

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**

**Université Mohamed Khider de Biskra**  
**Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie**  
**Département d'informatique**

N° d'ordre :.....



Thèse présentée en vue de l'obtention du diplôme de  
**Doctorat en sciences en : Informatique**

***Acquisition 3D et visualisation d'objets culturels  
pour les applications de la réalité augmentée.***

Présentée par :

**ROUABHIA Djaber**

Soutenue publiquement le : 07/05/2023

Devant le jury composé de :

BABAHENINI Mohamed Chaouki	Professeur	Président	Université de Biskra
DJEDI NourEddine	Professeur	Rapporteur	Université de Biskra
LAOUAR Mohamed Ridda	Professeur	Examineur	Université de Tébessa
BENDJENNA hakim	Professeur	Examineur	Université de Tébessa

Année universitaire 2022/2023

---

# ***Table des matières***

---

<b>PROBLEMATIQUE</b>	<b>2</b>
<b>MOTIVATION</b>	<b>3</b>
<b>ORGANISATION DE LA THESE</b>	<b>4</b>
<b>INTRODUCTION</b>	<b>5</b>
<b>1. SYSTEMES DE COORDONNEES</b>	<b>5</b>
COORDONNEES HOMOGENES	5
<b>2. TRANSFORMATIONS USUELLES</b>	<b>6</b>
2.1 TRANSLATION	6
2.2 CHANGEMENT D'ECHELLE	6
2.3 ROTATION	6
<b>3. MODELISATION DE LA CAMERA</b>	<b>7</b>
3.1 TRANSFORMATION SCENE-CAMERA	7
3.2 PROJECTION CAMERA-RETINE	7
3.3 TRANSFORMATION RETINE-IMAGE	8
<b>4. FORMATION DE L'IMAGE</b>	<b>8</b>
<b>5. RELATIONS GEOMETRIQUES ENTRE LES IMAGES.</b>	<b>9</b>
GEOMETRIE EPIPOLAIRE.	9
<b>6. RELATION ENTRE VIDEO ET IMAGES</b>	<b>10</b>
<b>CONCLUSION</b>	<b>11</b>
<b>INTRODUCTION</b>	<b>12</b>
<b>1. METHODES DE MODELISATION ET RENDU A BASE D'IMAGE : IBMR</b>	<b>12</b>
<b>2. QUELQUES APPROCHES D'IBMR</b>	<b>14</b>
2.1 FORME A PARTIR X ( <i>SHAPE FROM X</i> )	14
2.1.1 <i>Forme à partir d'ombres (Shape from shading)</i>	14
2.1.2 <i>Forme à partir silhouette (Shape from silhouette)</i>	14
2.1.3 <i>Forme à partir de textures (Shape from texture)</i>	15
2.1.4 <i>Forme à partir du mouvement (Shape from motion)</i>	15
2.1.5 <i>Forme à partir Focus/ Defocus (Shape from focus/defocus)</i>	16
2.2 LES IMPOSTEURS	16
2.2.1 <i>Imposteurs statiques</i>	17
2.2.2 <i>Imposteurs dynamiques</i>	17
<b>3. CLASSIFICATION MULTICRITERES DES METHODES DE RECONSTRUCTION 3D</b>	<b>18</b>
3.1 NIVEAU D'AUTOMATICITE	19
3.1.1 <i>Reconstruction automatique</i>	19
3.1.2 <i>Reconstruction semi-automatique</i>	19
3.1.3 <i>Reconstruction manuelle</i>	19
3.2 NIVEAU DE COMPLEXITE DE LA SCENE	20
3.2.1 <i>Grandes scènes</i>	20
3.2.2 <i>Scènes moyenne</i>	20
3.2.3 <i>Petites scènes</i>	20
3.3 MOUVEMENT DE LA SCENE	21
3.3.1 <i>Scènes statiques</i>	21
3.3.2 <i>Scènes dynamiques</i>	21
3.4 TYPES D'APPROCHES UTILISEES	21
3.4.1 <i>Approche ascendante</i>	21
3.4.2 <i>Approche descendante</i>	21

3.4.3 Approche hybride	21
3.5 PASSIVE OU ACTIVE	22
3.5.1 Les méthodes actives	22
3.5.2 Les méthodes passives	23
3.6 MULTIPLICITE DES OBJETS	23
3.6.1 Reconstruction mono-objet	23
3.6.2 Reconstruction multi-objet	24
3.7 DEGRE DE PRECISION	24
3.7.1 L'apparence visuelle	24
3.7.2 Exigences de description	24
3.7.3 Exactitude du modèle géométrique	24
3.7.4 Plusieurs aspects à la fois	24
3.8 LE DEGRE DE GENERALITE DE LA MODELISATION 3D	25
3.8.1 Reconstruction projective	25
3.8.2 Reconstruction euclidienne	25
3.8.3 Reconstruction affine	25
3.8.4 Reconstruction métrique	26
<b>CONCLUSION</b>	<b>27</b>
<b>INTRODUCTION</b>	<b>28</b>
<b>1. L'ENVELOPPE VISUELLE</b>	<b>28</b>
1.1 PROBLEMES LIES A L'ENVELOPPE VISUELLE	29
1.2 EXTRACTION DES SILHOUETTES	30
1.2.1 Méthodes basées différence de pixels	30
1.2.2 Méthodes basées régions	30
1.2.3 Méthodes basées contours	30
1.3 METHODES DE CALCUL D'ENVELOPPE VISUELLE	31
1.3.1 Les approches surfaciques	31
1.3.2 Les approches volumiques	31
<b>2. LA STEREOVISION</b>	<b>32</b>
2.1 CALIBRATION	32
2.1.1 Paramètres intrinsèques	33
2.1.2 Paramètres extrinsèques	33
2.1.3 Matrice essentielle	33
2.1.4 Matrice fondamentale	33
2.1.4.1 Calibrage classique (hors ligne)	33
2.1.4.2 Auto-calibration (en ligne)	34
2.2 LA MISE EN CORRESPONDANCE	34
2.2.1 Primitives stéréoscopiques	34
2.2.2 Notion de la disparité	34
2.2.2.1 Vecteur de disparité	34
2.2.2.2 Carte de disparité	35
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
2.2.4.2 Contraintes figurales	39
2.2.5 Techniques de mise en correspondance	40
2.2.5.1 Méthodes d'appariement denses	40
2.2.5.2 Méthodes d'appariement éparées	40
2.2.5.3 Méthodes d'appariement locales	40
2.2.5.3 Méthodes d'appariement globales	42
2.3 RECONSTRUCTION 3D	42
2.3.1 La triangulation	43

