux vues d'une même scène prises de points de vue différents.

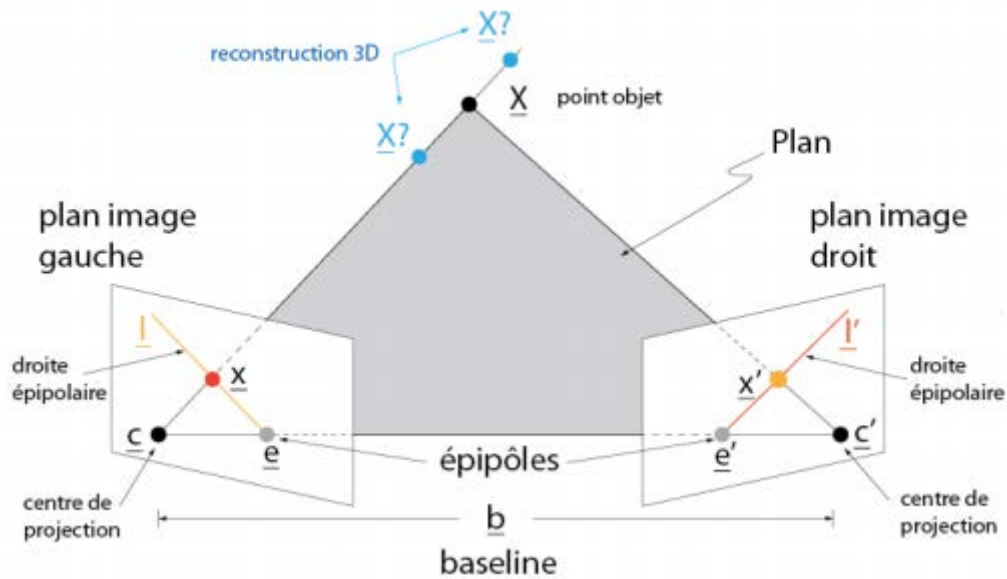


Figure [I.5] : Géométrie épipolaire.

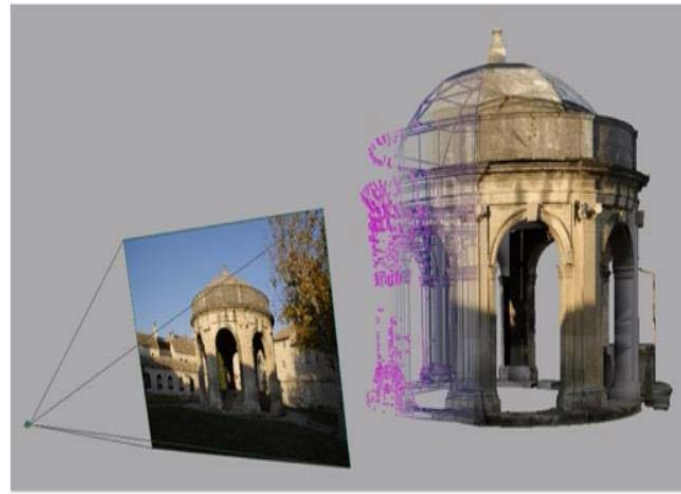
## 7.Méthodes de modélisation et rendu à base d'images : IBMR

Les méthodes de modélisation et de rendu à base d'images (IBMR) [10] reposent sur un ensemble d'images en deux dimensions d'une scène pour générer un modèle en trois dimensions, puis restituer de nouvelles vues de cette scène. La profondeur est perdue au cours de la projection de ces images, et la reconstruction 3D consiste ici à restaurer la dimension perdue.

Il y a ceux qui font la différence entre les méthodes de modélisation et les méthodes de rendu, les termes IBR et IBM sont utilisés respectivement pour (Image Based Rendering et Image Based Modeling).

- L'approche IBM est une technique consistant à prendre un ensemble de photographies d'une scène et à extraire un modèle tridimensionnel (Voir figure I.6).

- L'approche IBR est une technique qui utilise des images pour colorier ou concevoir une texture 3D (Voir figure I.7).



**Figure [I.6] : Modélisation basée images. (Florian Moreno - Avril 2009 - Portal of Architectural Image-Based Modeling).**



**Figure [I.7] : Rendu basée image. (Fujitaand, Masahiro, Takashi, Kanai - hardware assisted relief texture mapping)**

Des échantillons de l'environnement sont capturés sous forme d'ensembles de photographies (ou de séquences vidéo) et rééchantillonnés pendant le rendu.

On peut classifier les techniques de IBMR selon leurs positions relatives le long du spectre géométrie/image (indiqué par un point) sur le diagramme présenté dans la (Voir tableau I.1)

	Images	Hybride	Géométrie
IBR	Panoramas cylindriques. Mosaïques concentriques. Rendu Light-Field. Assemblage Plénoptique.	Images avec profondeur · LDIs · Objets bases image · Textures de relief.	Lumigraph · Light-Field Placage surfacique Light-Field. Enveloppes visuelles



IBM	Méthodes projectives.	Tour dans l'image.	Façade · Coloration Voxel.
-----	-----------------------	--------------------	-------------------------------

**Tableau [I.1] : Quelques techniques de rendu basées sur l'image.**

## 8. Quelques approches d'IBMR

Il existe des méthodes de reconstruction 3D basées principalement sur des images qui transforment les images en trois dimensions et rappellent les plus utilisées dans la reconstruction 3D

### 8.1. Forme à partir X (Shape from X)

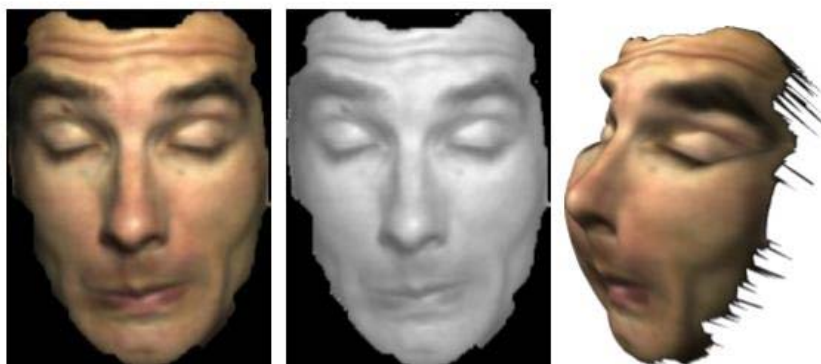
#### 8.1.1 Forme à partir d'ombres (Shape from shading)

L'approche shape from shading [11] est une technique de reconstruction 3D établie en premier temps par Horn au début des années 1970. Ce dernier fut le premier qui a posé ce nouveau concept [12] dont le principe est d'exploiter un ombre pour trouver le relief d'une scène à partir d'une seule photographie, puis estimez la forme de l'objet éclairé à partir des changements d'éclairage observés.

Il est nécessaire que l'image ait une échelle de gris et que toutes les informations concernant le résultat exact et la source de lumière, pour permettre de connaître une position dans le l'espace.

Cela produit des modèles 3D puissants, mais il existe deux problèmes :

- Le premier est que l'algorithme du Shape from shading ne peut donner des résultats en temps réel (Voir figure I.8) [13].
- Le second consiste à calculer la forme tridimensionnelle de la surface à partir des données de luminosité de l'image en noir et blanc de cette surface (Voir figure I.9) [14].



**Figure [I.8] : Résultat de *shape from shading*. A gauche l'image source, au centre l'image en niveau de gris et le résultat dans l'image de droite.**

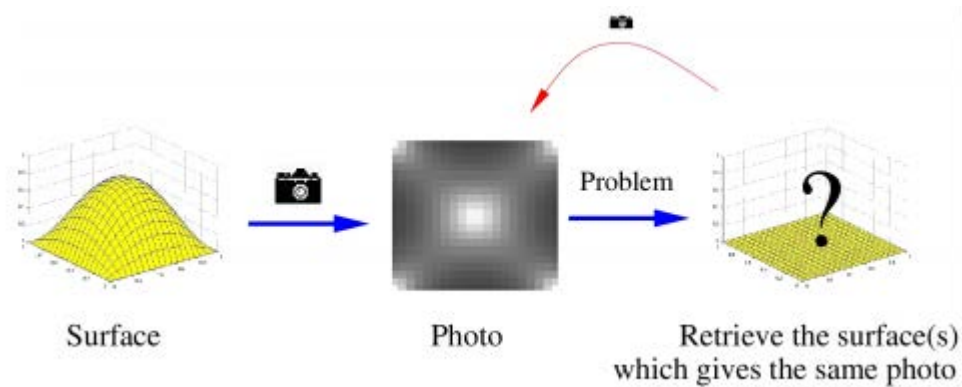


Figure [I.9]: problème de Shape from shading

### 8.1.2 Forme à partir silhouette (Shape from silhouette)

La Shape forme silhouette, consiste à récupérer la forme d'un objet à partir des informations contenues dans une image.

Cette méthode utilise les silhouettes de l'objet pour obtenir la forme 3D [15], et vise à capturer un ensemble d'images à partir de différentes références et points de vues et de les utiliser pour aboutir à la forme de l'objet.

Les algorithmes « Shape from Silhouette (SfS) » calculent la forme 3D (**approximative**) d'un objet à partir de plusieurs projections 2D en ne considérant que le contour de l'objet dans les projections. La classe la plus importante de méthodes SfS sont les algorithmes **Visual Hull**.

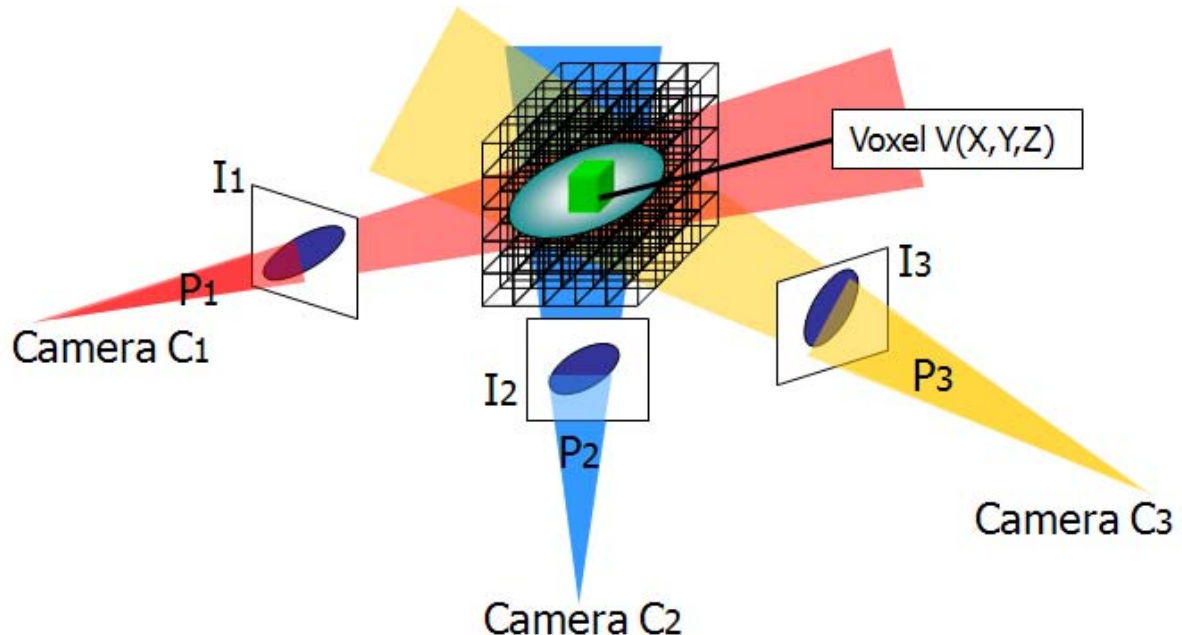


Figure [I.10] : Shape from silhouette partir de 3 vues différentes.

### 8.1.3 Forme à partir de textures (Shape from texture)

La forme à partir d'une texture est une technique de vision par ordinateur dans laquelle un objet 3D est reconstruit à partir d'une image 2D [16]. Bien que la perception humaine soit capable de réaliser des motifs, d'estimer la profondeur et de reconnaître des objets dans une image en

utilisant la texture comme repère, la création d'un système capable d'imiter ce comportement est loin d'être réalisable.

La texture est reconnue par la répétition d'un élément ou l'apparition d'un modèle spécifique sur une surface. Cet élément ou surface est appelé texel (TEXTure ELeMENT).



Figure [L.11] : Quelques exemples de textures.

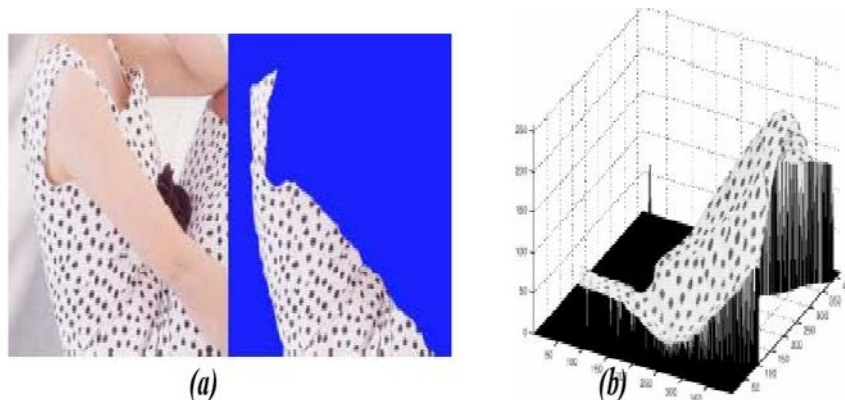


Figure [L.12] : (a) Texture extraite d'une robe, (b) surface reconstruite de la texture.

#### 8.1.4 Forme à partir du mouvement (Shape from motion)

Est de capturer un ensemble d'images d'un objet avec une condition de mouvement [17]. On utilise alors une séquence d'images pour calculer la troisième dimension perdue lors du processus de capture des images.

Le principe des algorithmes « Structure from motion SfM » se basent sur une technique d'imagerie par intervalle photogrammétrique destinée à estimer la structure 3D de quelque chose à partir de la capture de son mouvement dans une période de temps .

#### 8.1.5 Forme à partir Focus/ defocus (Shape from focus/defocus)

La profondeur de la focalisation / défocalisation est le problème de l'estimation de la surface 3D d'une scène à partir d'un ensemble d'au moins deux images de cette scène. Les images sont obtenues en modifiant les paramètres de l'appareil photo (généralement le réglage de la focale ou de la position axiale du plan de l'image) et prises selon le même point de vue (voir la figure ci-dessous).

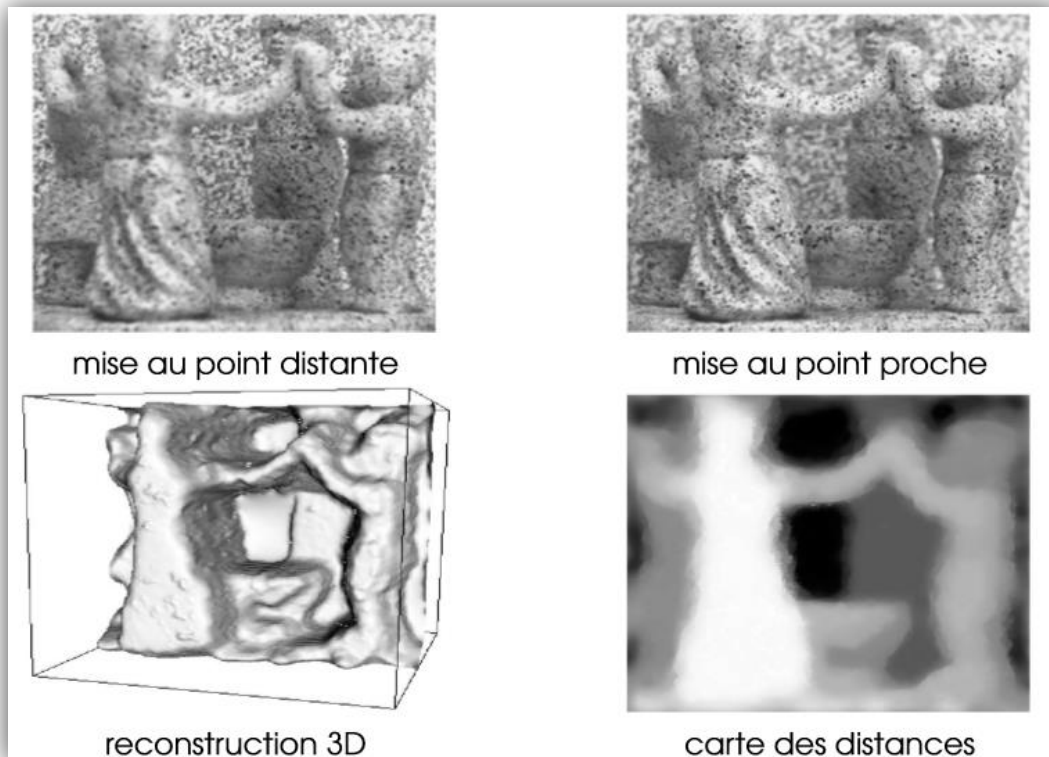


Figure [I.13] : Forme à partir de focus/defocus.

## 8.2 Les imposteurs

Il existe une classe de techniques regroupées sous le terme d'imposteurs [18]. Et cette fonctionnalité dépend principalement de la prévisualisation dans l'écran hors ligne de la part de la scène. L'image créée est utilisée par la suite comme texture plaquée sur un rectangle positionné au sein de la scène, ainsi, à la place de l'objet complexe est affiché son image : un imposteur

Ils sont plus couramment utilisés dans les panneaux d'affichage et les sprites. Typiquement, les billboards sont employés pour représenter des objets tels que des nuages ou des arbres, très difficiles à modéliser explicitement. Ceux toujours parallèles à l'écran ou orientés vers l'observateur servent par exemple pour représenter des nuages ou de la fumée. Un arbre, par contre, est souvent approximé par deux billboards perpendiculaires positionnés dans la scène.