<b>CONCLUSION</b>	<b>44</b>
<b>INTRODUCTION</b>	<b>45</b>
<b>2. TRAVAUX DE ERIC NOWARK ET AL</b>	<b>45</b>
AVANTAGES ET INCONVENIENTS	47
<b>3. TRAVAUX DE MING LI ET AL</b>	<b>48</b>
AVANTAGES ET INCONVENIENTS	49
<b>4. TRAVAUX DE MATSUMOTO ET AL</b>	<b>50</b>
AVANTAGES ET INCONVENIENTS :	56
<b>5. CONCLUSION</b>	<b>56</b>
<b>INTRODUCTION :</b>	<b>57</b>
<b>1. REALITE AUGMENTEE :</b>	<b>59</b>
<b>2. TECHNOLOGIE DERRIERE LE FONCTIONNEMENT DE LA REALITE AUGMENTEE :</b>	<b>60</b>
S.L.A.M	61
<b>3. TYPES DE REALITE AUGMENTEE :</b>	<b>62</b>
3.1 AR A BASE DE MARQUEURS (RECONNAISSANCE)	62
3.2 AR SANS MARQUEURS	62
3.3 AR PAR PROJECTION	63
3.4 AR BASEE SUR LA SUPERPOSITION	63
3.5 APPROCHE FONDEE SUR L'EMPLACEMENT :	63
3.6 TECHNOLOGIE DE SUIVI DE PROFONDEUR :	64
<b>4. COMPOSANTES COMMUNES DE LA AR</b>	<b>64</b>
<b>5. AVANTAGES DE LA REALITE AUGMENTEE</b>	<b>64</b>
<b>6. EXEMPLES D'APPLICATIONS DE LA REALITE AUGMENTEE :</b>	<b>65</b>
<b>7 PATRIMOINE VIRTUEL, HERITAGE CULTUREL VIRTUEL</b>	<b>66</b>
<b>7. TRAVAUX CONNEXES :</b>	<b>68</b>
7.1 RECONSTRUCTION TRIDIMENSIONNELLE	68
7.2 PATRIMOINE NUMERIQUE	70
7.3 MUSEES VIRTUELS	71
7.4 EXPERIENCE UTILISATEUR	73
7.5 EDUCATION	74
7.6 TOURISME	76
7.7 JEUX	78
7.8 PATRIMOINE CULTUREL INTANGIBLE	80
<b>CONCLUSION</b>	<b>80</b>
<b>1 INTRODUCTION</b>	<b>84</b>
<b>2 TRAVAUX CONNEXES</b>	<b>85</b>
<b>3 METHODE PROPOSEE</b>	<b>86</b>
3.1 EXTRACTION D'IMAGES ET DE SILHOUETTES	87
3.2 ISOLATION D'OBJET ET EXTRACTION DE SES SILHOUETTES	89
3.3 RESTRICTION DE LA ZONE DE CORRESPONDANCE	89
3.4 ÉTAPE FINALE	91
<b>4 RESULTATS EXPERIMENTAUX</b>	<b>92</b>
4.1 ENVIRONNEMENT MATERIEL ET DE PROGRAMMATION	92

## Table des matières

4.2 EXTRACTION D'IMAGES DE LA VIDEO	93
4.3 EXTRACTION DE SILHOUETTES	93
4.4 STEREOSCOPIE SANS NOTRE METHODE DE RESTRICTION	93
4.5 APPLICATION DE NOTRE APPROCHE RESTRICTIVE	94
4.6 COMPARAISON	94
4.7 CARTE DES DISPARITES	96
<b>5 DISCUSSION</b>	<b>96</b>
<b>6 CONCLUSIONS</b>	<b>97</b>
<b>REFERENCES</b>	<b>80</b>

---

# *Liste des figures*

---

Figure 1 : Système de coordonnées.	5
Figure 2 : Les trois transformations élémentaires du modèle sténopé, et les repères associés.	7
Figure 3 : Modèle de la caméra.	8
Figure 4 : Géométrie épipolaire.	9
Figure 5: Relation géométrique entre deux images : Géométrie épipolaire.	10
Figure 6 : Relation géométrique entre trois images : Trilinearité.	10
Figure 7 : Modélisation basée images.	13
Figure 8 : Rendu basé image.	13
Figure 9 : 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.	14
Figure 10 : Shape from silhouette à partir 3 vues différentes.	15
Figure 11 : (a) Texture extraite d'une robe, (b) surface reconstruite de la texture.	15
Figure 12 : Forme à partir de focus/defocus.	16
Figure 13 : De gauche en droite Modèle original, ensemble optimal de rectangles, plaquage de texture, Panneau d'affichage rendu.	17
Figure 14 : Imposteur. A gauche, l'objet original, à droite, l'imposteur.	17
Figure 15 : Exemple d'une voiture reconstruite automatiquement.	19
Figure 16 : Reconstruction semi-automatique : Exemple de détection des vertices et arrêtes pour obtenir le modèle final.	19
Figure 17 : Exemple de reconstruction 3D d'une ville.	20
Figure 18 : Modèle 3D d'un bâtiment service publiques.	20
Figure 19 : Modèle de reconstruction 3D d'une petite scène.	21
Figure 20 : Télémètre Laser.	22
Figure 21 : Lumière structurée illuminant un visage.	22
Figure 22 : Simple exemple du système de stéréovision.	23
Figure 23 : Exemple de reconstruction mono-objet.	23
Figure 24 : Regroupement des éléments reconstruits.	24
Figure 25 : Résultat d'une reconstruction projective.	25
Figure 26 : Reconstruction euclidienne.	25
Figure 27 : Résultat d'une reconstruction affine.	26
Figure 28 : Résultat d'une reconstruction métrique.	26
Figure 29 : Un objet 2D représenté avec une vue, sa silhouette et le cône correspondant.	28
Figure 30 : $v^\infty$ et $V_r$ pour un objet 2D : $v^\infty$ est l'enveloppe convexe de l'objet.	29
Figure 31 : Exemple d'objet avec une zone concave.	29
Figure 32 : Exemple d'objet avec fantôme.	30
Figure 33 : Illustration d'une coupe de la géométrie estimée d'un objet $O$ , en rouge enveloppe visuelle surfacique et bleue une enveloppe visuelle volumique.	31
Figure 34 : (a) Géométrie épipolaire quelconque. (b) Géométrie épipolaire rectifiée.	32
Figure 35 : Modèle géométrique de prise de vues.	33
Figure 36 : Vecteur de disparité.	35
Figure 37 : Exemple de carte de disparité.	35
Figure 38 : Zone occultée – Le pixel marqué d'un point rouge a son correspondant occulté dans l'image de droite-	36
Figure 39 : 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.	36
Figure 40 : Zone homogène –Manque d'information qui caractérise le pixel-.	36
Figure 41 : Zone avec textures répétitives –Pour un pixel il existe plusieurs candidats dans l'autre image.	37
Figure 42 : Géométrie épipolaire.	37
Figure 43 : Exemple de la rectification épipolaire.	38
Figure 44 : Un point image correspond à un point objet est un seul.	38
Figure 45 : Contrainte d'ordre.	39
Figure 46 : Contrainte de continuité figurale.	40
Figure 47 : Mise en correspondance locale.	41
Figure 48 : Système de projection des scène –Représentation de triangle formé par le système.-.	43
Figure 49 : Exemple des résultats obtenus par l'approche de Nowark et al.	47
Figure 50 : Exemple de modélisation des zones concaves.	47



<i>Figure 51 : Trois paires de caméras sont disposées le long d'un arc.</i>	48
<i>Figure 52 : Démarche de l'approche proposée. (a) : Extraction des silhouettes. (b) : Création d'enveloppe visuelle. (c) Estimation de profondeur à partir de la stéréovision.</i>	48
<i>Figure 53 : Résultat de modélisation de la méthode de Li et al (Exemple 1).</i>	49
<i>Figure 54 : Résultat de modélisation de la méthode de Li et al (Exemple 2).</i>	50
<i>Figure 55 : Shape from silhouette basé vote.</i>	52
<i>Figure 56 : Vote sur l'espace voxel : la densité montre le score de chaque voxel.</i>	52
<i>Figure 57 : (a) Estimation de surface partielle. (b) Génération totale de modèle.</i>	53
<i>Figure 58 : (a) Estimation de surface partielle (b) Génération totale du modèle.</i>	53
<i>Figure 59 : Présentation du système.</i>	54
<i>Figure 60 : Exemples d'images originales et de résultats de modélisation (Ecureuil)</i>	55
<i>Figure 61 : Exemples d'images originales et de résultats de modélisation (Château).</i>	55
<i>Figure 62 : Positionnement de l'AR et la VR</i>	59
<i>Figure 63 : Une vue d'environnement réelle est complétée par des informations numériques.</i>	60
<i>Figure 64 : Un utilisateur naviguant dans un environnement virtuel VR.</i>	60
<i>Figure 65 : Principe de la réalité augmentée.</i>	61
<i>Figure 66 : Une application mobile d'ameublement Réalité augmentée basée sur marqueurs.</i>	62
<i>Figure 67 : Exemple de réalité augmentée sans marqueurs.</i>	63
<i>Figure 68 : Un système d'affichage par projection (hologramme).</i>	63
<i>Figure 69 : Exemple de réalité augmentée basée sur la superposition.</i>	63
<i>Figure 70 : Exemple d'application Réalité augmentée mobile pour l'exploration de la reconstruction 3D</i>	70
<i>Figure 71 : Visualisation et interaction avec l'ensemble de la scène 3D (a) et un artefact virtuel (b). Avec l'aimable autorisation de Popovici et al.</i>	73
<i>Figure 72: Pipeline de la méthode proposée.</i>	86
<i>Figure 73 : Exemple de mini-studio</i>	87
<i>Figure 74 : Objets réels, (a) Minion (b) Schtroumpf</i>	88
<i>Figure 75 : Un sous-ensemble d'images originales</i>	88
<i>Figure 76 : Exemple de silhouettes obtenues</i>	89
<i>Figure 77 : Délimitation verticale</i>	89
<i>Figure 78 : Détection des bords</i>	90
<i>Figure 79 : La zone de recherche effective</i>	91
<i>Figure 80 : Résultats obtenus pour l'objet 'Minion'</i>	94
<i>Figure 81 : Résultats obtenus pour l'objet 'Schtroumpf'</i>	94
<i>Figure 82 : Détails du temps obtenu lors de l'application de la stéréoscopie à chaque image de les 120 images, résultats (a) pour l'objet Minion et (b) pour l'objet Schtroumpf</i>	95
<i>Figure 83 : Cartes de disparités obtenues</i>	96

---

# *Résumé*

---

Au cours des dernières décennies, le domaine de la reconstruction 3D a connu une croissance rapide et de nombreuses techniques ont été présentées par les scientifiques. L'enveloppe visuelle et la stéréovision sont deux de ces techniques, et sont classées parmi les techniques IBMR.

Les enveloppes visuelles utilisent la forme d'un objet dérivée des images prises sous différents angles pour créer un modèle 3D approximatif de l'objet. La stéréovision calcule la dimension 3D (profondeur) en comparant deux images de la même scène qui ont été prises sous deux angles différents. Les pixels correspondants dans les deux images sont ensuite calculés pour déduire une carte de disparité qui est utilisée pour déterminer les profondeurs.

Les deux techniques présentent des inconvénients lorsqu'elles sont utilisées seules, et l'un des objectifs de cette recherche est de surmonter ces problèmes et de développer une technique de reconstruction 3D efficace et robuste. Notre étude a abouti à la proposition d'une méthode innovante pour estimer l'enveloppe visuelle. Deux approches ont été combinées, à savoir, « la correspondance de caractéristiques » et « l'approche par bloc ». Ce qui a permis de reconstruire des objets avec des gains considérables de temps et sans perte de qualité.

Nous avons utilisé les informations géométriques présentes dans les images pour réduire l'espace de recherche des algorithmes stéréoscopiques, réduisant ainsi le temps d'exécution de plus de la moitié du temps initial.

Dans le cadre de cette recherche, nous avons eu l'occasion de reconstruire des objets culturels réels et d'intégrer leur modèles 3D à des applications de réalité augmentée.

**Mots clés :** Reconstruction 3D, IBMR, enveloppe visuelle, stéréovision et mise en correspondance, Héritage culturel, Réalité augmentée.

---

ملخص

---

في العقود الأخيرة، نما مجال إعادة النمذجة ثلاثية الأبعاد بسرعة كبيرة وقدم العلماء العديد من التقنيات. المغلف البصري والتجسيم ثلاثي الأبعاد هما اثنان من هذه التقنيات، ويتم تصنيفهما على أنهما تقنيات IBMR.

تستخدم المغلفات البصرية شكل جسم مستخرج من الصور المأخوذة من زوايا مختلفة لإنشاء نموذج ثلاثي الأبعاد تقريبي للجسم الحقيقي. يحسب البعد ثلاثي الأبعاد (العمق) من خلال مقارنة صورتين لنفس المشهد مأخوذتين من زاويتين مختلفتين. ثم يتم حساب البكسلات المقابلة في كلتا صورتين لاستنتاج خريطة تفاوت تستخدم لتحديد الأعماق.

كلا التقنيتين لهما عيوب عند استخدامهما منفردتين، وأحد أهداف هذا البحث هو التغلب على هذه المشاكل وتطوير تقنية إعادة نمذجة ثلاثية الأبعاد فعالة وقوية. أسفرت دراستنا عن اقتراح طريقة مبتكرة لتقدير المغلف البصري. وجرى الجمع بين نهجين هما «المناطق المميزة» و «المحيطات». وقد مكن ذلك من إعادة نمذجة ثلاثية الأبعاد لأجسام مع توفير كبير في الوقت ودون فقدان الجودة. استخدمنا المعلومات الهندسية الموجودة في الصور لتقليل مساحة البحث في خوارزميات الستيريو، مما أدى إلى تقليل وقت التنفيذ بأكثر من نصف الوقت الأولي.

كجزء من هذا البحث، أتاحت لنا الفرصة لإعادة بناء قطع ثقافية حقيقية ودمج نماذجها ثلاثية الأبعاد في تطبيقات الواقع المعزز.