Les sprites se différencient du billboard par l'utilisation qui en est faite et leur mise en œuvre : un sprite est généralement plus petit, en mouvement dans la scène et appliqué directement à l'écran.

En résumé, un billboard est un plan 3D tandis qu'un sprite est un rectangle 2D toujours parallèle à l'écran quelle que soit l'orientation de la caméra.

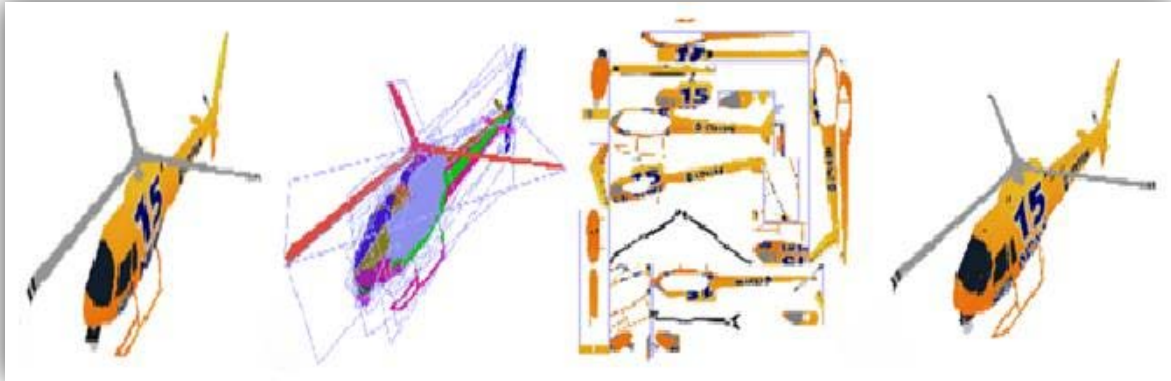


Figure [I.14] : De gauche en droite Modèle original, ensemble optimal de rectangles, plaquage de texture, Panneau d'affichage rendu.

## 9. Classifications multicritères des méthodes de reconstruction 3D

Une classification, proposée dans [19], des différentes approches de reconstructions 3D, permet de choisir la méthode appropriée pour travailler sur un contexte particulier. Pour éviter toute confusion possible, le tableau suivant résume les différentes méthodes présentées dans cette classification selon des critères bien définis :

Critère de classification	Méthodes existants
Niveau d'automatisme	Automatique
	Semi-automatique
	Manuelle
Niveau de complexité de scène	Grande scène
	Scène moyenne
	Petite scène
Mouvement de la scène	Scène statique
	Scène dynamique
Type d'approche utilisée	Approche ascendante
	Approche descendante
	Approche hybride
Passive/active	Méthode passive
	Méthode active
Multiplicité d'objet	Mono-objet
	Multi-objet
Degré de précision	Exigences de description

	Apparence visuelle
	Exactitude du modèle géométrique
	Plusieurs aspects à la fois
Degré de généralité de modélisation 3D.	Projective
	Euclidienne
	Raffine
	Métrique

Tableau [I.2] : Critères de classification des méthodes de reconstruction 3D.

## 9.1 Niveau d'automatisme

### 9.1.1 Reconstruction automatique

L'un des principes de base de la reconstruction en trois dimensions [20] consiste à reconstruire un paysage en trois dimensions en utilisant la machine. Cela dépend de la restructuration de la scène et de sa préparation pour la reconstruction et des matériaux utilisés pour l'acquérir.

La reconstruction purement automatique est purement automatique, et concerne généralement des objets simples.



Figure [I.15] : Exemple d'une voiture reconstruite automatiquement.

### 9.1.2 Reconstruction semi-automatique

Cette méthode nécessite des interventions de la part de l'utilisateur qui devra gérer toutes les tâches à contrôler et automatiser, ce qui nécessite que les points des images ou droites soient clairement précisés pour obtenir un prototype de reconstruction 3D (Voir figure I.21). On peut même utiliser des modèles génériques disponibles à traiter jusqu'à l'obtention d'une bonne estimation du modèle 3D.



**Figure [I.16] : Reconstruction semi-automatique : Exemple de détection des vertices et arrêtes pour obtenir le modèle final.**

### **9.1.3 Reconstruction manuelle**

Cette technique est similaire dans sa reconstruction au classique de modélisation tridimensionnel, mais la reconstruction elle dépend de la conception d'objets sur à base d'objets réels. La reconstruction manuelle dépend de la conception d'objets synthétisés, et est largement utilisée dans le cas d'indisponibilité des informations pertinentes à la reconstruction envisagée.

## **9.2 Niveau de complexité de la scène**

### **9.2.1 Grandes scènes**

La reconstruction 3D de scènes de grande taille nécessite des technologies sophistiquées et prend beaucoup de temps, par exemple, la reconstruction d'une ville ou d'un aéroport ....etc. Il sera nécessaire de mettre en œuvre des algorithmes spécifiques basés sur le plaquage de textures afin de détecter le format 3D, la reconstruction 3D est adoptée pour les scènes de grande taille les satellites, les données GPS ..... etc.



**Figure [I.17] : Exemple de reconstruction 3D d'une ville.**

### 9.2.2 Scènes moyenne

La reconstruction de scènes de taille moyenne nécessite l'utilisation de technologies lui permettant d'acquérir des données à partir d'images et de les assembler pour créer des scènes de reconstruction de taille moyenne, telles que des bâtiments, des petites entreprises, etc.

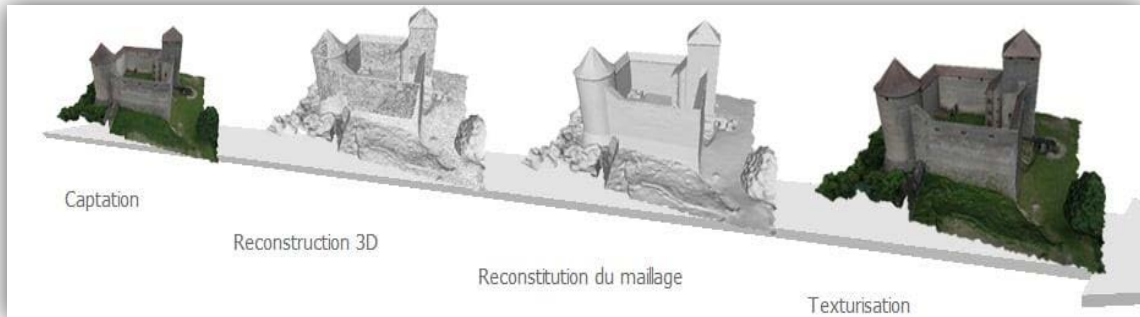


Figure [I.18] : Modèle 3D d'un bâtiment.

### 9.2.3 Petites scènes

Une technique de reconstruction à petite échelle prenant des projections bidimensionnelles de l'objet sur lequel nous appliquons n'est pas difficile, et il faut relativement peu de temps pour collecter et manipuler les informations à partir des images acquises pour restituer l'objet en question.

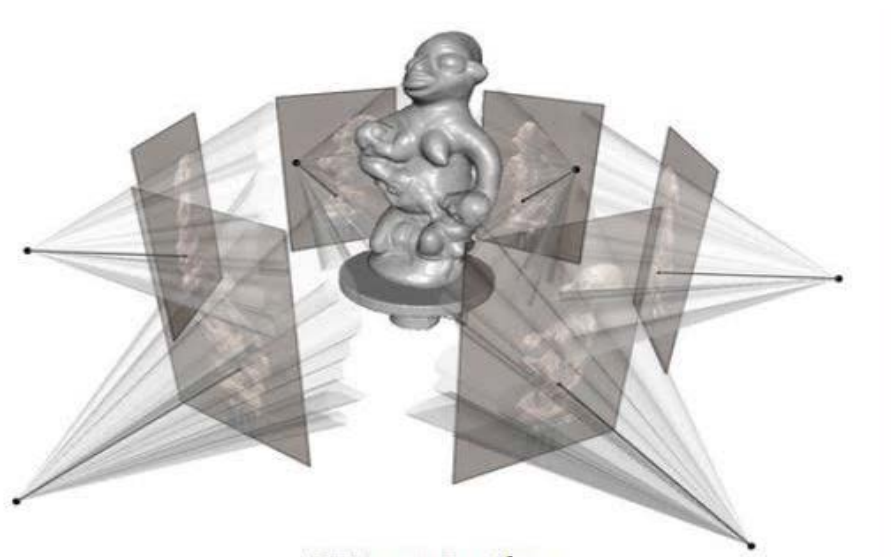


Figure [I.19] : Modèle de reconstruction 3D d'une petite scène.

## 9.3 Mouvement de la scène

### 9.3.1 Scènes statiques

Pour reconstituer une scène toujours immobile, il doit exister certaines conditions pour que l'objet soit statique et sans lien avec le temps. La plupart des études menées dans ce domaine visent principalement à obtenir des résultats réalistes et à partir de photos d'objets stables.

### 9.3.2 Scènes dynamiques

La reconstruction dépend de plusieurs conditions de la scène en mouvement, car il y a suffisamment d'informations sur le mouvement et la vitesse de l'objet et sur la façon de changer de direction.



## 9.4 Types d'approches utilisées

Il existe plusieurs types de plateformes sur lesquelles le système fonctionne. On parle alors d'une approche ascendante (dite bottom-up) ou descendante (dite top-down) et hybride et se caractérise par le principe général de fonctionnement.

### 9.4.1 Approche ascendante

Une approche dite ascendante, ou approche bottom-up (« de bas en haut » en français), se caractérise par une suite de processus qui apportent chacun une partie fondamentale de l'édifice qu'elle cherche à produire, à partir d'éléments de base, il peut s'agir d'une synthèse (ascendante) où l'on part du détail, du « bas », c'est-à-dire l'échelon le plus fin, pour consolider progressivement et opérer une synthèse, il peut s'agir d'un pilotage participatif (ascendant) où le fil directeur de l'animation démarre des perceptions et initiatives de l'échelon le plus « bas » (au sens hiérarchique) ou le plus « terrain » (au sens opérationnel) pour être répercutées, déclinées et prises en compte par les échelons supérieurs .



Figure [I.20] : Reconstruction ascendante.

### 10.4.2 Approche descendante

Une approche dite descendante, ou approche top-down (« de haut en bas » en français), implique des processus qui, à partir d'un apport de matière première brute, visent à forger celle-ci, à la transformer par étapes (raffinement) en vue d'y apporter une valeur ajoutée en augmentant la complexité de l'ensemble.

Par analogie, dans les domaines techniques, cette approche consiste à concevoir le sujet d'études ou le produit dans les grandes lignes, puis, itérativement, à s'intéresser à des détails de plus en plus fins. L'approche descendante permet de délimiter et de conceptualiser rapidement le projet et de le diviser en sous-parties aisément manipulables. Elle permet donc d'avoir une vue globale du projet final et de donner une estimation rapide, bien qu'approximative, de sa complexité et de son coût.



**Figure [I.21] : Reconstruction descendante.**

### **9.4.3 Approche hybride**

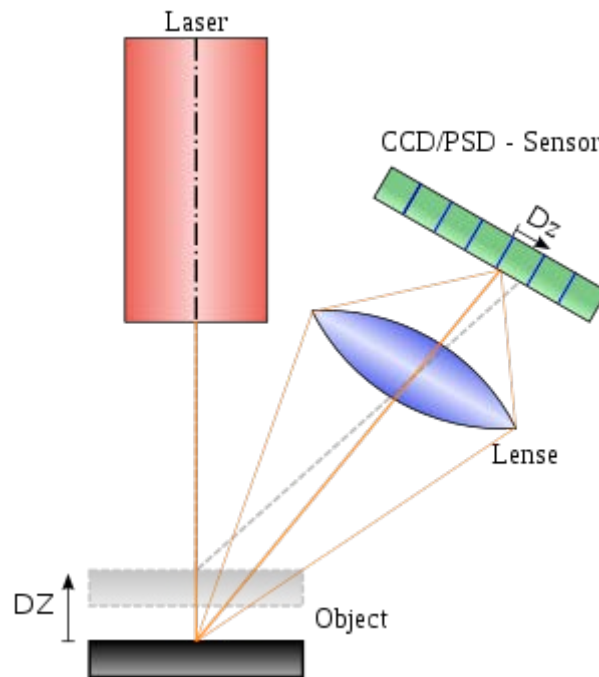
Cette approche utilise les deux approches mentionnées ci-dessus simultanément. D'abord, l'approche ascendante commence par proposer des hypothèses, puis vient le temps de la descendante de s'assurer que les hypothèses sont correctes et de modifier les résultats au cas contraire ou changer les hypothèses.

Le problème de ces approches est qu'il existe des cas où certains objets de la scène (surfaces et limites de surfaces courbe) ne peuvent être décrites par la grammaire de formes basiques, ce qui se reflète sur la forme 3D finale de cette scène.

## **9.5 Passive ou active**

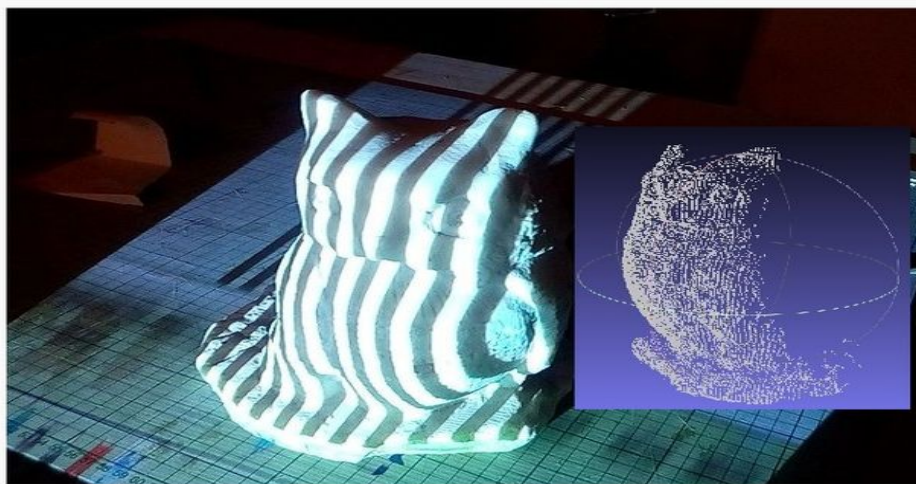
### **9.5.1 Les méthodes actives**

Ces méthodes interfèrent activement avec l'objet reconstruit, de manière mécanique ou radiométrique, à l'aide de télémètres, afin d'acquérir la carte de profondeur, par ex. lumière structurée, télémètre laser et autres techniques de détection active. Un exemple simple de méthode mécanique utiliserait une jauge de profondeur pour mesurer une distance par rapport à un objet en rotation placé sur une plaque tournante. Des méthodes radiométriques plus applicables émettent une luminance énergétique vers l'objet puis mesurent sa partie réfléchie. Les exemples incluent des sources de lumière en mouvement, de la lumière visible colorée, des lasers à temps de vol, des micro-ondes ou des ultrasons [22].



**Figure [I.22] : Exemple de méthode active de reconstruction 3D, Télémètre Laser.**

[23] Un autre type de méthodes actives est la reconstruction basée sur la lumière structurée [24,25]. Cette dernière se réalise à base d'un éclairage actif projeté sur la scène et qui permet d'obtenir un nuage de points 3D dense. Le principe de cette méthode sert à observer les déformations de la scène à l'aide d'une caméra décalée légèrement à un projecteur (La source lumineuse), le résultat de l'éclairage permet de calculer les profondeurs par une triangulation.



**Figure [I.23] : Exemple de méthode active de reconstruction 3D, Lumière structurée illuminant un objet.**

### 9.5.2 Les méthodes passives

Les méthodes passives de reconstruction 3D n'interfèrent pas directement avec l'objet reconstruit, elles utilisent uniquement un capteur pour mesurer la luminance réfléchie ou émise par la surface de l'objet afin d'en déduire sa structure 3D grâce à la compréhension de l'image. En règle générale, le capteur est un capteur d'image comme une caméra qui est sensible à la lumière. Dans ce cas, on parle de reconstruction à base d'images dont le résultat est un modèle

3D. En comparaison avec les méthodes actives, les méthodes passives peuvent être appliquées à un plus grand nombre de situations.

Le principe de la reconstruction 3D avec des méthodes passives [24] consiste à ne pas avoir de contact directe avec l'objet, la scène est alors décrite par des capteurs (tels que des caméras). Les capteurs utilisés dans ce cas sont généralement des caméras. La stéréovision est basée sur l'utilisation de deux caméras qui capturent la même scène, où chaque caméra a un angle particulier. La phase la plus difficile dans la méthode de stéréovision est la mise en correspondance, ce qui la rend inefficace pour reconstruire des scènes complexes.

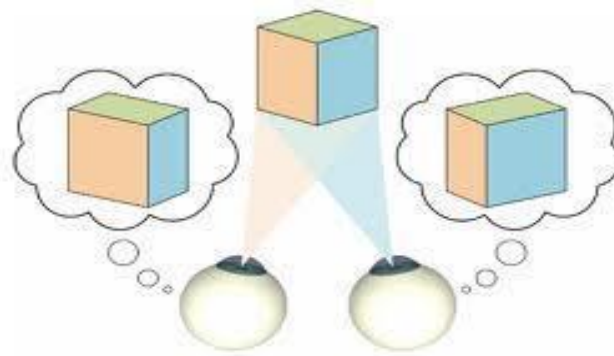


Figure [I.24] : Simple exemple du système de stéréovision.