**الكلمات المفتاحية:** إعادة نمذجة ثلاثية الأبعاد، IBMR، المغلف البصري، الستيريو والخرائط البعدية، التراث الثقافي، الواقع المعزز.

---

# *Abstract*

---

The field of 3D reconstruction has grown rapidly and many techniques have been presented by scientists. Visual envelope and stereovision are two of these techniques, and are classified as IBMR techniques.

Visual envelopes use the shape of an object derived from images taken from different angles to create an approximate 3D model of the object. Stereovision calculates the dimension (depth) by comparing two images of the same scene taken from two different angles. The corresponding pixels in both images are then calculated to deduce a disparity map that is used to determine depths.

Both methods have disadvantages when used alone, and one of the objectives of this research is to overcome these problems and develop an efficient and robust 3D reconstruction technique. Our study resulted in the proposal of an innovative method for estimating the visual envelope. Two approaches have been combined, namely, “characteristic matching” and “block approach. This made it possible to rebuild objects with considerable time savings and without loss of quality.

We used the geometric information present in the images to reduce the search space of the stereoscopic algorithms, thus reducing the execution time by more than half of the initial time.

As part of this research, we had the opportunity to reconstruct real cultural objects and integrate their 3D models into augmented reality applications.

**Key words:** 3D reconstruction, IBMR, visual hull, stereovision, bloc matching, cultural heritage, augmented reality

---

# *Introduction générale*

---



La vision artificielle vise à reproduire certaines fonctions de la vision humaine en analysant des images. Pour les ordinateurs, c'est un problème assez délicat, puisque l'information disponible dans ce cas est une représentation en deux dimensions d'un environnement en trois dimensions. La troisième dimension est perdue lors de génération d'image par projection.

L'image est utilisée comme support riche en informations dans de nombreux domaines (design, perception, aide à la décision, communication, robotique, etc.). En effet, la projection 3D-2D implique une importante perte d'informations. Le raisonnement 3D est difficile et nécessite une compréhension approfondie des phénomènes visuels. La littérature regorge de travaux différents et surtout de méthodes distinctes, et plusieurs projets de recherche sont en cours à cet égard.

L'analyse d'images regroupe plusieurs disciplines qu'on classe en deux catégories (Baumgart & Bruce G, 1974):

- Processus de bas niveau nécessitant peu d'informations sur le contenu de l'image. Il s'agit de processus de filtrage, d'amélioration et de restauration d'images regroupés sous les termes de traitement d'images et d'extraction d'index.
- Processus de haut niveau en aval des processus de bas niveau qui peuvent nécessiter des informations sur le contenu de l'image. Ce sont des processus cognitifs généraux, la reconstruction tridimensionnelle, la reconnaissance de formes...

Notre thèse intitulée "Acquisition 3D et visualisation d'objets culturels pour les applications de la réalité augmentée." entre dans la deuxième catégorie. Le but est de pallier les problèmes liés à la perte d'informations afin de reconstituer en 3D des objets réels dans un délai satisfaisant sans affecter la qualité des résultats obtenus qui seront par la suite utilisés dans le cadre de la réalité augmentée. La reconstruction 3D est un domaine actif de longue date visant à générer des modèles 3D à partir d'informations provenant de scènes réelles, et la recherche est basée sur l'obtention de formes d'objets assez réalistes. Le principe est de déterminer les coordonnées 3D de n'importe quel point de l'objet.

Il existe plusieurs techniques de reconstruction 3D, selon le type d'objet à reconstruire (forme, dimensions, textures, etc.) et les conditions environnementales (éclairage, scène structurée, etc.). Ces méthodes ont plusieurs critères de classification dont les détails sont clairement présentés dans notre démarche de recherche. La reconstruction 3D touche plusieurs axes de recherche tels que : traitement d'image, isolement d'objets de leur environnement, étalonnage de caméra, appariement, triangulation, vision par ordinateur, réalité virtuelle ou augmentée... etc. Pour cela, un grand nombre de connaissances théoriques doivent être acquises et bien maîtrisées pour aboutir à la solution automatique de tels problèmes. Dans ce travail, nous souhaitons combiner deux méthodes bien connues. À savoir, « Correspondance de caractéristiques » et « Blocs ». L'intérêt est d'obtenir des modèles 3D d'objets réels avec une qualité visuelle très proche de la réalité dans un temps acceptable.

De plus, nous avons eu l'opportunité de reconstruire des objets culturels réels dont les modèles 3D ont été intégrés dans des applications de réalité augmentée. En fait, la réalité augmentée est une superposition temps réel de réalité et d'éléments (son, images 2D, 3D, vidéos, etc.). Elle fait souvent référence aux diverses méthodes qui permettent d'insérer de manière réaliste des objets virtuels dans une séquence d'images. Cela s'applique à la fois à la vision (la superposition d'images virtuelles et réelles) et aux perceptions proprioceptives telles que le toucher et l'ouïe. Ces applications sont diverses et incluent les jeux vidéo, l'éducation par jeu, les chasses au trésor virtuelles, le cinéma, la télévision (post-production, studios virtuels, retransmissions sportives, etc.), le tourisme, l'industrie (conception, design, maintenance, assemblage, navigation, robotique et télé-robotique, transplantation, études d'impact, etc.) ou dans le domaine médical.

## **Problématique**

La modélisation 3D fait partie des approches dites classiques de synthèse d'images. Dans ce cas, le processus de création d'une forme 3D suit deux phases principales :

- **Phase de modélisation 3D :**

Consiste à créer manuellement des modèles 3D simples ou complexes, en les dessinant directement à l'aide de logiciel de modélisation 3D. Peut faire appel à des outils plus complexes et nécessite une phase d'apprentissage considérables ; et le modèle 3D peut également être créé à partir de formes primitives simples telles que des sphères, des polygones, des cônes et des courbes.

- **Phase de rendu :**

La phase de modélisation est suivie par une phase de rendu qui est un processus créatif qui ressemble à la photogrammétrie<sup>1</sup> ou la cinématographie<sup>2</sup>. Les objets obtenus par modélisation sont recouverts par la suite avec des couleurs ou des textures (Séverinie, 2008). Il est même possible d'ajouter des sources lumineuses, d'ombres, de l'atmosphère, ou des caméras virtuelles pour que les résultats semblent plus proches de la réalité. On peut aussi appliquer des effets optiques tels que la réfraction de la lumière ou du mouvement flou sur les objets en mouvement.

Cette phase implique beaucoup de calculs et prend beaucoup de temps, même sur des ordinateurs puissants, puisqu'on applique un effet particulier pour chaque pixel de l'objet.

La modélisation 3D donne des résultats très satisfaisants pour des objets synthétiques, mais est moins robuste pour des scènes réelles avec beaucoup de détails à modéliser (les forêts et les villes par exemple). Aujourd'hui, les ordinateurs sont de plus en plus puissants, mais le calcul d'une seule image peut prendre beaucoup de temps et c'est pour cette raison que ces méthodes sont considérées gourmandes en temps de calcul.

Dans tous les cas, ce genre d'approches nécessite une phase d'apprentissage longue et fastidieuse et doit être réalisée par des experts et spécialistes qui maîtrisent fortement les logiciels de modélisation et fait intervenir plusieurs artistes dans ce domaine. De plus, les durées y afférentes sont trop volumineuses, sans parler de la qualité des objets obtenus qui sont synthétiques et clairement artificiels.

Les problèmes associés à la modélisation 3D ont limité son utilisation et, alternativement, des recherches peuvent être menées sur d'autres types de modélisation 3D tels que les méthodes IBMR. En particulier, on peut citer la reconstruction 3D à partir d'images qui englobe un vaste éventail de théories et de techniques. Malgré cette avancée technologique, on a constaté que ces approches utilisent une grande masse d'informations (dont une partie n'est pas nécessaire) ; ce qui a abouti à l'augmentation des temps de calculs et l'apparition d'artefacts dans les images obtenues (nécessitant eux même des temps de calculs considérables pour leurs élimination). Cette constatation nous a permis de poser la problématique suivante : « Comment peut-on réduire les temps de calculs de telles approches tout en préservant une qualité visuelle acceptable ? »

---

<sup>1</sup> **Photogrammétrie** : Technique qui vise à effectuer des mesures dans une scène, en utilisant la parallaxe obtenue entre des images acquises selon des points de vue différents.

<sup>2</sup> **Cinématographie** : Ensemble des méthodes mis en œuvre pour obtenir la reproduction du mouvement par l'intermédiaire d'un film.

En réponse à cette problématique, nous nous sommes focalisés sur la réduction du temps de calculs des modèles 3D.

La valorisation patrimoniale fait appel à plusieurs techniques et surtout à aux représentations numériques. Les méthodes utilisées visent à renforcer l'idée « de l'application potentielle des différentes technologies pour la préservation, la promotion et la mise en valeur des objets à caractère culturel ». La réalité augmentée (RA) est un domaine de recherche très dynamique, et il est nécessaire de souligner l'état de la recherche et de fournir les dernières tendances qui permettront le développement d'études futures. Les outils de la réalité augmentée facilitent l'accès au patrimoine culturel de manière interactive et engageante (Angeline & Richard, 2005). Le présent document vise à mener une étude concluante sur les efforts déployés au cours des dix dernières années pour le développement des applications de la réalité augmentée utilisées pour la préservation et la promotion de l'héritage culturel (HC). Nous analysons comment la Réalité augmentée est impliquée dans les applications du patrimoine culturel en menant des recherches bibliométriques sur les documents publiés sur les applications de la réalité augmentée en HC au cours des dix dernières années. Nous avons, à cet occasion, pu utiliser des modèles 3D dans le cadre d'applications de la réalité augmentée. Nous examinons la documentation pertinente afin d'examiner les questions de recherche suivantes :

Quelles sont les tendances dans l'utilisation de la réalité augmentée pour les applications du patrimoine culturel ? Quels sont les principaux résultats de la recherche ?

## **Motivation**

Il a été constaté que le calcul de la carte de disparité implique l'utilisation de l'image entière pour un large éventail de méthodes stéréoscopiques, ce qui conduit à un calcul de temps supplémentaire et à des artefacts visuels qui apparaissent dans les résultats. Sur la base de cette observation, nous avons dérivé une approche qui réduit considérablement le temps de calcul des algorithmes stéréoscopiques et évite les bruits pouvant apparaître dans le résultat. Dans ce travail, nous proposons une approche de reconstruction 3D basée sur le principe de l'hybridation de deux méthodes. Appliquée à des vidéos capturant des objets en rotation, nous espérons que l'approche proposée contribuera à la réduction des temps nécessaires pour la reconstruction d'objets culturels en vue de leur valorisation et préservation.

La réalité augmentée est une technologie mature qui utilise le monde réel comme support et l'étend en superposant des informations générées par ordinateur. Elle a été appliquée à plusieurs domaines. En particulier, la technologie s'est avérée très utile pour la gestion et la préservation du patrimoine culturel. Huit sujets très intéressants de l'application de la technologie de la réalité augmentée dans le contexte du patrimoine culturel peuvent être mentionnés : la reconstruction 3D d'artefacts culturels, patrimoine numérique, musées virtuels, expérience des utilisateurs, éducation, tourisme, patrimoine culturel immatériel et jeux. Notre but est de donner un aperçu des applications existantes et des tendances de recherche pour chaque domaine d'application.

## **Organisation de la thèse**

Dans ce travail nous nous intéressons particulièrement à la reconstruction 3D d'objets culturels, tout en proposant une méthode qui permet la réduction des temps de calculs y afférant.

Dans la suite de ce manuscrit, nous tenterons de présenter ce domaine de recherche. La description du contenu de chaque chapitre est comme suite :

Dans le chapitre I, nous présentons divers concepts de base nécessaires pour aborder notre travail. La première partie du chapitre II couvre les méthodes les plus courantes de modélisation 3D à base d'images. Nous présentons dans sa deuxième partie notre classification multicritère des différentes méthodes de reconstruction 3D.

Le chapitre III décrit les concepts préliminaires liés aux différentes approches utilisées dans notre recherche, la stéréovision, ainsi que certaines techniques existantes pour mettre en œuvre l'approche que nous proposons.

Un ensemble de travaux connexes sont présentés dans le chapitre IV. Le chapitre V, quant à lui, est consacré à la présentation des notions liées à la réalité augmentée, où nous présentons aussi quelques travaux similaires.

Dans le chapitre VI, nous expliquons l'approche de reconstruction 3D que nous proposons et nous présentons les résultats expérimentaux obtenus.

Nous terminons ce travail par une conclusion et un ensemble de perspectives pour nos futures recherches.

# Chapitre I

---

## *Notions de base*

---

## Introduction

La vision par ordinateur sert à munir une machine de capacités pour réaliser des tâches semblables à la vision humaine. Le processus de la vision se base sur deux phases importantes, la première étant la phase d'acquisition des images, qui sont généralement prises par des caméras et représentées par une matrice à deux dimensions où chaque élément de cette matrice est appelé pixel<sup>3</sup>. La deuxième phase est le traitement de ces images.

Dans notre travail, nous nous intéressons à un traitement d'images de haut niveau dont le but est de récupérer la troisième dimension perdue au cours de la capture des images. Il s'agit de la reconstruction 3D qui est un domaine très actif où de nombreuses méthodes sont proposées, et dont chacune est dédiée à un type particulier d'objet à reconstruire. Dans cette thèse, nous étudions la reconstruction 3D à partir d'images.

Ce chapitre est consacré pour la présentation de quelques notions de base liées à ce domaine de recherche. Nous allons voir comment on peut passer d'une scène réelle à une scène représentée par une image à l'aide d'une projection issue de caméras. Nous allons présenter le modèle de caméra le plus connu ainsi que la formation d'images.