## 9.6 Multiplicité des objets

### 9.6.1 Reconstruction mono-objet

L'objet de cette opération est de reconstruire la scène une fois pour toutes si elle contient un ensemble d'objets indivisibles, mais cette méthode ne fonctionne pas avec des objets complexes dans leur conception.

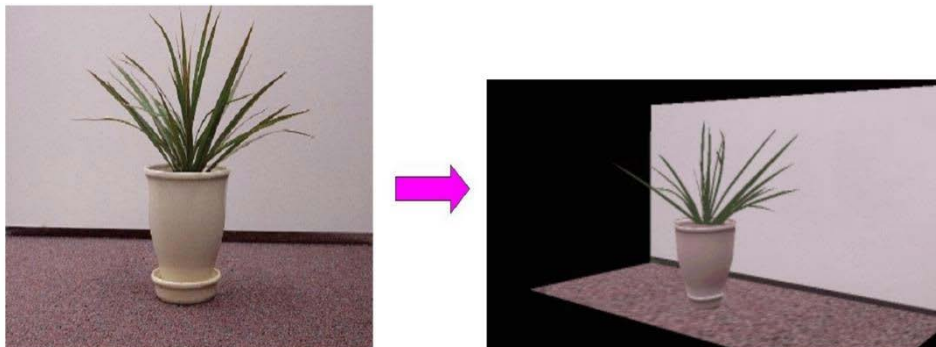


Figure [I.25] : Exemple de reconstruction mono-objet.

### 9.6.2 Reconstruction multi-objet

Pour concevoir une scène complexe de reconstruction 3D, on doit analyser les images, puis séparer les objets de la scène pour faciliter le travail et reconstruire chaque forme et ses caractéristiques indépendamment des autres. Lorsque toutes les formes sont finies individuellement, nous les combinons dans une scène.

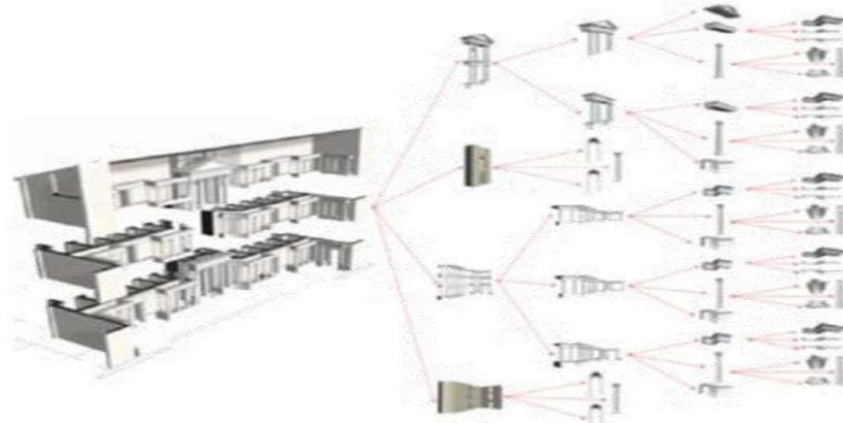


Figure [I.26] : Regroupement des éléments reconstruits.

## 9.7 Degré de précision

### 9.7.1 L'apparence visuelle

Plusieurs techniques de synthèse d'images sont disponibles aujourd'hui pour simuler l'apparence visuelle des formes basées sur le calcul de l'interaction entre sources lumineuses, caractéristiques des matériaux et géométrie des objets. Le but de la restitution de l'apparence visuelle dans un contexte de numérisation d'artefacts patrimoniaux consiste au contraire à restituer les véritables caractères qui décrivent l'aspect des surfaces en cohérence avec la géométrie de l'objet.



Figure [I.27] : Restitution de l'apparence visuelle d'un édifice.

### 9.7.2 Exigences de description

Ces approches permettent de regrouper un ensemble de points liés pour fournir des informations sur la description d'objet d'un aspect spécifique tel que la lecture de la forme, la récupération des caractères constructifs, etc.

Ce type d'approches est caractérisé par le traitement conçu en cohérence avec un objectif de représentation.

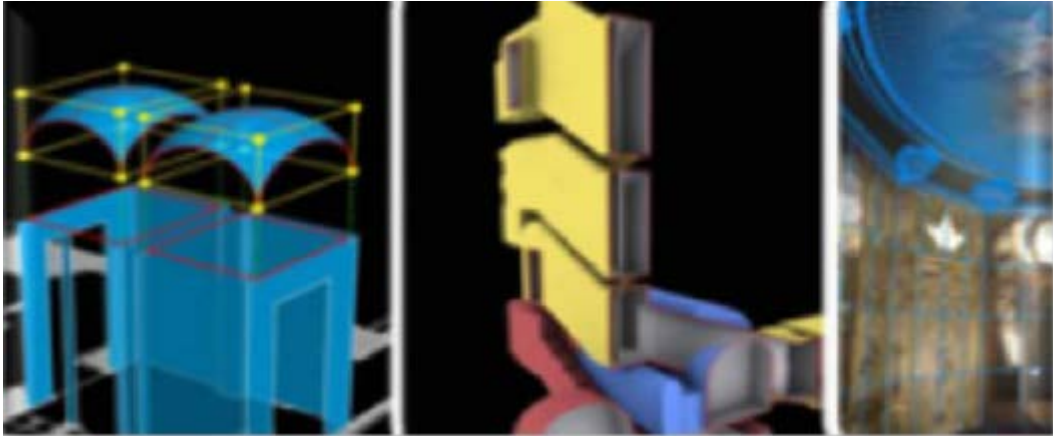


Figure [I.28] : Exemples de reconstructions descriptives.

### 9.7.3 Exactitude du modèle géométrique

Ces méthodes reposent principalement sur une procédure de balayage laser 3D pour le processus de reconstruction 3D. Mais il y a plusieurs problèmes, à savoir la difficulté de calculer la composante de l'objet entier. Il existe également un problème de qualité des surfaces qui rend difficile l'étude de la surface, qui est contenue des artefacts de la géométrie (Des trous, une discontinuité, etc.). Il y a aussi le type de matériau qui forme l'objet qui affecte parfois le faisceau laser, ce qui affecte l'intensité du signal et provoque donc la perte d'informations, empêchant ainsi la traduction des informations vers l'objet souhaité.



Figure [I.29] : exemple scanne l'objet.

### 9.7.4 Plusieurs aspects à la fois

La représentation de l'objet étudié dans ce cas nécessite l'intervention de plusieurs méthodes qui sont utilisées pour la restitution des aspects reliés à cet objet. Les informations métriques et les mesures optiques sont utilisées pour déduire les différents niveaux de complexité d'ingénierie montrés par l'objet étudié et les limites spécifiques du système de représentation choisi.

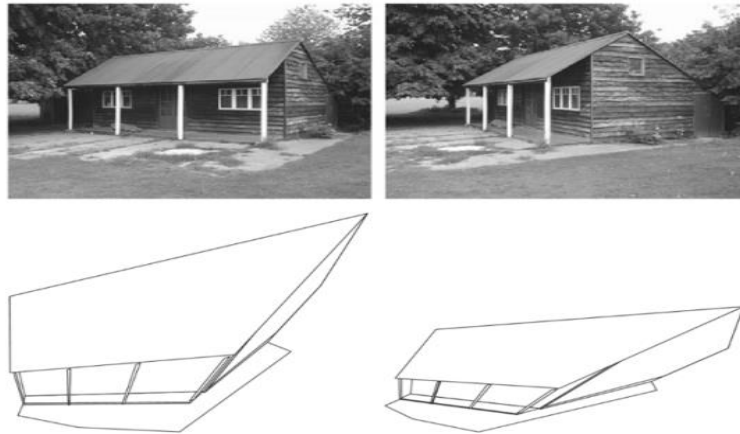
## 9.8 Le degré de généralité de la modélisation 3D

### 9.8.1 Reconstruction projective

La reconstruction de la projection consiste à prendre de nombreuses photos de la scène et les points correspondants dans différentes photos. Ces coordonnées doivent être représentées dans

un cadre de projection. Cette méthode ne repose pas sur la structure de la scène ni sur la direction de la caméra à partir de laquelle les images sont prises pour la reconstruction.

Les points sont alors définis dans un repère projectif, et une telle reconstruction ne contient donc pas d'information critique (absence de la notion d'angle, de longueur, parallélisme ... ).



**Figure [I.30] : Résultat d'une reconstruction projective. (Kassem Al Ismaeil - Structure from Motion & Camera Self-Calibration, 2011)**

### 9.8.2 Reconstruction euclidienne

Quand les points sont définis dans un repère euclidien, la reconstruction est plus riche, elle contient les informations que l'homme a l'habitude de manipuler. On parle dans ce cas de reconstruction euclidienne, car le facteur d'échelle absent ici ne pose pas de problème particulier.



**Figure [I.31] : Une reconstruction euclidienne à droite.**

### 9.8.3 Reconstruction affine

Quand les points sont définis dans un repère affine. Une telle reconstruction contient des informations affines (rapport de longueurs, parallélisme).

Cette méthode est appliquée lorsque les points sont définis dans un repère affine. Mais il est difficile d'estimer ce repère. Pour réaliser cette reconstruction il suffit alors de prendre une quelconque paire de points à l'infini dans les deux images et de reconstruire le point 3D pour obtenir le troisième point qui permet de déterminer le plan à l'infini.

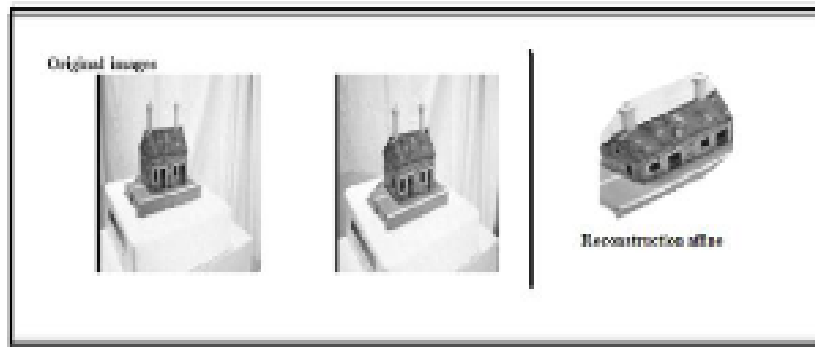


Figure [I.32] : Une reconstruction affine à droite.

#### 9.8.4 Reconstruction métrique

Cette méthode est utilisée lorsque les coordonnées d'un point sont définies dans un repère métrique. Cela fonctionne dans la reconstruction en convertissant une métrique à proximité et contient toutes les informations métriques (Changement d'échelle).

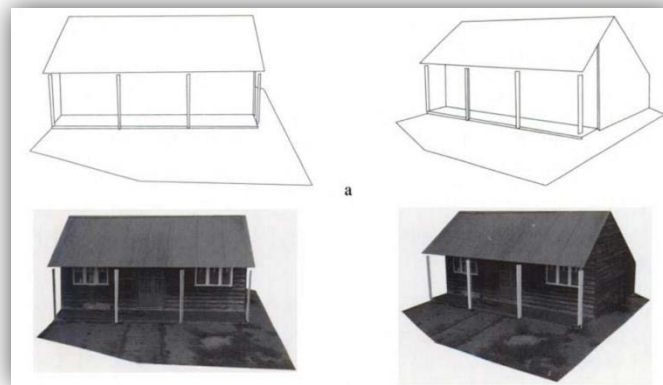


Figure [I.33] : Résultat d'une reconstruction métrique. (Kassem Al Ismaeil - Structure from Motion & Camera Self-Calibration, 2011)

## Conclusion

Dans ce premier chapitre, nous avons présenté quelques concepts de base pour débiter dans le domaine de la reconstruction 3D, et décrire certains des aspects utilisés dans le reste du document. Parmi ces concepts, nous avons présenté la caméra et sa modélisation permettant de décrire un objet filmé dans l'espace et affiché sur une image. Nous avons donc détaillé le processus de formation de l'image.

Nous avons également fourni la relation entre les images prises par l'appareil photo et utilisé cette relation ultérieurement dans le traitement de ces images. Par exemple, extraire des informations. Dans notre cas, les informations seront utilisées pour calculer la troisième dimension par images stéréoscopiques.

L'objectif de notre travail est d'appliquer une reconstruction 3D basée uniquement sur la carte de disparités obtenue par stéréovision. Dans ce chapitre, nous avons également présenté une classification générale de diverses techniques de reconstruction 3D à base de plusieurs critères. Cette classification vise à simplifier la sélection de la méthode d'utilisation, évitant ainsi une mauvaise reconstruction 3D.



Après avoir présenté les généralités de la reconstruction tridimensionnelle, nous présenterons dans le chapitre suivant une description générale de la technique de stéréoscopie.

# **CHAPITRE 02 :**

# **STEREOSCOPIE**

## Introduction

Nous avons vu dans le chapitre précédent que les méthodes IBMR utilisaient les informations disponibles dans les photographies pour calculer la troisième dimension perdue au cours du processus de projection. Ce processus de reconstruction est effectué sur la base des points visibles de l'ensemble des images disposées. Ce problème est suffisamment sensible pour être résolu par une machine et obtenir un résultat proche de la réalité.

La stéréovision est une des méthodes IBMR les plus connues, elle désigne en fait l'ensemble des techniques mises en œuvre pour reproduire une perception du relief à partir de deux images planes. Dans cette section, nous allons d'abord présenter la méthode stéréoscopique et son principe de fonctionnement.

Cette méthode fait partie majeure de notre travail, et sera utilisée ultérieurement dans la mise en œuvre de notre méthode.

L'être humain utilise ces deux yeux pour acquérir une perception du monde réel. En fait, chaque œil perçoit une image bidimensionnelle du monde réel. Le décalage entre les deux images, sera utilisé par le cerveau pour inférer la troisième dimension afin de repositionner la scène acquise dans l'espace.

Ce même principe a été projeté dans le domaine de la vision par ordinateur, ainsi, on utilise deux images de la scène prises de points de vue différents. En appliquant un algorithme de mise en correspondance entre les différents points de l'image, on peut obtenir une matrice qui contient des distances de déplacement des points d'une image à l'autre. Cette matrice est dite matrice de disparité et est utilisée par la suite pour inférer la dimension perdue par triangulation.

Dans ce chapitre, nous essayons de donner une vue générale sur la notion de stéréovision, qui sera utilisé par la suite pour réaliser notre approche. La compréhension des différentes notions liées à la stéréoscopie est primordiale pour aborder notre sujet, et nous essayons alors de les éclaircir au lecteur de façon résumée.

### 1. Définition de la stéréoscopie

La stéréoscopie consiste à élaborer une collection d'images de la même scène prises à partir de points de vue différents et pouvant être distribuées au hasard. Elle vise essentiellement à reconstruire un modèle en trois dimensions de la scène [26].

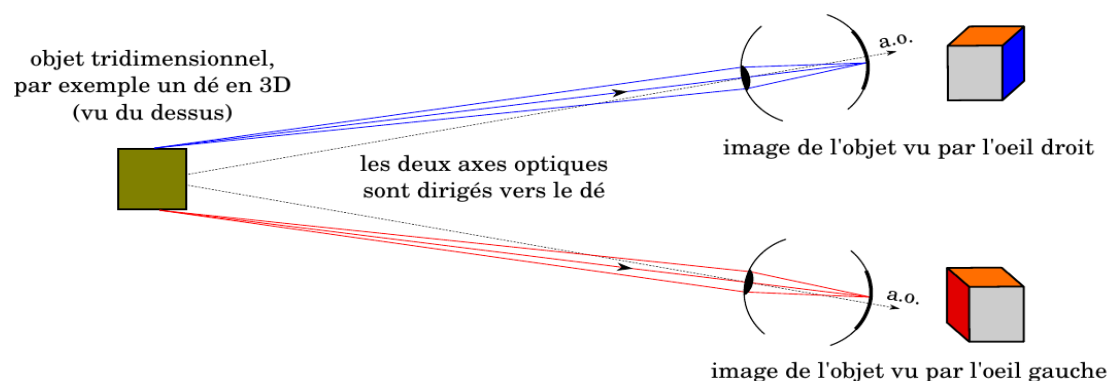


Figure [II.1] : Principe de la vision stéréoscopique.

## 2. La stéréovision

La vision stéréoscopique par ordinateur consiste à extraire des informations 3D à partir d'images numériques, telles que celles obtenues par une caméra. En comparant les informations relatives à une scène à partir de deux points de vue différents, des informations 3D peuvent être extraites en examinant les positions relatives des objets dans les deux panneaux. Ceci est similaire au processus biologique Stéréopsie.

Sur la base de la Stéréovision, on peut récupérer la troisième dimension de la scène [27], ceci est possible en cliquant deux images de deux différents points de vue de la même scène. Étant donné que les caméras seront liées par la projection centrale, et en prenant les rayons lumineux du même pixel ; on obtient des droites dites Epipoles. L'épipole crée une ligne à partir des points  $p$  et  $p'$ , et on utilise cette méthode pour calculer la position  $P''$  dans l'espace par triangulation, tout en utilisant la position relative de l'objet et la direction de la caméra.

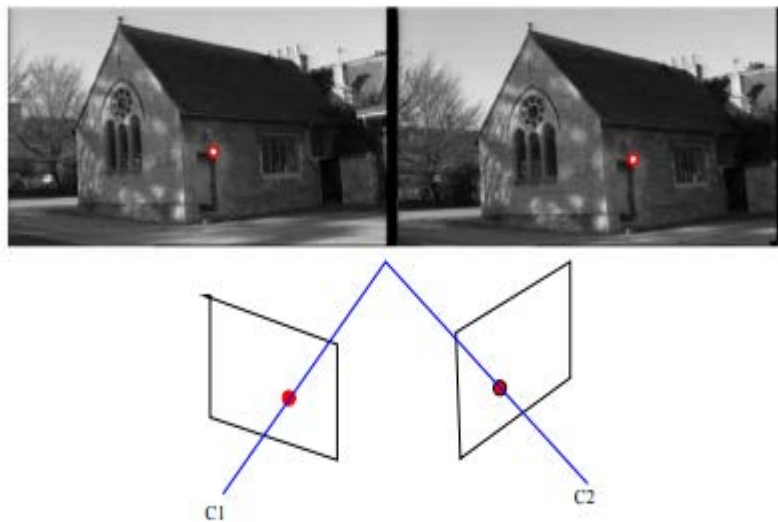


Figure [II.2] : Exemple d'objet avec fantôme.