## 1. Systèmes de coordonnées

Il existe plusieurs systèmes de coordonnées qui permettent de représenter la position de la caméra ainsi que celle de l'objet, cette position est représentée à l'aide d'un référentiel. Trois types de référentiels différents existent (Beury, 2014) :

Référentiel « Monde » lié à la scène  $R_S (O_S, X, Y, Z)$ .

Référentiel « Caméra »  $R_C (O_C, x, y, z)$ .

Référentiel « Image »  $R_I (O_i, u, v)$ .

Dans le cas le plus général, la position de la caméra et celle de l'objet sont représentées dans le système  $R_C$ .

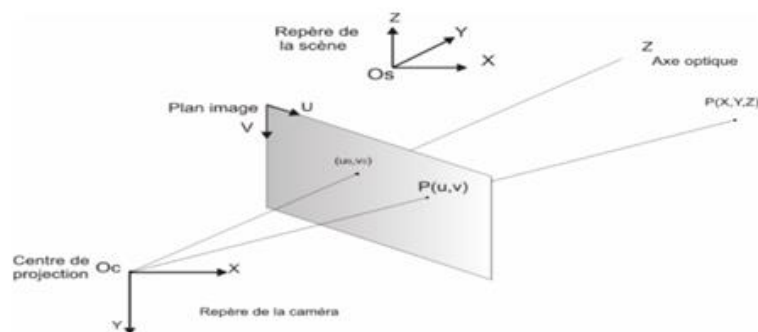


Figure 1 : Système de coordonnées.

## Coordonnées homogènes

Les systèmes de coordonnées homogènes (Crowley, 2010) sont utilisés pour faire une transformation entre deux référentiels différents. Le principe de base est d'utiliser un vecteur de  $n+1$  coordonnées dans un espace de représentation de dimension  $n$  ;  $n$  coordonnées plus une coordonnée

<sup>3</sup> **Pixel** : « élément d'image », dont le nom provient de la locution anglaise « picture element ».

supplémentaire initialisée à  $I$  utilisée pour représenter un vecteur de position ou bien initialisée à  $0$  pour représenter un vecteur de déplacement invariant pour la translation (Le zéro désigne le point à l'infini).

Les coordonnées homogènes sont représentées par :  $[x, y, z, h]$ .

Tels que :  $x, y, z$  représentent la position d'un point dans un espace 3D et  $h$  est la direction du point vers l'infini.

Le passage des coordonnées usuelles vers les coordonnées homogènes est réalisé par :

$$[x, y, z, h] \rightarrow [x/h, y/h, z/h] \text{ Si } h \neq 0.$$

$$(x, y, z) \rightarrow [x, y, z, I]$$

## 2. Transformations usuelles

Les transformations usuelles sont généralement des translations, des rotations et des changements d'échelle, elles sont représentées sous forme de matrices.

### 2.1 Translation

La translation est une transformation géométrique qui permet de glisser et de décaler l'objet sans changer son angle ou ses dimensions. Cette transformation est obtenue par la matrice suivante :

$$T(d_x, d_y, d_z) = (1 \ 0 \ 0 \ d_x \ 0 \ 1 \ 0 \ d_y \ 0 \ 0 \ 1 \ d_z \ 0 \ 0 \ 0 \ 1)$$

Tel que :  $d_x, d_y$  et  $d_z$  sont les coordonnées en axe  $x, y$  et  $z$  respectivement.

### 2.2 Changement d'échelle

C'est une transformation linéaire qui augmente ou réduit le volume d'un objet par un facteur. Le résultat d'une mise à l'échelle est similaire à la forme originale (Dans le sens géométrique). Cette transformation est obtenue par la matrice suivante :

$h_x, h_y$  et  $h_z$  sont les facteurs de mouvement selon les axes  $x, y, z$  respectivement.

$$H(h_x, h_y, h_z) = (h_x \ 0 \ 0 \ 0 \ 0 \ h_x \ 0 \ 0 \ 0 \ 0 \ 0 \ h_z \ 0 \ 0 \ 0 \ 1)$$

### 2.3 Rotation

La rotation permet de tourner un objet dans l'espace sur l'un des axes. Les matrices suivantes représentent respectivement la formule de rotation sur les axes  $x, y$  et  $z$  avec un angle  $\theta$ :

$$R_x(\theta) = (1 \ 0 \ 0 \ \cos(\theta) \ 0 \ 0 \ -\sin(\theta) \ 0 \ 0 \ \sin(\theta) \ 0 \ 0 \ \cos(\theta) \ 0 \ 0 \ 1)$$

$$R_y(\theta) = (\cos(\theta) \ 0 \ 0 \ 1 \ \sin(\theta) \ 0 \ 0 \ 0 \ -\sin(\theta) \ 0 \ 0 \ 0 \ \cos(\theta) \ 0 \ 0 \ 1)$$

$$R_z(\theta) = (\cos(\theta) \ -\sin(\theta) \ \sin(\theta) \ \cos(\theta) \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1)$$

### 3. Modélisation de la caméra

Le modèle de la caméra constitue plusieurs transformations (Figure 2) : La transformation entre le repère monde et le repère caméra, la projection du repère caméra vers le repère rétine et la transformation entre le repère rétine et le repère image (Brandou, 2008).

#### 3.1 Transformation scène-Caméra

Le processus de transformation du repère monde et le repère de la caméra peut être décrit d'une manière très simple si on utilise les coordonnées homogènes. Dans ce cas, les translations, les rotations et les changements d'échelle sont représentés par de simples produits matriciels.

Une transformation pour un point  $M : P_M(x_M, y_M, z_M, 1)^T$  dans l'espace est décrite par :

$$P_C = T_M^C \cdot P_M$$

Tel que le point  $P_C$  de coordonnées  $(x_c, y_c, z_c, 1)^T$  est exprimé dans le repère de la caméra.

La matrice  $T_M^C$  de transformation entre les deux repères a la forme suivante :

$$T_M^C = (R_M^C \ t_M^C \ 0 \ 1)$$

#### 3.2 Projection caméra-Rétine

La projection de la caméra vers la rétine est une projection perspective<sup>42</sup>. Les algorithmes de vision s'appuient généralement sur le modèle sténopé qui est connu aussi par « modèle en trou », ce modèle transforme un point 3D de l'espace  $M$  en un point image  $x$  et peut se décomposer en trois transformations élémentaires successives (Voir figure 2) :

- 1) Transformation entre repère monde et celui de la caméra.
- 2) Transformation entre repère de la caméra et le repère (Plan rétinien).
- 3) Transformation entre repère capteur et le repère d'image.

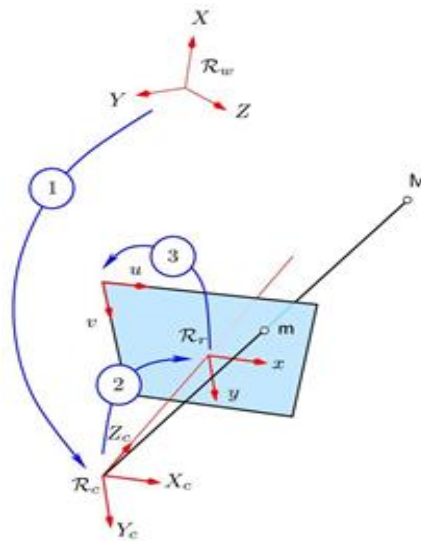


Figure 2 : Les trois transformations élémentaires du modèle sténopé, et les repères associés.

<sup>42</sup> **Projection perspective** : Transformation géométrique réalisée par un système optique qui forme l'image d'objets situés dans l'espace. Cette transformation est utilisée pour représenter des objets sur une image.



### 3.3 Transformation Rétine-image

L'image est une représentation bidimensionnelle  $(u, v)$  qui est composée d'un ensemble de pixels. La transformation du repère caméra au repère image s'écrit par :

$$(u \ v \ w) = (k_u \ 0 \ 0 \ 0 \ k_v \ 0 \ 0 \ 0 \ 0) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} (x' \ y' \ z') + (u_0 \ v_0 \ w_0)$$

Tels que :  $u_0$ ,  $v_0$  et  $w_0$  sont les coordonnées de F (centre de projection) et  $k_u$  est le facteur d'échelle vertical et  $k_v$  est le facteur d'échelle horizontal.

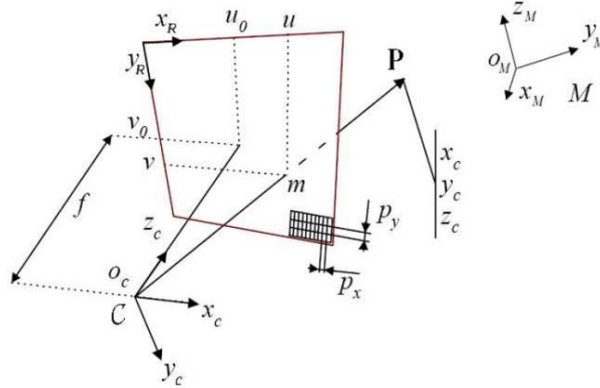


Figure 3 : Modèle de la caméra.

## 4. Formation de l'image

Une image obtenue avec une caméra est le résultat d'une transformation géométrique, cette transformation fait passer d'une représentation tridimensionnelle à une représentation bidimensionnelle.

La caméra est associée à un repère appelé centre projectif  $O_c$ ,  $(O_c, x)$  où  $(O_c, y)$  sont les axes parallèles au plan image et  $(O_c, z)$  est appelé l'axe optique et est orienté vers la scène, il est perpendiculaire au plan image.

La formation de l'image passe par trois phases (Boufama, 1994) :

- 1) Un déplacement tridimensionnel : Pour passer du repère de la caméra, il faut effectuer un changement aux points tridimensionnelles exprimés dans un repère de la scène, ce changement comporte des paramètres intrinsèques<sup>5</sup> (6 paramètres : 3 pour la rotation et 3 pour la translation).
- 2) Projection 3D-2D : Après le changement de repère, les points tridimensionnels sont projetés sur le plan image, dans ce cas de nouvelles coordonnées sont obtenues appelées coordonnées normalisées.
- 3) Changement de coordonnées : A cette étape, il faut transformer les coordonnées normalisées pour passer aux coordonnées pixel, cette transformation est appelée transformation affine<sup>6</sup> du plan et comporte 5 paramètres appelés paramètres intrinsèques<sup>7</sup> de la caméra. La formule de cette transformation peut être écrite par :

<sup>5</sup> **Paramètres intrinsèques** : Ce sont les paramètres liés à la caméra (Plus de détails dans le chapitre III).

<sup>6</sup> **Transformation affine** : Utilisée pour placer un objet dans l'espace ou bien pour faire des changements entre les repères.

$$A = (\alpha_u - \alpha_u \cos \theta \ u_0 \ 0 \ \alpha_v \ sin \ \theta \ v_0 \ 0 \ 0 \ 1)$$

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  $\frac{\pi}{2}$  la forme de la matrice de transformation devient :

$$A = (\alpha_u \ 0 \ u_0 \ 0 \ \alpha_v \ v_0 \ 0 \ 0 \ 1)$$

## 5. Relations géométriques entre les images.

L'ensemble des images prises par les caméras ne sont pas sans relation. La relation de transformation de la scène en une image est appelée homographie de plans, il est montré que deux images d'une scène plane sont liées par une homographie de plans (Frédéric, Noirfalise, & Sturm, 2011).

Considérons deux images (généralement une image gauche et une image droite) d'une même scène. Dans ce cas un point  $p$  d'une image gauche représente la projection de tous les points de l'espace appartenant à une droite  $OP$  (Voir la figure 4), donc les correspondants de  $P$  dans l'image de droite sont tous trouvés sur cette droite. Cette géométrie est appelée géométrie épipolaire.

### Géométrie épipolaire.

La géométrie épipolaire est un modèle mathématique qui vise à trouver une relation entre un ensemble d'images d'une même scène prises sous différents angles de vues, le principe général est de trouver pour chaque pixel d'une image son correspondant dans l'autre image, la relation obtenue par cette géométrie permet de réduire la zone de recherche des points correspondants qui seront utilisés pour la reconstruction 3D de la scène.

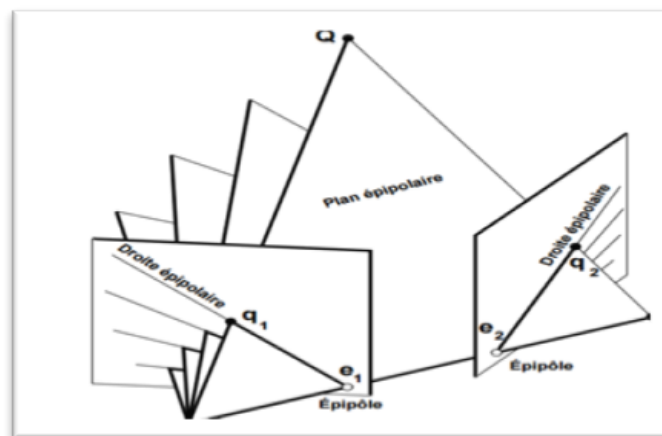


Figure 4 : Géométrie épipolaire.

Considérons maintenant trois images d'une même scène, ces images sont traitées deux à deux, donc on dispose de trois paires d'images. On considère aussi trois matrices fondamentales<sup>8</sup>  $F_{11}$ ,  $F_{13}$  et  $F_{23}$  (Voir figure 5). Il est supposé qu'un point  $P$  de la scène qui se projette en  $p$ ,  $p'$  et  $p''$  dans les images 1, 2, 3 respectivement. Une géométrie épipolaire lie la première image avec la troisième image, on déduit donc que  $p''$  se trouve sur la droite  $(F_{13}p)$ . De la même façon si on lie la deuxième

<sup>8</sup> **Matrices fondamentales** : C'est une matrice qui décrit la relation entre les coordonnées images et les coordonnées caméra.

image avec la troisième, dans ce cas  $p''$  se trouve sur la droite  $(F_{23}p')$ . Ce dernier est donc trouvé par l'intersection de ces deux droites épipolaires :

$$p'' = (F_{13}p) \wedge (F_{23}p')$$

Il est déduit donc que les points  $p$ ,  $p'$  et  $p''$  sont reliés par une relation trilineaire qui dépend des matrices fondamentales. Cette relation est très utile pour la poursuite des points dans une séquence d'images, elle a un grand intérêt aussi dans la stéréovision trinoculaire.