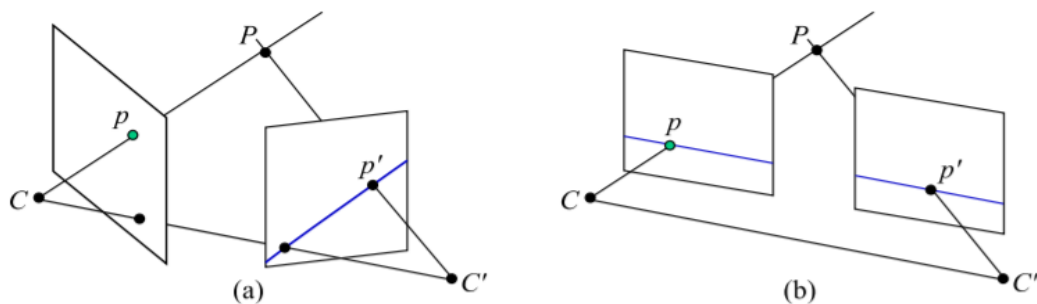


Figure [II.3] : (a) Géométrie épipolaire quelconque. (b) Géométrie épipolaire rectifiée.

La stéréovision s'effectue en trois étapes successives : Calibration, mise en correspondance et enfin la reconstruction 3D [28].

## 2.1 Calibration

La calibration d'un banc stéréoscopique s'intéresse à la connaissance des paramètres intrinsèques K1 et K2 de chacune des caméras ainsi qu'aux positions et orientation relatives des deux caméras [29]. Le but de la calibration est de rendre possible la triangulation afin d'effectuer des mesures sur l'environnement perçu.

La Calibration de la Caméra est très importante pour et constitue une étape importante dans la réussite de la reconstruction 3D, et nécessite une estimation des paramètres du modèle technique de la photographie, dans lequel il existe deux types de paramètres : les paramètres internes et les paramètres externes.

### 2.1.1 Paramètres intrinsèques

Observons les quatre éléments utilisés comme paramètres fondamentaux de la caméra et analysons-les comme suit :

- $\alpha_u$  et  $\alpha_v$  expriment la distance focale, en nombre de pixels (une fois en direction horizontale, une fois verticalement).
- $u_0$  et  $v_0$  sont les coordonnées du point principal, exprimées dans le repère pixels.
- La matrice K est alors :

$$K = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Nous appelons K la matrice de calibrage de la caméra.

### 2.1.2 Paramètres extrinsèques

Dans ce processus, les paramètres de la direction géométrique des caméras liés à la scène (le repère du monde). Les critères utilisés sont les suivants [30] :

- $R_{3 \times 3}$  : qui est la matrice de rotation permettant de passer du repère lié à l'espace de travail au repère lié à la caméra
- $t_x$ ,  $t_y$  et  $t_z$  : qui sont les composantes du vecteur de translation permettant de passer du repère lié à l'espace de travail au repère lié à la caméra.

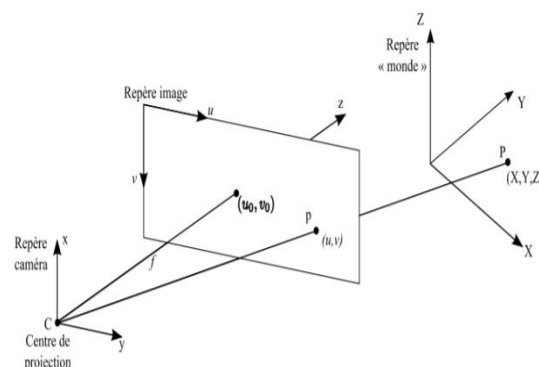


Figure [II.4] : Modèle géométrique de prise de vues.

Au total, cela fait 12 paramètres à estimer (la matrice de rotation R contient 9 éléments, mais la relation qui la définit comme matrice de rotation  $R^T R = 1$  réduit le nombre d'éléments indépendants à 3 : les 3 angles polaires).

Étalonner la caméra consiste à déterminer la valeur numérique des paramètres de ce modèle. Il est toutefois possible de les regrouper de manière différente, suivant la forme sous laquelle ce modèle doit ensuite être exploité, ceci conduit à différentes variantes possibles du modèle.

C'est le passage entre les différents repères (repère monde et repère de caméra) dans lequel les coordonnées sont prises de chaque point de vue observé à partir de ces projections.

### 2.1.2.1 Matrice essentielle

Sur la base de paramètres externes, la matrice décrit les coefficients de transition entre l'étiquette d'objet et l'étiquette de caméra, qui à leur tour représentent la position et la direction des capteurs, qui travaillent en parallèle avec le repère [31].

En substituant  $(x_1, y_1, 1)$  T dans cette expression on obtient :

$$A_2 x_2 + B_2 y_2 + C_2 = 0$$

Avec

$$A_2 = -(b_z r_{21} - b_y r_{31})x_1 - (b_z r_{22} - b_y r_{32})y_1 - b_z r_{23} - b_y r_{33}$$

$$B_2 = -(b_x r_{31} - b_z r_{11})x_1 - (b_x r_{32} - b_z r_{12})y_1 - b_x r_{33} - b_z r_{13}$$

$$C_2 = (b_x r_{21} - b_y r_{11})x_1 + (b_x r_{22} - b_y r_{12})y_1 + b_x r_{23} - b_y r_{13}$$

Ces équations peuvent se mettre sous forme matricielle :

$$\begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_z & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix} = E \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

La droite épipolaire s'écrit :

$$\begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix} = 0$$

La droite épipolaire s'écrit :

$$\begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} A_2 \\ B_2 \\ C_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} E \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = 0$$

$$m_2^T E m_1 = 0$$

E est la matrice essentielle.

### 2.1.2.2 Matrice fondamentale

La matrice est utilisée pour décrire les paramètres qui interviennent à la fois dans les coordonnées des points, de la caméra et de l'image, en passant par les paramètres des coordonnées de la matrice, qui sont convertis en fonction de la référence de la caméra, lesquels sont exprimés en fonction de la référence de l'image, en fonction des paramètres intrinsèques liés aux caractéristiques de chaque caméra.

la calibration nous avons pu connaître l'angle que nous obtenons à l'intersection de deux rayons de lumière émis par le point opposé, ce qui correspond essentiellement à deux points identiques, ce qui nous permet de connaître la position de ce point en fonction de la référence de l'objet sur lequel nous travaillons (repère de référence).

La relation entre les coordonnées image (p) et caméra (m) est :

$$p_1 = C_1 m_1 \text{ Pour la caméra gauche et}$$

$$p_2 = C_2 m_2 \text{ Pour la caméra droite avec}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f\alpha_u & 0 & \alpha_u p_x \\ 0 & f\alpha_v & \alpha_v p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

(À facteur multiplicatif près)

$$\text{or } m_2^T E m_1 = 0$$

$$p_2^T (C_2^{-1})^T E C_1^{-1} p_1 = 0$$

La matrice F telle que  $F = (C_2^{-1})^T E C_1^{-1}$  est la matrice fondamentale.

$$(A B)^T = B^T A^T$$

La calibration est classée en deux catégories globales.

#### 2.1.2.2.1 Calibrage classique (hors ligne)

Cette calibration est effectuée en analysant les images d'une cible avec une géométrie bien connue. Il est principalement utile de connaître la forme du motif à photographier. Ces formes sont ensuite découvertes dans les images afin de calculer les paramètres intrinsèques.

#### 2.1.2.2.2 Auto-calibration (en ligne)

Calibration automatique de l'appareil photo consiste à déterminer les paramètres internes de l'appareil directement à partir de plusieurs images non étalonnées de scènes non structurées. Contrairement à l'étalonnage classique de la caméra, calibration automatique ne nécessite aucun objet calibration spécial dans la scène. Dans le secteur des effets visuels, calibration automatique de la caméra fait souvent partie du processus de "correspondance en mouvement", dans lequel une trajectoire de caméra synthétique et un modèle de projection intrinsèque sont résolus pour reprojeter le contenu synthétique en vidéo. Dans ce cas on n'a pas besoin de photographier la mire, les paramètres intrinsèques sont calculés une fois pour toute, par contre les paramètres extrinsèques doivent être recalculés à chaque mouvement à l'aide de la géométrie épipolaire [32].

## 2.2 La mise en correspondance

La mise en correspondance stéréoscopique était un [33] problème complexe dans la recherche scientifique et la communauté scientifique de plus de 40 ans. Marr et Poggio, en 1979, a été en mesure de déterminer l'ambiguïté naturelle du problème de la mise en correspondance. Et a découvert deux solutions à cette ambiguïté et qui sont l'unicité et la continuité.

L'une des premières étapes de la reconstruction est la mise en correspondance. En fait, quel que soit le type de reconstruction envisagé, il est nécessaire de trouver dans les deux images les primitives homologues, c'est ce qui correspond à la même entité physique du monde réel [34].

En utilisant deux images et en prenant les positions des deux points dans lesquels les images, on peut reconstruire un point dans l'espace grâce à la triangulation, mais c'est là n'est pas où réside le principal problème de la mise en correspondance. En effet, l'être humain n'éprouve aucune difficulté à s'afficher des points, en revanche ce qui se passe du point de vue algorithmique et très compliqué.

### 2.2.1 Primitives stéréoscopiques

Avant de procéder à la mise en correspondance, on doit d'abord analyser la scène à reconstruire. L'image doit être traitée conformément aux méthodes utilisées, mais le travail est effectué sur les mêmes éléments, on s'intéresse alors aux types de primitives suivantes :

- Point : (pixels, points d'intérêt, éléments de contour, points caractéristiques le long d'un contour, jonctions etc...).
- Segment : (segments de droite, arcs de cercle, portions de conique etc...).
- Région : Les régions sont mal adaptées dans le processus de la stéréovision sauf quelques cas particuliers.

### 2.2.2 Notion de la disparité

La disparité est utilisée dans les domaines de la vision par ordinateur, elle détermine la différence de position de projection du même objet sur la rétine. La disparité dans le domaine de la mise en correspondance est définie par :

La disparité est la différence de position de l'image dans laquelle les points de la scène sont supprimés, ce qui fonctionne avec les systèmes de caméra gauche et droite.

$$d: \mathbb{N} \rightarrow \mathbb{R}^+$$

$$I(u, v), I(u, v') \mapsto \text{disp}(I(u, v)) = |v - v'|.$$

#### 2.2.2.1 Fonction de disparité

On peut analyser un problème de mise en correspondance par la recherche d'une fonction de disparité  $d$  qui représente chaque pixel en eux  $p_g^{i,j}$  [10]. Dans le cas général, on a :

$$d: \mathbb{N}^2 \rightarrow \mathbb{R}^2$$

$$p_g^{i,j} \rightarrow d(p_g^{i,j}) = (u - i \ v - j)^T$$

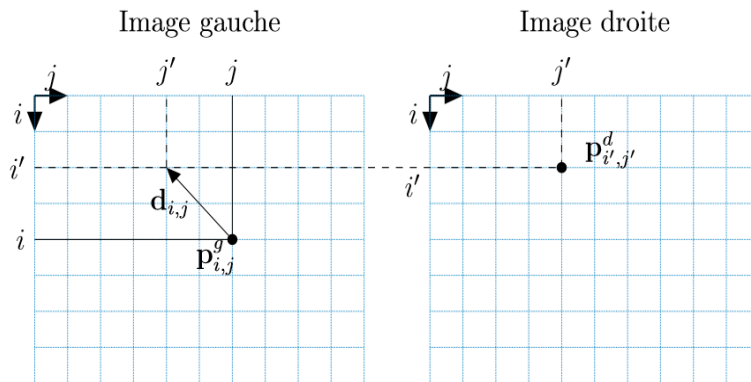
La différence réelle doit être calculée, car elle est responsable de l'appariement du sous-pixel. Dans la plupart des méthodes qui s'appuient sur une mise en correspondance de pixels, l'analyse est faite pixel par pixels, ce qui signifie qu'il faut compter les disparités entières.



**2.2.2.2 Vecteur de disparité**

Le résultat, est stocké dans un vecteur de disparité noté par  $d_{i,j}^{g \rightarrow d}$ , ce vecteur décrit la position de s deux pixels correspondants entre [35] deux images gauche  $g$  et droite  $d$ . On a  $i,j$  donc :  $d^{g \rightarrow d} = d_{i,j}$ ,

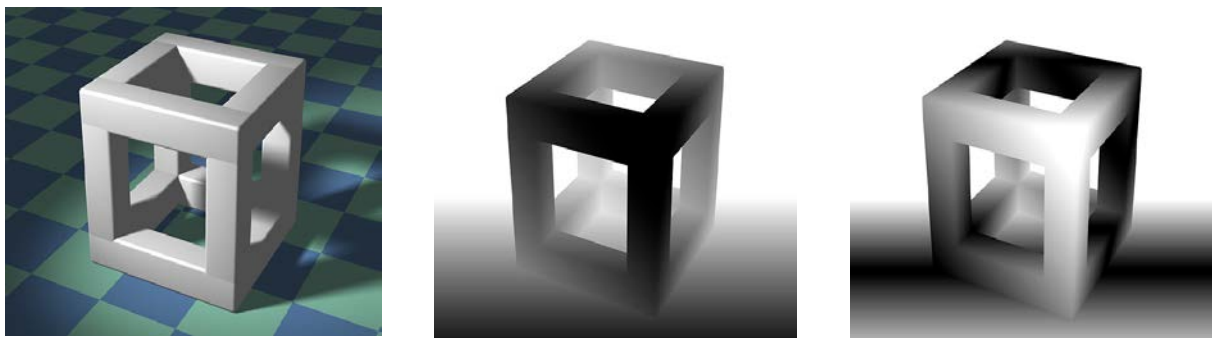
$$z d_{i,j} \in ([d_{i_{min}}; d_{i_{max}}] \times [d_{j_{min}}; d_{j_{max}}]) \cup \{OCC\}. \text{ Si } d_{i,j} \neq OCC \text{ alors : } \begin{bmatrix} i' \\ j' \end{bmatrix} = \begin{bmatrix} i \\ j \end{bmatrix} + d_{i,j}$$



**Figure [II.5] : Vecteur de disparité.**

**2.2.2.2 Carte de disparité**

En infographie 3D, une carte de disparité est une image ou un canal d'image contenant des informations relatives à la distance des surfaces des objets de la scène par rapport à un point de vue. Le terme peut être analogue à depth buffer, Z-buffer, Z-buffering and Z-depth. Le "Z" dans ces derniers termes concerne une convention selon laquelle l'axe de vision central d'une caméra est dans la direction de l'axe Z de la caméra, et non pas l'axe Z absolu d'une scène.



**Figure [II.6] : a) Structure cubique b) Carte de disparité : Plus proche est plus sombre c) Carte de disparité : Plus le plan focal est sombre.**

Vous pouvez voir ici deux Cartes de disparités différentes, ainsi que le modèle d'origine à partir duquel elles ont été dérivées. La première Carte de disparité montre la luminance proportionnellement à la distance de la caméra. Les surfaces les plus proches sont plus sombres ; les autres surfaces sont plus légères. La deuxième Carte de disparité montre la luminance en relation avec les distances par rapport à un plan focal nominal. Les surfaces plus proches du plan focal sont plus sombres ; les surfaces plus éloignées du plan focal sont plus claires (à la fois plus proches et plus éloignées du point de vue).

### 2.2.3 Problème lié à la mise en correspondance

Les principales difficultés liées au processus de la mise en correspondance sont dues essentiellement à deux raisons [36] :

- **Information manquante** : Il arrive souvent que les informations prises pour le travail de la mise sur la correspondance ne figurent pas sur la photo, et on les trouve dans :
  - **Les zones occultées** : Dans cette zone, on peut voir une partie de l'image qui n'apparaît pas dans l'autre. Les zones occultées doivent avoir une texture transparente.
  - **Les zones de raccourcissement** : Dans cette zone, différents points de la scène sont appliqués dans un ensemble de pixels de l'image, mais sur les mêmes pixels que l'autre image.
  - **Les zones de discontinuité de profondeur** : Ce sont les zones situées à la limite des différentes profondeurs. Cela peut se traduire par des voisinages de deux correspondants qui ne se ressemblent pas nécessairement.
- **Informations ambiguës** : des 'information dans l'image, mais on ne peut pas distinguer entre les correspondants.
  - **Régions homogènes** : Dans cette région, il est difficile de distinguer les points correspondants car tous leurs pixels sont homogènes et ressemblent à du bruit.
  - **Régions avec textures répétitives** : Il existe une ambiguïté lorsqu'un même motif se répète plusieurs fois.

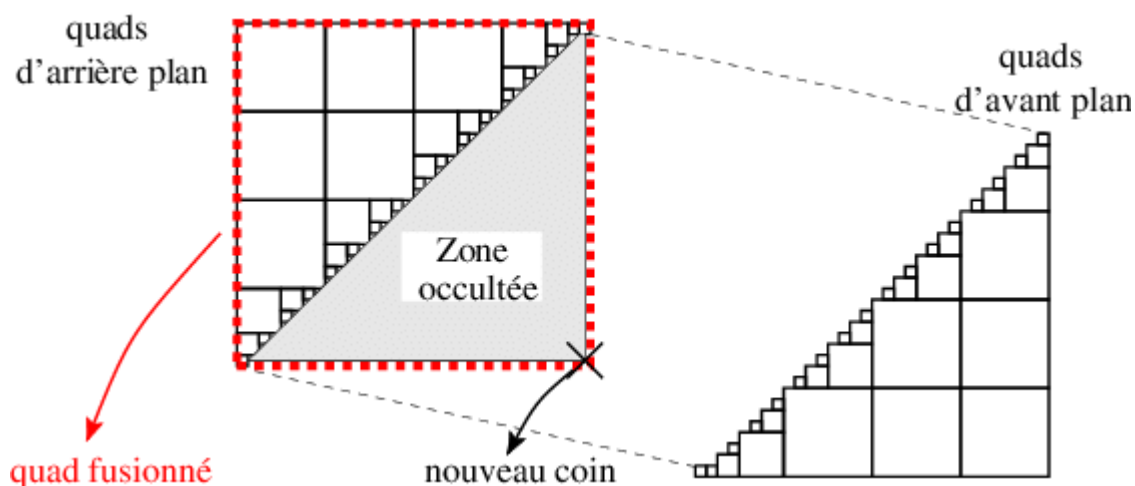


Figure [II.7] : Les zones occultées doivent avoir une texture transparente. La profondeur du nouveau coin doit être calculée.

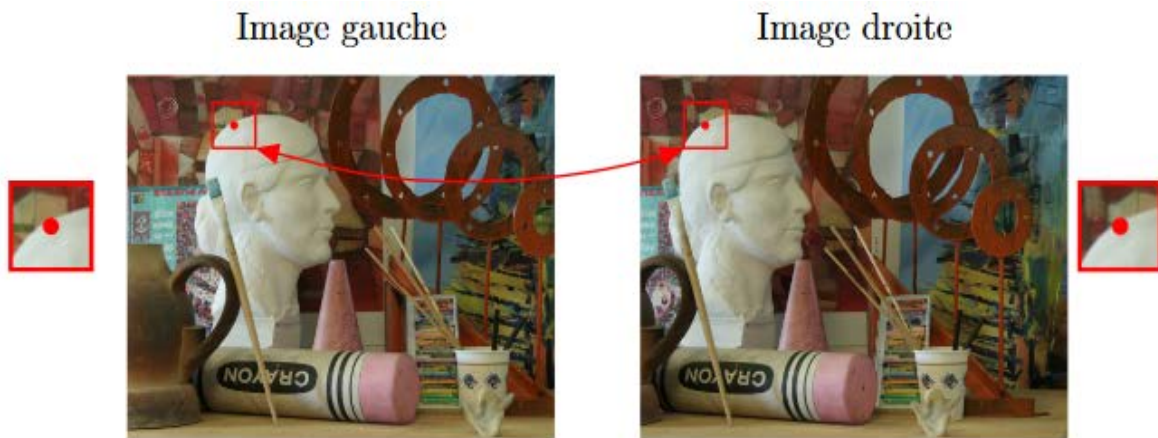


Figure [II.8] : Zone de discontinuité de profondeur – Les deux pixels d'un point rouge se correspondent mais leurs voisinages diffèrent car ils sont proches d'une discontinuité de profondeur-.

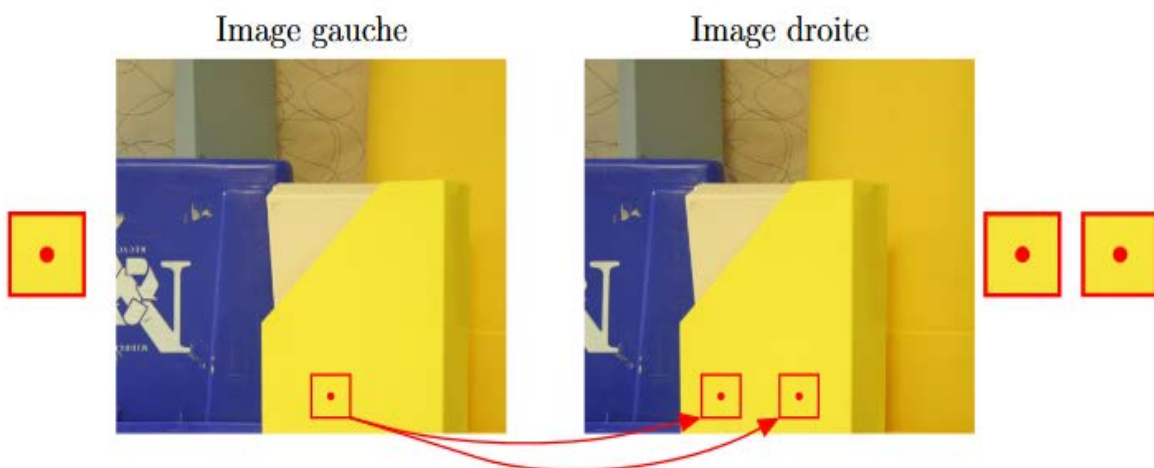


Figure [II.9] : Zone homogène –Manque d'information qui caractérise le pixel-.

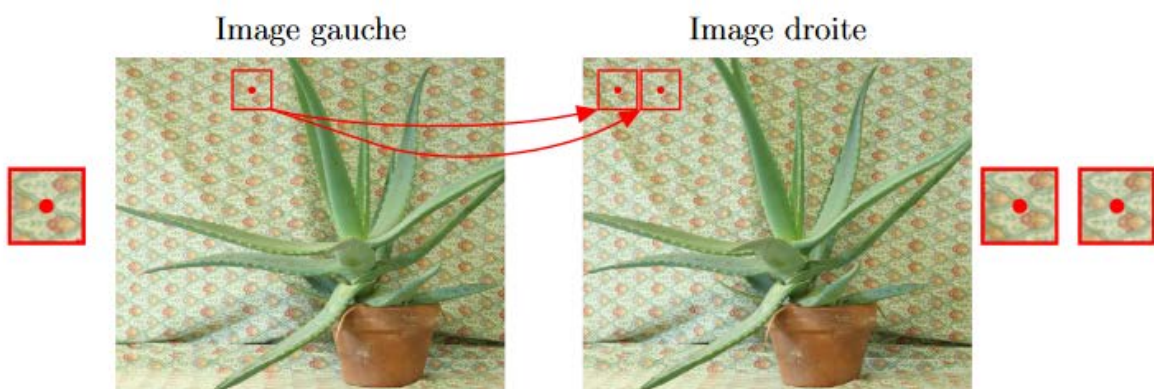


Figure [II.10] : Zone avec textures répétitives –Pour un pixel il existe plusieurs candidats dans l'autre image. (D. Scharstein and R. Szeliski - Middlebury Stéréo Vision dataset)

### 2.2.4 Les contraintes de l'appariement

Il existe d'autres règles régissant la mise en correspondance, un autre aspect est celui des contraintes de la mise en correspondance, qui est divisé en deux parties différentes de ces contraintes : Contraintes géométriques et contraintes figurales. Ces limitations permettent de réduire le nombre de correspondants, ou de supprimer certaines correspondances existantes.

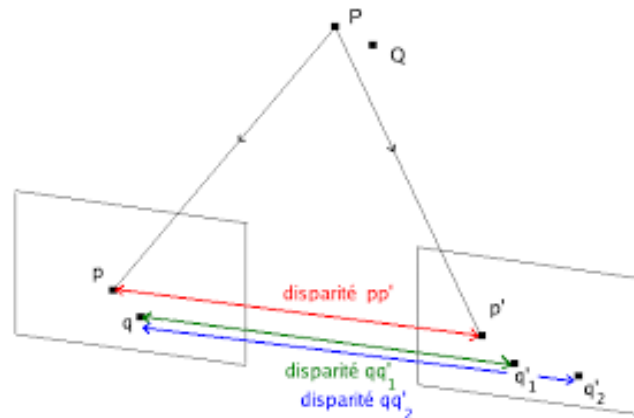


Figure [II.11] : l'appariement stéréoscopie.

#### 2.2.4.1 Contraintes géométriques

##### Contrainte épipolaire

Comme nous l'avons mentionné dans la première section sur la géométrie épipolaire et l'utilisation de ses propriétés, la zone de recherche peut être réduite à une ligne épipolaire.

Le fait que le point correspondant doit se trouver sur une droite est souvent appelé la « contrainte épipolaire » (cette expression deviendra plus claire dans la suite). Cette contrainte est très utile pour la mise en correspondance d'images : la recherche de correspondances [37] peut être accélérée considérablement (il ne faut plus chercher dans toute l'image pour trouver le point correspondant) et le risque d'erreur est réduit.

$P$  est un point de la scène, le point qui passe par  $P$  et les centres optiques  $O^g$  et  $O^d$  des deux caméras est appelé plan épipolaire. Ce plan coupe les images gauche et droite suivant deux droites appelées droites épipolaires. Pour renforcer cette contrainte, il faut que les deux images stéréo soient rectifiées. La rectification consiste à recalculer la position générale de deux images pour que la géométrie épipolaire de ces dernières soit simple, c'est-à-dire que les droites épipolaires soient horizontales

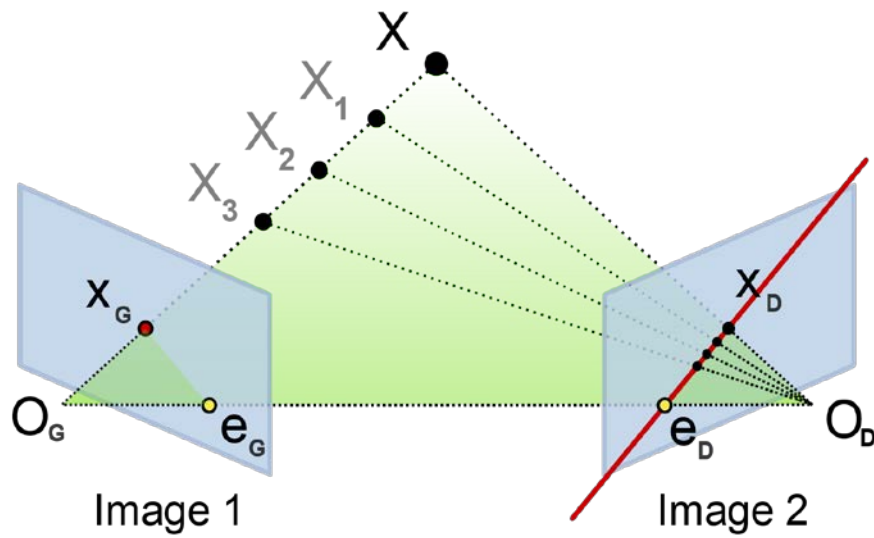


Figure [II.12] : Géométrie épipolaire.

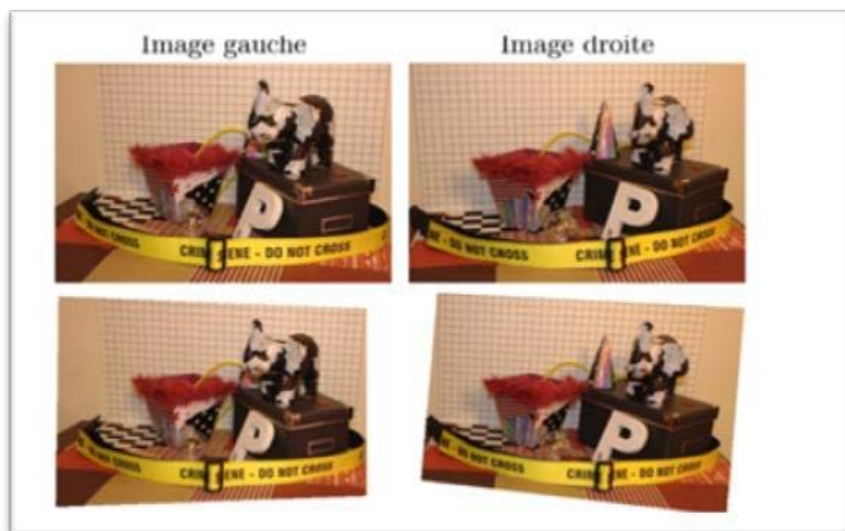


Figure [II.13] : Exemple de la rectification épipolaire. (G. Galles-Pixel matching for binocular stereovision by propagation of feature points matches and region-based randomized voting scheme)

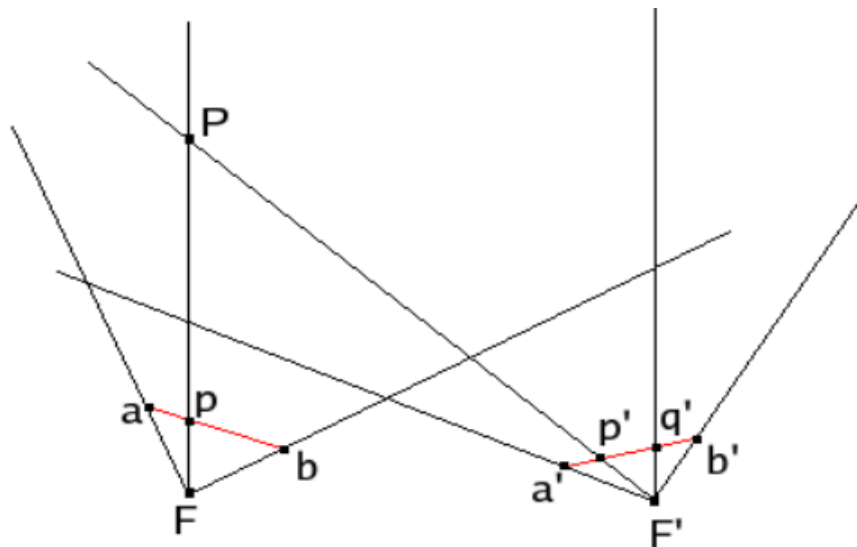


Figure [II.14] : Le point P visualisé e par deux cameras F et F'.

Le point P d'une scène se projette sur le segment a' q' pour l'image de centre de projection F'. Cette variation est due à la distance qui sépare P du capteur stéréoscopique.

### Contrainte d'unicité

Cette contrainte dépend du fait que les pixels de l'image de gauche peuvent avoir un seul pixel correspondant à l'image de droite. S'il y a un pixel de l'image de droite et plus d'un pixel symétrique dans l'image de gauche, cela signifie des régions occultées, c'est le cas du pixel de l'image de gauche qui n'a pas d'équivalent dans l'image de droite.

À chaque pixel de l'image de référence est associé au maximum un pixel de l'autre image, et inversement.

Cette contrainte peut être formulée par :

$$\forall I(u, j) \in P, \exists ! I(u, v) \in P / (disp(Ig(u, v)) = |v - j|) \text{ Si}$$

Cette contrainte n'est pas respectée, la formulation devient :

$$\forall I(u, j) \in P, \exists ! I(u, v), I(u, v') \in P / (disp(Ig(u, v)) = |v - j|; \\ (disp(Ig(u, v')) = |v' - j| \text{ et } v \neq v')$$

Deux points peuvent être projetés en un même point dans une image et en deux points dans l'autre image.

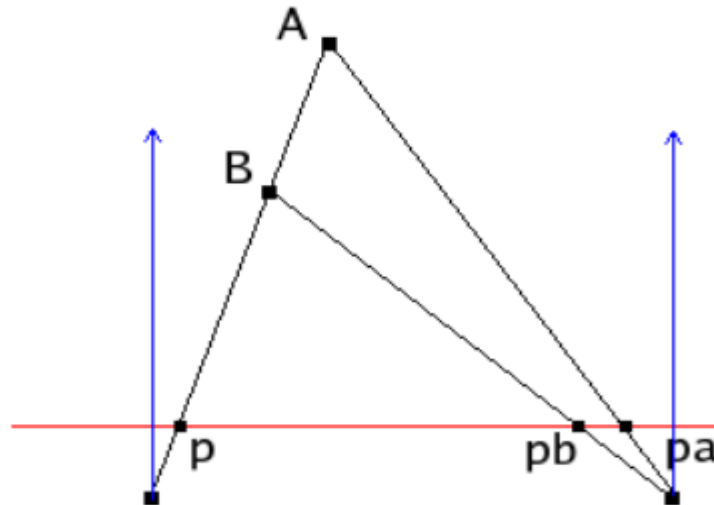


Figure [II.15] : deux caméras visualisent deux points.

### Contrainte de continuité de disparité

Cette contrainte est issue de la calibration de la caméra, elle implique que les primitives voisines spatialement ont des valeurs de disparité proches.

### Contrainte d'ordre

Les contraintes épipolaires et de disparité permettent de réduire le nombre d'appariements possibles entre les primitives de l'image gauche et les primitives de l'image droite.

Lorsque deux points sont à la surface d'un objet, l'ordre de leur projection est le même dans les deux images. Si l'objet est fortement incliné et qu'il est estompé, les projections n'apparaîtront pas en même temps dans les deux images. Par contre, si l'objet est transparent, l'ordre des projections ne sera pas respecté.

Cette contrainte peut être formulée par :

$$\forall I(u, v) \ I(u, v') \in P \text{ et } v > v', \exists I(u, j) \ I(u, j') \in P / (\text{disp}(I_g(u, v)) = |v - j|, (\text{disp}(I_g(u, v')) = |v' - j'|, \text{ et } j > j').$$

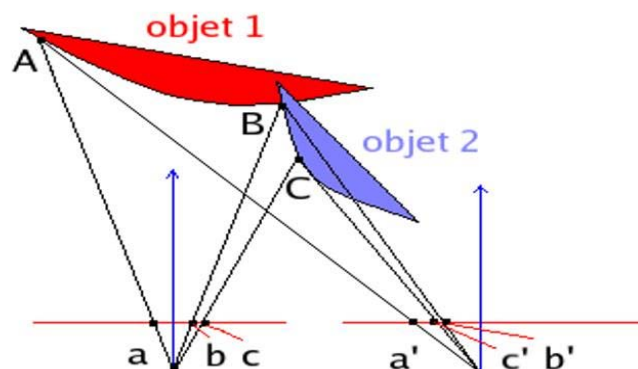


Figure [II.16] : Contrainte d'ordre.

### 2.2.4.2 Contraintes figurales

Les contraintes figurales sont différentes des contraintes géométriques en ce qu'elles imposent des restrictions sur les contraintes géométriques et trouvent où toutes les contraintes spécifiques à une primitive particulière. Et elles peuvent être aussi imposer une contrainte d'ordre géométrique.

#### Disparité locale constante

Lorsque la scène concerne des objets dont la surface est fortement inclinée par rapport au niveau des deux images, il existe un risque de manque de respect pour le système. Ce problème peut être évité si on impose une limite au gradient de disparité entre deux paires de points appariés consécutifs. Cependant, en fait, nous limitons la pente des surfaces des objets dans la scène et, en fait, nous limitons la diversité des objets à reconstruire.

#### Continuité figurale

Cette contrainte permet de limiter les différences de disparités le long de la frontière. Il est conseillé d'éviter de faire correspondre les points de contour de l'image de gauche avec plusieurs points de paramètres dans l'image de droite. Que le contour puisse appartenir à plus d'une surface, mais ne doit pas dépasser le bord de l'objet.

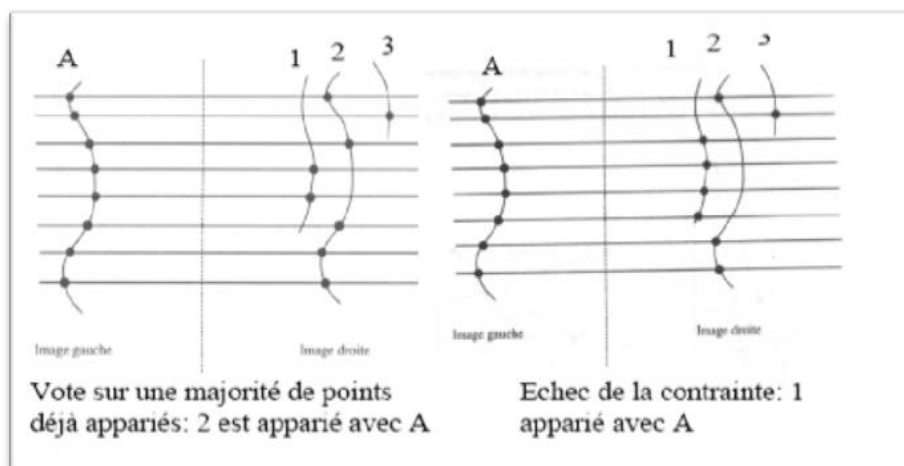
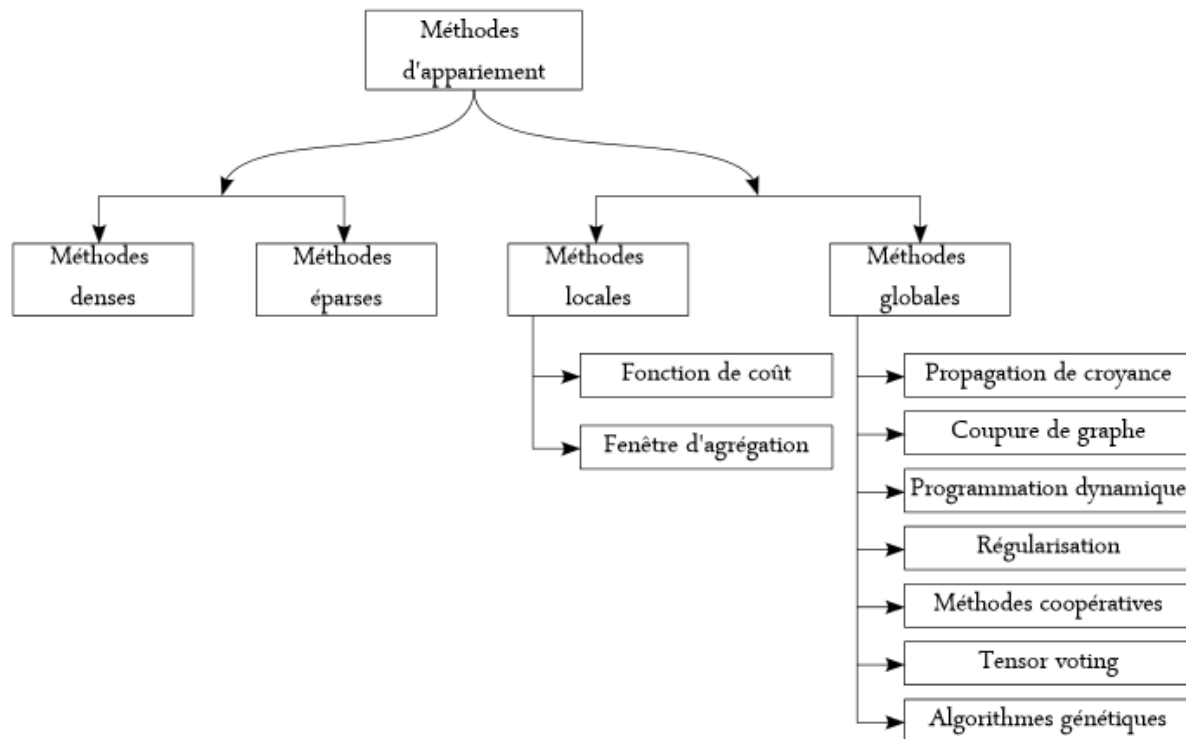


Figure [II.17] : Contrainte de continuité figurale.

### 2.2.5 Techniques de mise en correspondance

Il y a de nombreuses façons d'accomplir la mise en correspondance [38], et il en est de même pour les problèmes qui leur sont liés. La classification des méthodes d'appariement dépend de la complexité algorithmique, de la nature de la carte de disparité obtenue telles que les méthodes denses et éparées. Parmi les algorithmes récents d'appariement proposés dans la littérature, on trouve ceux qui sont basés sur la coupure de graphe, ils transforment le problème de la mise en correspondance stéréoscopique en un problème de minimisation d'une fonction d'énergie globale. Une autre classification a été largement étudiée et est classée par des méthodes locales et globales.





**Figure [II.18] : Classification possible des méthodes d'appariement.**

### 2.2.5.1 Méthodes d'appariement denses

Ce type de méthodes prend en compte tous les pixels composant l'image. La mise en correspondance tire parti des informations contenues dans l'environnement des pixels. La probabilité de mise en correspondance est mesurée en deux pixels avec la corrélation entre les voisins de chacun de ces deux pixels. Cette possibilité peut être estimée par des déviations des valeurs photométriques ou colorimétriques des pixels correspondants. Cette technique suppose la continuité des surfaces dans les scènes réelles, en les traduisant en continuité de disparité.

### 2.2.5.2 Méthodes d'appariement éparses

Le changement d'échelle dans les images résultant de transformations géométriques assure la stabilité des points d'intérêt. Le principal inconvénient de ce type de technique réside dans l'impossibilité de reconstruire une scène tridimensionnelle complète. L'appariement est soutenu par des points d'intérêt non ambigus, dans de nombreux algorithmes de mise en correspondance. Ces points correspondent à la primitive pouvant comporter plusieurs attributs. Plus la distinction entre les primitives est grande, plus le couplage est précis, puissant et fiable.

### 2.2.5.3 Méthodes d'appariement locales

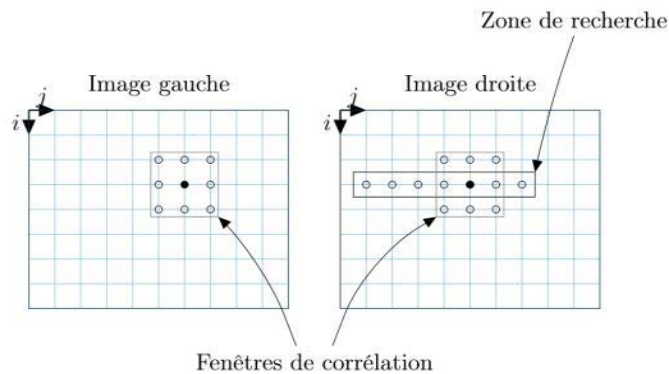
Les méthodes locales mesurent la similarité de voisinage entre deux pixels, Ce voisinage est appelé zone d'agrégation ou bien fenêtre de corrélation, Pour simplifier, on suppose que cette fenêtre de carré de corrélation est obtenue comme suit :

$$(n * n) = (2r + 1) * (2r + 1), \quad u, r \in \mathbb{N}^*$$

Cette fenêtre dépend des pixels sur lesquels nous travaillons et étudions. Il est également à noter qu'un vecteur est la dimension  $n^2$  qui contient des valeurs de pixels adjacentes aux pixels visualisés sur l'image de référence, et  $b$  pour l'autre image. Les méthodes locales supposent que les voisinages de deux correspondants se ressemblent, c'est-à-dire :

$$a \approx b$$

Dans la figure II.19 le pixel considéré pour la recherche de correspondant est marqué en noir, son pixel homologue est recherché dans la zone de corrélation. Dans cet exemple, la zone de recherche considérée est monodimensionnelle. Le but est de calculer pour chaque pixel candidat un score de corrélation est calculé. A la fin le candidat qui donne le meilleur score est retenu comme correspondant.



**Figure [II.19] : Mise en correspondance locale.**

Parmi les méthodes locales les plus connues on trouve les suivantes :

**NCC** : La méthode de corrélation croisée normalisée est une méthode couramment utilisée, elle est employée pour la mesure de similarité [39]. C'est la somme des produits entre les termes correspondants des fenêtres de corrélation normalisée par le produit des moyennes quadratiques calculées pour chacune des fenêtres. L'avantage de cette méthode est la capacité de corriger les variations de l'intensité et de la couleur qui change d'une caméra à une autre. Par contre le problème de la corrélation normalisée est la lenteur des calculs.

Pour un pixel (u, v) est une disparité d on a :

$$NCC(u, v, d) = \frac{(\sum_{i,j} I_1(u+i, v+j) - I_1) \cdot (\sum_{i,j} I_2(u+d+i, v+j) - I_2)}{\sqrt{(\sum_{i,j} I_1(u+i, v+j) - I_1)^2 + (\sum_{i,j} I_2(u+d+i, v+j) - I_2)^2}}$$

Tels que :  $I_1$  et  $I_2$  sont les deux images de la mise en correspondance,  $I_1$  et  $I_2$  correspondent aux moyennes de l'intensité des pixels sur la fenêtre de corrélation.

**SSD** : C'est la somme des différences carrées, utilisée pour la mise en correspondance en bloc [40].

$$SSD(u, v, d) = \sum_{i,j} (I_1(u+i, v+j) - I_2(u+d+i, v+j))^2$$

**SAD** : C'est la somme des valeurs absolues entre termes correspondants des deux fenêtres [14].

$$\text{SAD}(u, v, d) = \sum_{i,j} |I(u+i, v+j) - I(u+d, v+j)|$$

**ZSAD:** C'est une méthode analogue à la méthode SAD, mais on fait la différence entre écarts à la moyenne, elle représente un critère invariant aux variations uniformes de luminance dans une des images.

### 2.3.1 La triangulation

La triangulation permet de calculer la position 3D (X, Y, Z) d'un point P à partir d'une projection en perspective de P sur les plans d'image des caméras une fois que la position relative et l'orientation des deux caméras sont connues.

Pour obtenir une précision des mesures exactes, la position entre les capteurs doit être importante et les capteurs doivent être parallèles et orientés vers l'objet à analyser. Soit  $P$  un vecteur représentatif d'un point source dans le repère (O, X, Y, Z). L'orientation du repère caméra est définie par la direction de la normale  $N$  au plan du capteur. Et soient  $(\vec{v})$  les vecteurs des directions principales du capteur.

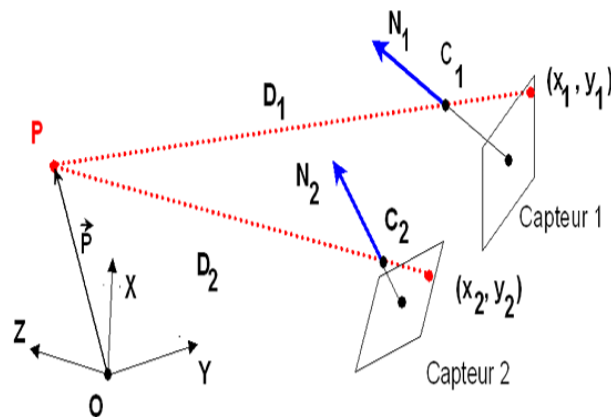


Figure [II.20] : Système de projection des scène –Représentation de triangle formé par le système.

La projection de P dans le capteur1 est donnée par :

$$x_1 = f \frac{\vec{D}_1 \cdot \vec{1}}{\vec{D}_1 \cdot \vec{N}_1} \quad y_1 = f \frac{\vec{D}_1 \cdot \vec{2}}{\vec{D}_1 \cdot \vec{N}_1} \quad \text{Avec : } \vec{D}_1 = \vec{P} - \vec{C}_1$$

Et la projection de P sur le capteur 2 est aussi donnée par une relation identique :

$$x_2 = f \frac{z_2}{z_2 - C_{D_2 \cdot N_2}} \quad y_2 = f \frac{D_2 \cdot z_2}{z_2 - C_{D_2 \cdot N_2}} \quad \text{avec } : \vec{D}_2 \rightarrow \vec{N}_2$$

Afin de déterminer les trois composantes (x, y, z) du point P, il faut résoudre cet ensemble d'équations au sens des moindres-carrés.

## Conclusion

Dans ce chapitre nous avons traité la stéréovision considérée comme l'une des méthodes de reconstruction 3D les plus connues. Nous avons détaillé la démarche générale, son utilité et les problèmes rencontrés au cours de son automatisation.

La stéréovision permet, l'extraction d'une carte de disparités qui représente le décalage des points de la première par rapport à la seconde. Cette carte est utilisée par la suite pour induire la forme tridimensionnelle de l'objet en question.

Notre but, est bien de tenter de reconstruire l'objet sans passer par le processus de triangulation. Et pour justifier notre idée, nous consacrons le chapitre suivant pour comprendre le principe de fonctionnement et identifier ses différentes lacunes.

**CHAPITRE03 :**  
**RECONSTRUCTION 3D**  
**SANS**  
**TRIANGULATION**

## Introduction

La reconstruction 3D est un sujet de recherche très actif ces dernières années, les algorithmes et les méthodes utilisés dans ce domaine continuent de progresser afin de produire des résultats plus proches de la réalité dans les meilleurs délais. Dans ce qui précède, nous avons présenté les différentes méthodes de reconstruction 3D, ainsi que leur classification multistandard.

La stéréovision est une technique qui permet de restructurer une scène tridimensionnelle à partir de deux images différentes prises sous deux perspectives différentes. Le but de cette méthode est de restaurer la troisième dimension manquante lors de la configuration de l'image en calculant l'ensemble de points correspondants. Ce processus est très sensible, en particulier si la scène contient des parties ou des zones recouvertes d'une texture homogène. De plus, le calcul des paires opposées est très complexe et nécessite un temps de traitement élevé.

En géométrie et trigonométrie, la triangulation est une technique permettant de déterminer la position d'un point en mesurant les angles entre ce point et d'autres points de référence dont la position est connue, et ceci plutôt que de mesurer directement la distance entre les points. Ce point peut être considéré comme étant le troisième sommet d'un triangle dont on connaît deux angles et la longueur d'un côté.

La combinaison de ces deux méthodes est un sujet déjà abordé, et le but est de surmonter les limites de chaque approche, mais dans ce mémoire, nous procédons différemment et nous essayons de ne pas utiliser la triangulation pour reconstruire l'objet étudié.

Dans ce qui suit nous allons traiter notre approche proposée sous le titre : « la reconstruction 3D sans triangulation ».

Notre objectif est de d'utiliser la stéréovision seule et sans faire intervenir la triangulation. Nous proposons ainsi une technique qui permet de reconstruire partiellement l'objet à partir de la carte de disparité seulement et ce en calculant les différentes profondeurs par l'application d'un facteur particulier. Par la suite, nous présenterons quelques résultats obtenus par l'application de notre idée.

## 1. Système d'acquisition de l'image

La méthode de reconstruction 3D proposée dans le cadre de notre travail est appliquée sur deux images acquises par la même caméra avec angles un petit angle de rotation ; de manière qu'on puisse appliquer la technique de stéréoscopie.

Le système d'imagerie peut être constitué d'un appareil photo qui prend deux images sous deux angles différents, mais proches l'une de l'autre. L'idée de base est de prendre deux images d'un objet et de les traiter pour créer un modèle tridimensionnel.

Dans ce cas, nous avons utilisé deux objets différents, avec 2 images de chaque objet

## 2. Matériels et langage de programmation

Les images que nous avons utilisées sont capturées par un smartphone de type LG G5 avec une caméra (de 16 mégapixels d'une longueur focale de 28 mm, la taille de l'ouverture est 1.9f, la taille du capteur est 1 / 2.6).

L'ordinateur que nous utilisons est un LENOVO 80QQ qui est doté d'un processeur Intel core i3 5005-U 2.00 GHz, une RAM de 4 GO et une carte graphique Intel HD graphique 5500. Le système d'exploitation installé sur cette machine est WINDOWS 8 Pro 64 bits.

Nous avons choisi l'environnement de programmation MATLAB pour l'implémentation des différents algorithmes, la version utilisée est celle de 2018.

### 2.1. Programmation avec MATLAB

Le nom MATLAB vient du mot anglais MATrix LABoratory. Ce logiciel prend son nom où la plupart de ces fonctions prédéfinies sont des grandeurs matricielles, et par extension pour les données vectorielles. Les domaines d'applications qui peuvent être traités par MATLAB sont très variés, on trouve par exemple : Le traitement du signal, le traitement d'images, l'automatisme...

Ce langage peut être exécuté sur différentes plateformes : LINUX, WINDOWS, ou bien MAC OS. Les programmes sont écrits sous forme de scripts. Lorsque MATLAB démarre, l'IDE (Integrated Développement Environnement) apparaît à l'écran, c'est l'interface graphique du MATLAB (Voir figure V.2). Cette interface est assez claire pour un utilisateur débutant.

Nous avons choisi ce langage de programmation parce qu'il est très utile pour les calculs numériques sur les vecteurs car nous sommes en train de traiter des images qui sont représentées sur machine sous forme de matrices, ce langage est sacrément évolué et contient beaucoup de fonctionnalités dédiées au domaine de traitement d'images et vision par ordinateur. De plus MATLAB aide beaucoup pour tracer les courbes de données.

## 3. Démarche proposée du processus de reconstruction 3D

Comme mentionné ci-dessus, notre approche est basée sur la méthode de stéréovision (en évitant l'utilisation de la triangulation) pour reconstruire partiellement un objet. La stéréoscopie produit une carte de disparité, à partir de laquelle nous tirons la troisième dimension des images, en utilisant un simple facteur de mise à l'échelle. L'objet est donc reconstruit par l'association des coordonnées x et y d'un point issue directement de l'une des deux images avec la valeur correspondante à une profondeur x qui est la valeur  $D[i, j]$  après application de la transformation en utilisant le facteur cité ci-dessus (où D est la matrice des disparités) ; de cette façon, on obtient une forme partielle de l'objet à reconstruire.

Le processus de la reconstruction 3D se déroule en six phases essentielles :

- 1) L'acquisition des images.
- 2) L'isolation de l'objet de son arrière-plan

- 3) Application stéréoscopie et obtention de la carte de disparité.
- 4) Application du facteur de transformation choisis.
- 5) Association des coordonnées x, y avec z qui le résultat de la phase précédente.
- 6) Positionnement des points 3D dans l'espace.

### 3.1. Acquisition des images

Il est connu que chaque image est constituée d'un ensemble de propriétés telles que le type d'image, les dimensions de l'image et le nombre de pixels qu'elle contient, toutes en relation directe à l'appareil photo utilisé pour prendre les photos.

La technique utilisée pour capturer des images consiste à prendre deux photos d'un objet avec un petit décalage horizontal entre les différents points de vue.

Les images que nous avons utilisé sont de taille 450\*450.



Figure [III.1] : Une image utilisée (gauche ; Objet 1).





Figure [III.2] : Une image utilisée (gauche ; Objet 2 ).

### 3.2. Isolation d'objet de son arrière-plan

Les images d'entrée, spécifiée en tant qu'image tricolore  $M$ -par- $N$ -par-3 ou image en niveaux de gris 2-D  $M$ -par- $N$ .  $I_1$  et  $I_2$  doivent être réels, finis et non séparés, et les images doivent avoir la même taille.

Nous avons alors utilisé un seuil pour isoler l'objet de son arrière-plan. Si la couleur d'un pixel est supérieure (inférieure) au seuil choisit ; alors appartient au fond et est recoloré en blanc, et si ce n'est pas le cas, alors le pixel appartient à l'objet et préserve sa même couleur.

On a appliqué ce traitement ; pour ne prendre en considération que les informations pertinentes. Et Les résultats obtenus après application de l'algorithme d'isolation sont les suivants :



Figure [III.3] : résultat d'Isolation d'objet 1

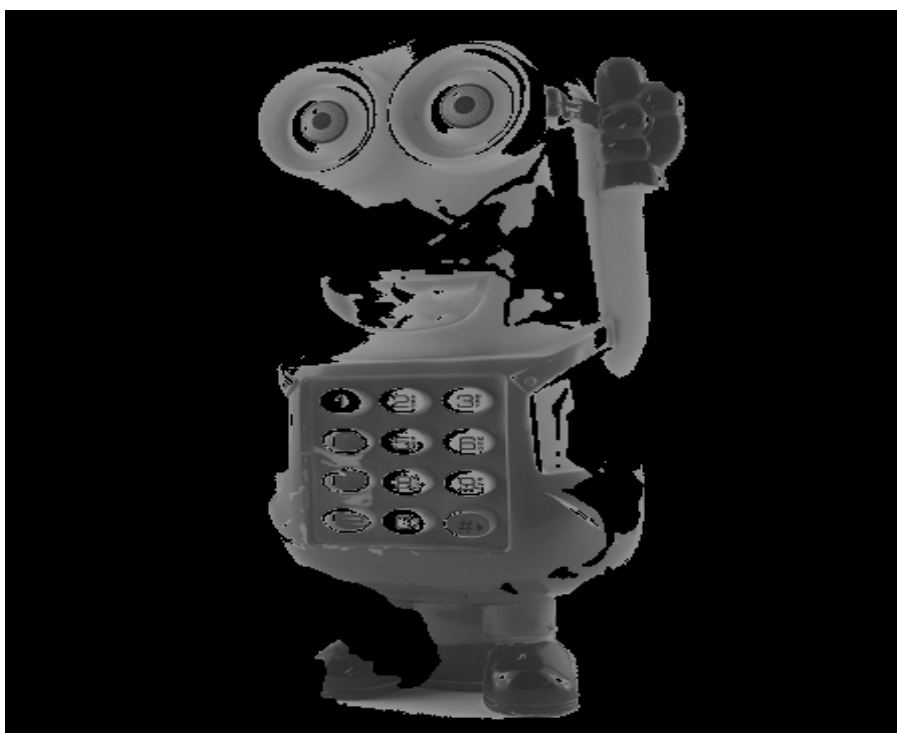


Figure [III.4] : résultat d'Isolation d'objet

2

### 3.3. Application de la stéréoscopie

La méthode stéréoscopique que nous avons utilisé calcul les correspondances en utilisant ma l'approche des moindres carrés.

Le code de stéréoscopie :

```
%% Basic function - Creates stereoscopic image and waits for key press.
global resultImage M Position ignore
ignore = 1;
% Saving function inputs to workspace.
assignin('base','image1',image1);
assignin('base','image2',image2);
assignin('base','fil1',filter1);
assignin('base','fil2',filter2);
% Reading the images.
img1 = imread(image1);
img2 = imread(image2);
% Adjusting the colors in both images, based on input filters.
switch filter1
    case 'red'
        img1(:,:, 2:3) = 0;
    case 'green'
        img1(:,:, 1:2:3) = 0;
    case 'blue'
        img1(:,:, 1:2) = 0;
    case 'cyan'
        img1(:,:, 1) = 0;
    case 'magenta'
        img1(:,:, 2) = 0;
end
switch filter2
    case 'red'
        img2(:,:, 2:3) = 0;
    case 'green'
        img2(:,:, 1:2:3) = 0;
    case 'blue'
        img2(:,:, 1:2) = 0;
    case 'cyan'
        img2(:,:, 1) = 0;
    case 'magenta'
        img2(:,:, 2) = 0;
end
```

```

% Creating 3 important variables in workspace.
assignin('base','rotation',0);
assignin('base','verticalShift',0);
assignin('base','horizontalShift',0);
% Our first image is just a sum of input images with adjusted colors.
resultImage = img1+img2;
% Getting screen and image information.
set(0,'Units','pixels');
screenSize = get(0,'ScreenSize');
imgSize = size(resultImage);
% Determining figure size, based on screen and image aspect ratios.
figureAR(1) = imgSize(2) / screenSize(3);
figureAR(2) = imgSize(1) / screenSize(4);
AR = figureAR(1) / figureAR(2);
Position = [0 0 0 0];
if (AR < 1)
    Position(4) = 0.5;
    Position(3) = Position(4) * AR;
    Position(1) = 0.5 - Position(3)/2;
    Position(2) = 0.5 - Position(4)/2;
else
    Position(3) = 0.5;
    Position(4) = Position(3) / AR;
    Position(1) = 0.5 - Position(3)/2;
    Position(2) = 0.5 - Position(4)/2;
end
set(0,'Units','norm');
% Drawing the picture, pressing any key launches "edit" function.
M = figure('menubar','none',...
    'Units','norm', ...
    'Position',Position, ...
    'name','Anaglyph 3D',...
    'ResizeFcn',@scaler,...
    'numbertitle','off',...
    'keypressfcn',@edit);
image(resultImage);

'image. The resulting 3D anaglyph stereo image can be adjusted';
'by shifting and/or rotating one of the input images.';
'After such adjustment, the result can be saved to a JPG image file.';
'';
'Function syntax: anaglyph3d(image1, image2, filter1, filter2)';
'';
'The following keys are used for adjustments and saving:';
'';
'arrowkey "up" - shifts "image2" upwards,';
'arrowkey "down" - shifts "image2" downwards,';
'arrowkey "left" - shifts "image2" to the left,';
'arrowkey "right" - shifts "image2" to the right,';
' "+"(plus) or "x" - rotates "image2" anti-clockwise,';
' "-"(minus) or "c" - rotates "image2" clockwise,';
' "space" - swaps the filters (e.g. "red-blue" to "blue-red"),';
' "s" - saves the stereo anaglyph image to "output.jpg",';
' "h" - opens help window.';
'';
'Author: Matej Mikulsky, October 2016';
'email: matej.mikulsky@student.tuke.sk';
'Student, Technical University, Kosice, Slovakia'};
% Getting screen information.
set(0,'Units','norm');
screen_dim=get(0,'ScreenSize');
bl=screen_dim(3)/2 - 0.15;
bb=screen_dim(4)/2 - 0.2;
% Drawing a window.
Saved = figure('Units','norm', ...
    'Color',[0.8 0.8 0.8], ...
    'MenuBar','none', ...
    'Name','Help', ...
    'NumberTitle','off', ...
    'Position',[bl bb 0.35 0.5], ...
    'RendererMode','manual', ...
    'Tag','Saved');
% Drawing a text from our string variable helpString.

```



Figure [III.5] : résultats de la stéréoscopie de l'objet 1.



Figure [III.6] : résultats de la stéréoscopie de l'objet 2.

### 3.4. L'estimation de carte de disparité

La méthode utilisée pour exécuter la phase de mise en correspondance dépend de la méthode SAD (différences absolues totales) et la similarité entre deux blocs différents est obtenue en calculant la différence absolue entre chaque unité de pixels de la première image et les pixels correspondants de la seconde image. Ce processus est répété pour calculer la différence absolue entre le bloc d'image Le premier et tous les blocs d'une même ligne polaire dans la deuxième image, à ce moment entre les différentes valeurs obtenues, la valeur minimale de la différence signifie que ces deux blocs sont identiques.

L'algorithme d'appariement fournit une carte de disparité, une image numérique contenant des informations de pixel en pixel entre deux images. Le contenu de cette carte signifie que les pixels avec des contrastes importants sont plus proches de la caméra (très loin du fond de l'objet) et que les pixels avec des disparités plus faibles sont éloignés de la caméra (plus près du fond de l'objet).

La carte de disparité est ainsi calculée par stéréoscopie, et nous avons obtenus les résultats suivant :

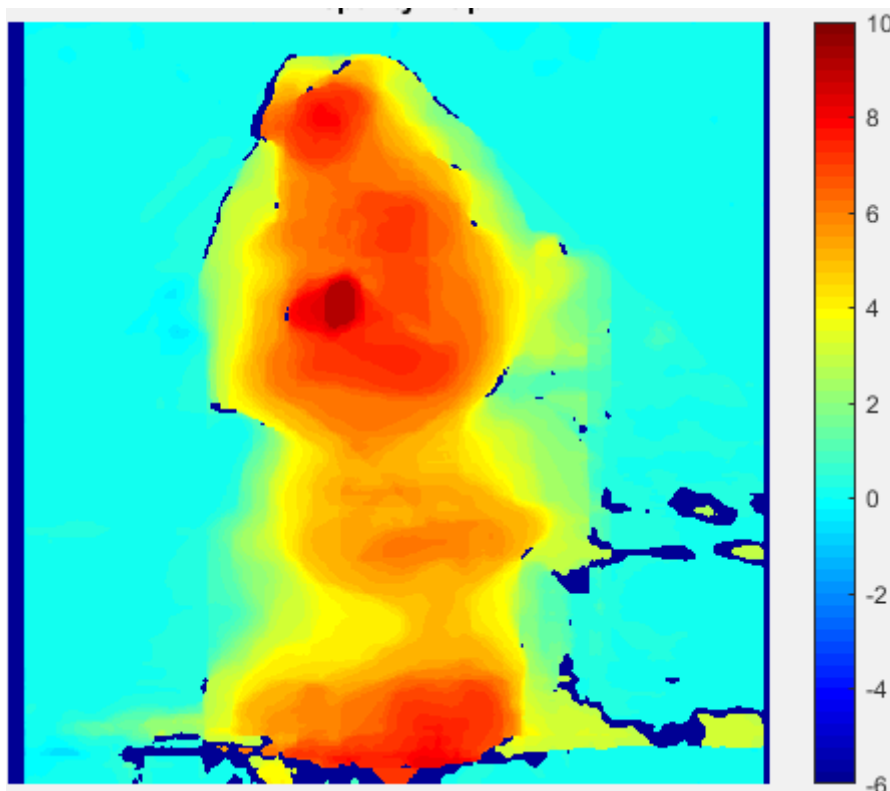


Figure [III.7] : Cartes de disparité résultantes des objets utilisés.

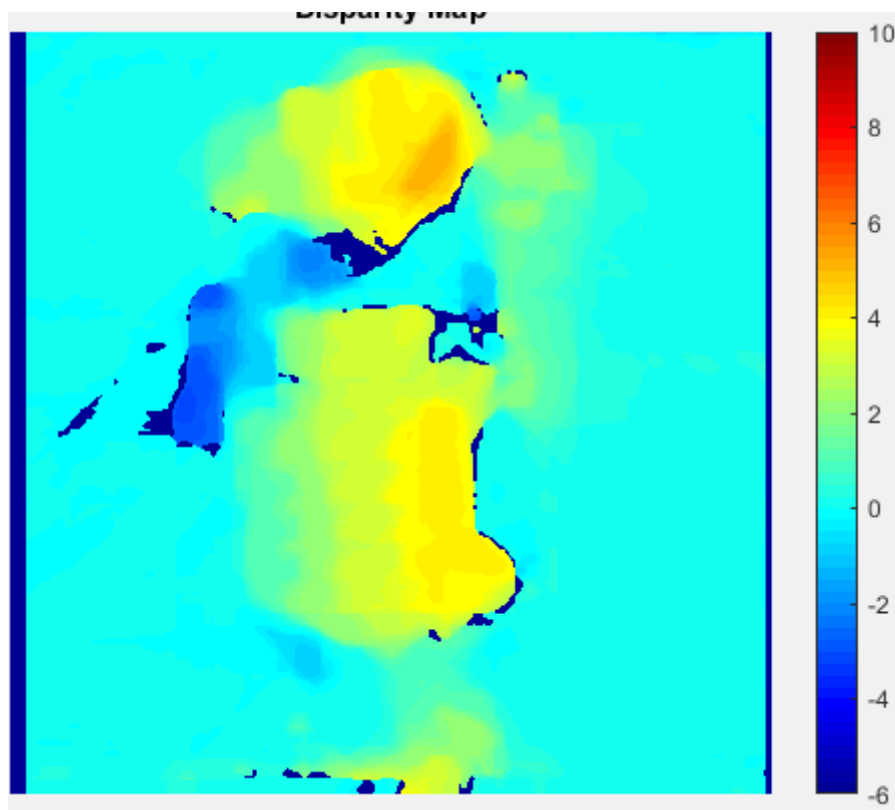


Figure [III.8] : Cartes de disparité résultantes des objets utilisés.

Le degré de couleur représenté dans la figure [III.7,8] signifie que l'objet est plus proche de la caméra pour les régions rouges, le jaune jusqu'à la couleur verte veut dire que la zone est plus encore éloignée. En arrivant à la couleur bleue qui désigne que cette partie est la plus écartée par rapport à la caméra.

### 3.5. Inférence de la troisième dimension à partir de la carte de disparité

Comme nous l'avons mentionné précédemment, la reconstruction tridimensionnelle nécessite deux étapes : stéréoscopie et triangulation, mais nous essayons dans ce mémoire de réaliser une reconstruction tridimensionnelle sans passer par la triangulation.

Ici, nous utilisons principalement la carte de disparité, qui résulte après l'application de la stéréoscopie sur les deux images sources.

En premier lieu, nous remplaçons la couleur du fond (dans la matrice de disparité) par la valeur 0, pour éviter par la suite de calculer leurs relatives profondeurs.

Et puis, tout le reste de la matrice sera multiplié par un facteur pour obtenir un ensemble de profondeurs approximatives de tous les pixels appartenant à l'objet. Nous obtenons ainsi la profondeur perdue de chaque pixel de coordonnées  $(x,y)$  dans l'image originale.

Cela nous donne une nouvelle matrice qui représente qui contient bien la troisième dimension que nous cherchons pour l'utiliser pour former une image en trois dimensions. L'algorithme de cette étape est comme suit :

```

C=disparityMap;
[x,y]=size(C);
M=zeros(x,y);

for i=1:x
    for j=1:y

        if (C(i,j)>0)

            M(i,j) = C(i,j);

        end
    end
end

O=M*23;

[m,n]=size(O);

Z=[];

for i=1:m
    for j=1:n

        %Z(i,j)=((O(i,j)*47)/C(i,j));

        Z(i,j)=((O(i,j)*76)/9);
    end
end

```

### 3.6. Positionnement des points dans l'espace :

À ce stade, nous utilisons les trois dimensions extraites de la carte de disparité et des images sources ; pour positionner chaque point résultant dans l'espace. Ce qui signifie la reconstruction de toute la partie apparente de l'objet à reconstruire. En Matlab, l'utilisation de fonction prédéfinis facilite beaucoup la tâche. Nous pouvons utiliser les fonctions « plot » pour représenter notre objet dans l'espace, et nous avons bien obtenus les résultats suivants :



```
i1=[];  
j1=[];  
z1=[];  
for i=1:x  
    for j=1:y  
        if (Z(i,j)~=0)  
            i1=[i1 i];  
            j1=[j1 j];  
            z1=[z1 Z(i,j)];  
        end  
    end  
end  
figure  
plot3(i1,j1,z1);  
figure  
scatter3(i1,j1,z1)
```

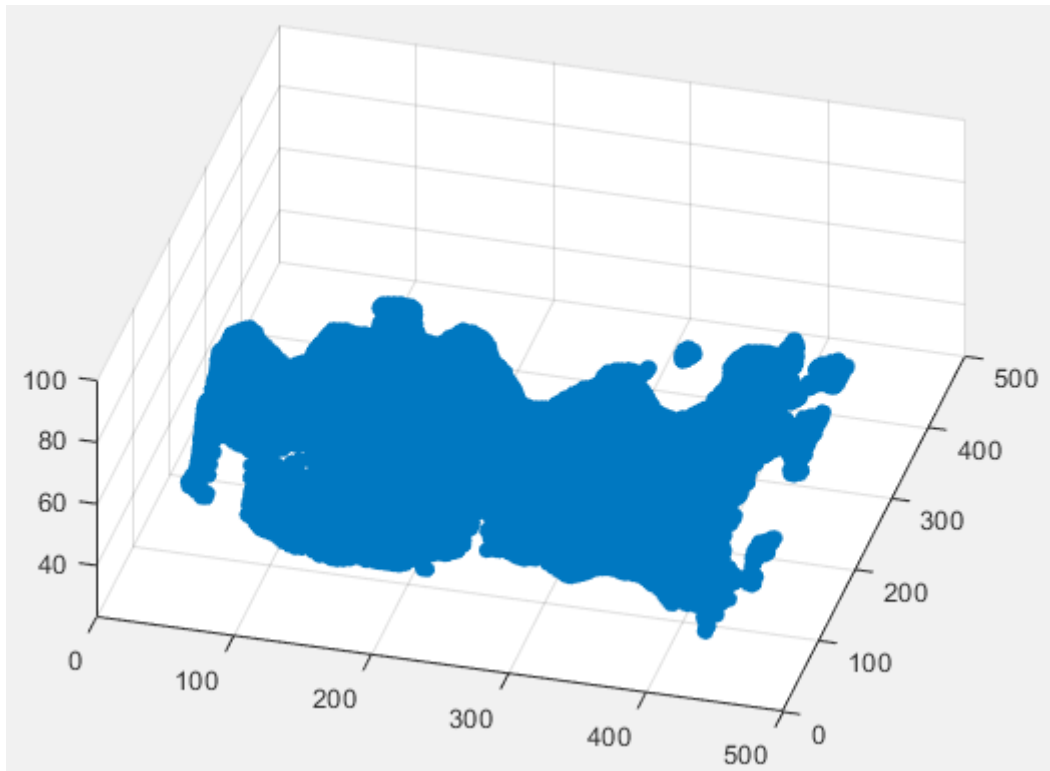


Figure [III.8] : les résultats de reconstruction 3d sans triangulation de l'objet 1.

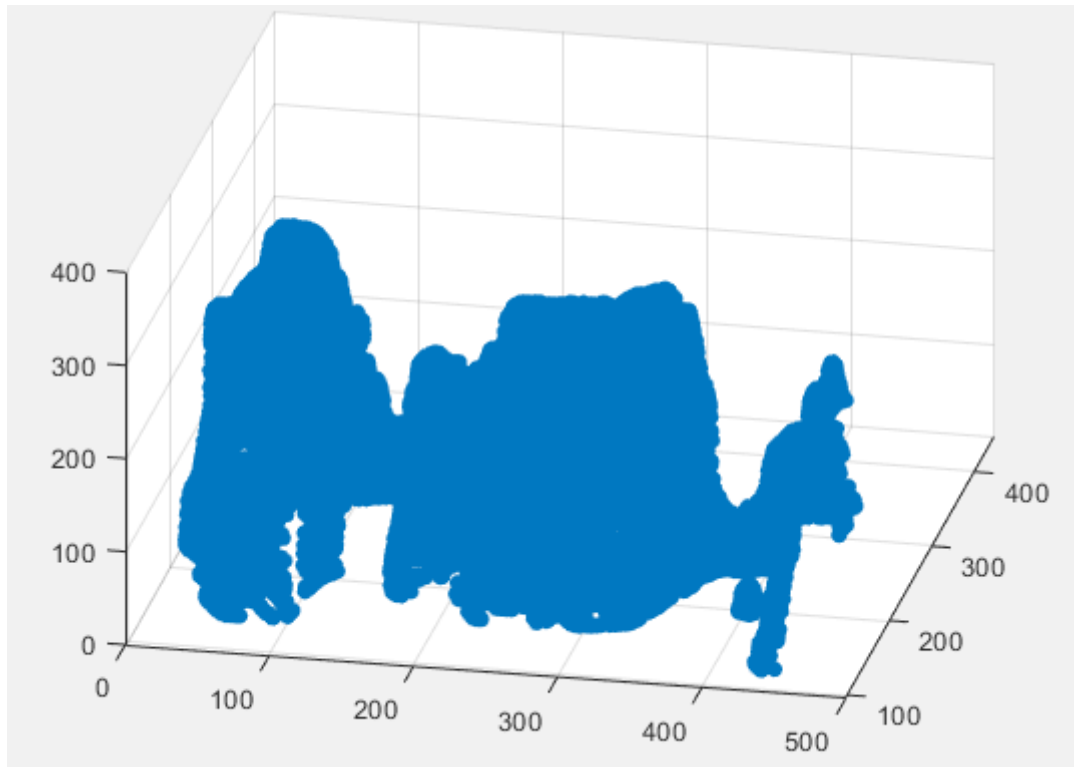


Figure [III.9] : les résultats de reconstruction 3d sans triangulation de l'objet 2.

Le tableau suivant résume les différentes valeurs comparatives des temps de calculs obtenus pour ces deux approches :

[III.1] : Tableau comparatif de temps de calcul de reconstruction 3d des deux objets.

Tâche	Temps de calculs en secondes	
	Minion	Smurf
StereoAnaglyph	0.135 s	0.098 s
DisparityMap	0.135 s	0.103 s
Scatter3	0.011 s	0.127 s
Total temp	96.240 s	50.237 s

## Conclusion

Dans ce chapitre, nous avons proposé une méthode de reconstruction tridimensionnelle, mais sans triangulation et n'utilisant que la carte de disparité qui le résultat de la stéréoscopie

Nous avons expliqué en détail l'approche de notre travail, allant de la prise de vue à la reconstruction de l'objet visé, en passant par l'extraction des informations significatives.

Nous avons ainsi utilisé des images d'objets réels, sur lesquelles nous avons appliqué la technique de stéréoscopie. Et nous avons pu obtenir des reconstructions 3D partielles des objets capturés. Les résultats obtenus, montrent bien qu'il est possible de reconstruire partiellement des objets directement en utilisant la carte de disparité (bidimensionnelle) et sans nécessité de passer par une phase de triangulation ardue et gourmande en ressources. Il suffit pour cela d'établir une relation directe entre la carte de disparité et les profondeurs réelles de scène capturée. Le bon choix de la formule ou paramètre, permet d'aboutir à des résultats acceptables dans le contexte de la vision par ordinateur.

La qualité des résultats que nous avons obtenus ont un aspect visuel médiocre, mais peuvent être améliorés par études plus approfondies et avec la consécration du temps suffisant maîtriser tous les aspects inhérents au domaine de recherche. En plus, la rigidité du Matlab en ce qui concerne les couleurs des plots utilisés pour la visualisation des courbes, affaiblit beaucoup l'aspect visuel de l'objet reconstruit (l'objet est représenté comme étant un seul lot, tandis qu'en réalité il s'agit d'un ensemble de points 3D).

Pour conclure ce chapitre, nous disons que ce travail nous a bien permis de s'introduire dans le domaine vaste de reconstruction 3D, et nous avons plusieurs aspects tels-que : la reconstruction 3D et ces diverses approches, la stéréovision, la triangulation...etc.

# **Conclusion Générale**

Nous nous sommes intéressés à la problématique de la reconstruction 3D à partir d'images qui appartient au domaine d'infographie et de vision par ordinateur, et qui représente un sujet de recherche très actif dans les dernières années. La modélisation 3D classique consiste à générer des modèles 3D en suivant deux phases essentielles, une phase de modélisation et une phase de rendu. Cette approche est assez complexe, surtout dans le cas des grandes scènes ; de plus les logiciels de modélisation sont coûteux et nécessitent une phase d'apprentissage et des utilisateurs experts.

Dans ce mémoire nous avons présenté quelques notions de base liées au domaine tel que : le système de coordonnées, le processus de la formation de l'image et le modèle de caméra.

Nous avons aussi présenté le principe de la reconstruction 3D à partir d'images qui est connue par le terme IBMR (Image Based Modeling and Rendering). Ce type d'approches consiste à calculer la troisième dimension à partir des informations disponibles dans les photographies avec le moindre de coût et avec une meilleure qualité.

En nous initiant à ce domaine, nous avons présenté une classification multistandard des différentes méthodes de reconstruction tridimensionnelle existantes, et cette classification permet de choisir la méthode de reconstruction tridimensionnelle appropriée pour le cas étudié.

Les méthodes actuelles d'IBMR incluent la stéréoscopie et la triangulation. Après avoir étudié ces deux méthodes, nous avons constaté qu'elles sont largement utilisées en adjonction, induisant ainsi à des temps de calcul très élevés. Et nous supposons qu'elles peuvent être séparées pour obtenir une reconstruction 3D qu'à partir des résultats de la stéréovision de façon plus rapide (sans nécessité d'application d'une phase de stéréoscopie).

Dans notre cas, nous avons proposé une méthode de reconstruction 3D sans triangulation qui consiste aussi à utiliser directement la carte de disparité étant le résultat d'une applique stéréoscopique. Cette réduction, permet de réduire des calculs supplémentaires et bien compliqués.

Pour vérifier l'exactitude de notre approche, nous avons pris deux photos (décalées) de différents objets, d'en calculer les correspondants de chaque point par stéréoscopie et finalement d'inférer une forme 3D partielle et approximative de l'objet étudié. Cela permet non seulement de réduire temps de calcul, mais la forme 3D obtenus ne contient pas de grandes parties hors l'objet reconstruit.

Après avoir extrait la matrice sur laquelle nous avons travaillé et renvoyé les valeurs négatives, nous utilisons une équation sur la matrice résultante. Cette équation est représentée en appliquant la règle des trois points à la matrice, nous multiplions les valeurs de la carte de disparités par un facteur relatif au nombre de pixels décalant les deux points. De cette façon, nous obtenant une nouvelle matrice  $Z$  qui contient les profondeurs approximatives de l'objet à reconstruire, et il ne reste que d'afficher tous les points  $p(x,y,z)$  dans l'espace par l'utilisation de courbe 3D.

# Références

- 1- Web site : [https://en.wikipedia.org/wiki/3D\\_reconstruction](https://en.wikipedia.org/wiki/3D_reconstruction)
- 2- Olivier Godin, 'cour de Transformations géométriques', Université de Sherbrooke, 22 janvier 2014
- 3- Olivier Faugeras, Three-Dimensional Computer Vision: A Geometric Viewpoint, The MIT Press, 1993.
- 4- Dr. Gabriele Bleser, Didier Stricker, '3D Computer Vision', 2002
- 5- R. I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2000.
- 6- Site web: [http://www.optique.ingenieur.org/fr/cours/OPI\\_fr\\_M04\\_C01/co/OPI\\_fr\\_M04\\_C01\\_web\\_1.html](http://www.optique.ingenieur.org/fr/cours/OPI_fr_M04_C01/co/OPI_fr_M04_C01_web_1.html)
- 7- James L. Crowley, 'Formation et Analyse d'Images', ENSIMAG 3 - MMIS Option MIRV, 19 Oct 2010
- 8- François Cabestaing, 'Éléments de géométrie projective application à la calibration de caméra', université de Lille, 2008
- 9- Brice Michoud, 'Reconstruction 3D à partir de séquences vidéo pour l'acquisition du mouvement de personnages en temps réel et sans marqueur', Doctorat de l'Université de Lyon délivré par l'Université Claude Bernard -Lyon, 27 juin 2011
- 10- B. BOUFAMA, 'Reconstruction tridimensionnelle en vision par ordinateur' : cas des caméras non étalonnées, L'institut national polytechnique de Grenoble, 2004.
- 11- M. C. BABAHENINI, A. HEMIDI & N. DJEDI, 'Rendu Volumique efficace par une représentation à base de couches d'images', Département d'Informatique, Université Mohamed Khider de Biskra, 2005
- 12- M. C. BABAHENINI, A. HEMIDI & N. DJEDI, 'Rendu Volumique efficace par une représentation à base de couches d'images', Département d'Informatique, Université Mohamed Khider de Biskra, 2005
- 13- Manuel M. OLIVEIRA, 'Image-Based Modeling and Rendering Techniques: A survey', UFRGS, Porto Alegre, RS, Brazil, 2002
- 14- R. ZHANG, Ping-sing. TSAI, J. Edwin et M. SHAH, 'Shape from shading: A survey', University of central Florida, 1999
- 15- B.K.P. Horn. 'Obtaining shape from shading information'. In P.H. Winston, éditeur, The Psychology of Computer Vision. McGraw-Hill, New York, 1975.
- 16- A. MEYER, H. BRICENO PULIDO et S. BOUAKAZ, 'User-guided Shape from Shading to Reconstruct Fine Details from a Single Photograph'. Asian Conference on Computer Vision, 2007.
- 17- Emmanuel Prados, Olivier Faugeras. Shape From Shading. N. Paragios, Y. Chen and O. Faugeras. Handbook of Mathematical Models in Computer Vision, Springer, 2006,
- 18- G. HARO, 'Shape from Silhouette Consensus', Université Pompeu Fabra, Espagne, 2014

- 19- A.M. LOH et R. HARTLEY, 'Shape from non-homogeneous, non-stationary, anisotropic, perspective texture'. Proc. Vision Conference, pages 69–78, 2005.
- 20- A. BOTTINO, A. LAURENTINI, 'A silhouette based technique for the reconstruction of human movement', 2001.
- 21- P. DAMIEN, 'les imposteurs', 2005.
- 22- MATTHIEU DUJARDIN. Le scanner laser 3D : reconnaissance de formes et modélisation de déformations. Engineering Sciences. 20130
- 23- D. ROUABHIA, M.C. BABAHENINI, 'Multi-criteria classification of 3D reconstruction approaches', Mohamed Khider University, Biskra, Algeria, 2009.
- 24- ALBAN BAJARD. Numérisation 3D de surfaces métalliques spéculaires par imagerie infrarouge. Autre [cs.OH]. Université de Bourgogne, 2012.
- 25- A. BOULCH Reconstruction automatique des maquettes numériques 3D Thèse d'université Paris. 2014.
- 26- N. CHAHATA, F. JUNG, M. Pierrot-Deseilligny, J. STAMON, 'Détection et reconstruction de facettes 3D par approche hiérarchique par régions à partir de couples d'images satellite en milieu urbain', Conférence internationale (Science of electronic), Tunisie, 2004.
- 27- M. LEVOY, 'The Digital Mic Project: 3D Scanning of Large Statues', Proc. Siggraph, ACM, 2000.
- 28- D. SCHARSTEIN, R. SZELISKI, 'High-accuracy stereo depth maps using structured light', Middlebury Coll, VT, USA, 2003
- 29- D. NISTÉR, 'Automatic passive recovery of 3D from images and video' 3DPVT, 2004
- 28- Pascal Granger, La Stéréoscopie ou la photographie en 3D, Vice-président du Groupe Stéréoscopie Franco-Suisse de Genève, 7 février 2014.
- 29- D. MARR, 'A computational investigation into the human representation and processing of visual information', Freeman, 1982
- 30- R. HORAUD, O. MONGA Vision par ordinateur outils fondamentaux. Hermès, 1995.
- 31- Z. hengyou ZHANG, 'A flexible new technique for camera calibration', IEEE Trans. Pattern Anal. Mach. Intell, 2000.
- 32- E. R. Davies, Machine Vision: Theory, Algorithms, Practicalities. Academic Press, 1997.
- 33- Site web : <http://perso.univ-lemans.fr/~berger/CoursStereoVision/co/essentFond.html>
- 34- P. DEBEVEC, C. J. TAYLOR, J. MALIK. Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. SIGGRAPH, 1996.
- 35- Christophe Rabaud. Une nouvelle approche de mise en correspondance stéréoscopique dense par méthodes possibilistes. Traitement du signal et de l'image [eess.SP]. Université Montpellier II - Sciences et Techniques du Languedoc, 2005. Français.



- 36- G. GALES, 'Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions', Université Toulouse 3, 2011.
- 37- Guillaume Gales. Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions. Interface homme-machine [cs.HC]. Université de Toulouse, 2011. Français.
- 38- G. GALES, 'Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions', Université Toulouse 3, 2011.
- 39- N. FAKHFAKH, ' Détection et localisation tridimensionnelle par stéréovision d'objets en mouvement dans des environnements complexes : application aux passages à niveau', Ecole centrale de Lille, 2012.