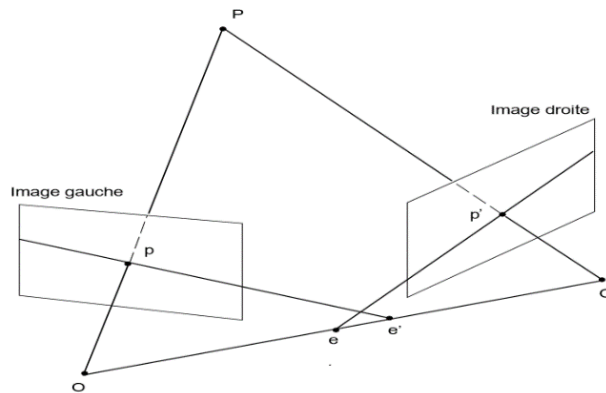


Figure 5: Relation géométrique entre deux images : Géométrie épipolaire.

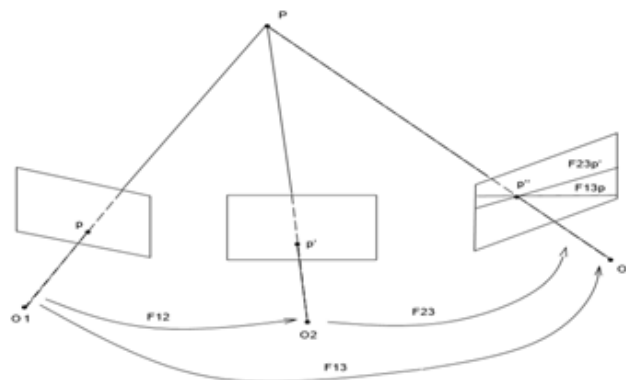


Figure 6 : Relation géométrique entre trois images : Trilinearité.

## 6. Relation entre vidéo et images

La vidéo est une succession rapide d'images qui sont mises en œuvre à une certaine cadence. Pour obtenir une vidéo ; une caméra est utilisée pour prendre un ensemble d'images dans une unité de temps (seconde).

Le nombre d'images prises par seconde représente une unité de mesure qui correspond au nombre d'images affichées par un dispositif, plus ce nombre est élevé plus l'animation semble fluide. Cette mesure est notée par **FPS** (*Frames per second*) c'est-à-dire le nombre des images par seconde en français.

Pour le codage d'une vidéo, le terme **GOP** (Group of image) est utilisé, ce dernier signifie l'ordre dans lequel les images internes sont disposées. Les images qui appartiennent aux **GOP** permettent de générer l'ensemble d'images visibles dans la vidéo.

La reconstruction 3D à partir d'images peut être utilisée aussi dans le cas de vidéo. Le principe est très simple, c'est l'opération réciproque de prendre une vidéo : c'est-à-dire il suffit de prendre cette vidéo et d'y extraire l'ensemble d'images qui la compose. La méthode choisie pour la reconstruction sera appliquée sur la séquence d'images obtenues.

## **Conclusion**

Ce premier chapitre a introduit quelques concepts de base dans le but d'introduire le domaine de recherche et d'expliqué certains aspects utilisés dans le reste de ce travail pour lever les ambiguïtés. Parmi ces termes, nous avons introduit les caméras et leur modélisation. Cela a permis de décrire les objets photographiés dans l'espace et projetés sur l'image, et en particulier le processus de formation de l'image.

Nous avons également montré la relation entre les images capturées par la caméra. Cette relation sera utilisée plus tard lors du traitement de ces images. Par exemple, extraire des informations utilisées pour calculer les trois dimensions de la vue stéréoscopique. L'objectif de notre travail est l'application de méthodes basées sur la reconstruction 3D et les notions y afférant seront utilisées tout au long de notre travail.

Après avoir décrit les principales caractéristiques de la reconstruction 3D, le chapitre suivant présente une classification générale des différentes méthodes de reconstruction 3D.

# Chapitre II

---

## *Préliminaires*

---

## **Introduction**

La modélisation 3D classique consiste à créer des objets 3D en commençant par un ensemble de primitives de base (cubes, cylindres, courbes ...), elle nécessite des utilisateurs experts et un temps de traitement assez important (Pouvant être chiffré en Homme-Année<sup>9</sup>). Cette modélisation n'est pas utile pour la réalisation de scènes complexes, malgré le grand progrès dans ce domaine.

Les résultats finaux sont, dans la plupart des cas, artificiels et les objets obtenus sont synthétiques. Les problèmes détectés dans la modélisation 3D classique ne résident pas seulement dans la complexité du travail, on trouve aussi que le matériel nécessaire pour la modélisation est très coûteux : Les logiciels de modélisation 3D sont assez cher et nécessitent des machines très puissantes.

Contrairement à la synthèse d'images classique, les méthodes IBMR (*Image Based Modeling and Rendering*) permettent de reconstruire la forme 3D à partir d'images d'une scène réelle. Ces approches ont attiré beaucoup d'attention dans la communauté graphique au cours de ces dernières années. Le principe général des méthodes IBMR est d'utiliser l'information présente dans les photographies pour la reconstruction 3D.

Dans ce chapitre, nous allons présenter en première partie plus de détails sur ces approches : Le fonctionnement, les méthodes les plus connues dans la littérature ...etc. Dans la deuxième partie nous allons présenter une classification des méthodes de reconstruction 3D qui regroupe toutes les méthodes existantes.

## **1. Méthodes de modélisation et rendu à base d'image : IBMR**

Il est connu que l'image est une représentation bidimensionnelle d'un monde tridimensionnel, la profondeur est perdue au cours de la projection de ces images. Le but des méthodes de modélisation et de rendu à partir d'images (*IBMR*) est de récupérer cette dimension perdue, en se basant sur les informations photométriques (couleur et intensité de chaque pixel) et géométriques (paramètres des caméras et disparité de chaque pixel) disponibles dans les images.

Certains auteurs font la séparation entre les méthodes de modélisation et les méthodes de rendu, dans ce cas les termes *IBR* et *IBM* sont utilisés respectivement pour (*Image Based Rendering* et *Image Based Modeling*). La technique *IBM* consiste à générer le modèle 3D à partir d'un ensemble de photographies d'une scène (Voir figure 1), par contre le rendu basé sur l'image (*IBR*) utilise des images pour colorer ou texturer un modèle 3D (Voir figure 2).

Les techniques d'IBMR se différencient selon la façon dont les informations disponibles dans les images sont utilisées. Certaines utilisent, interpolant<sup>10</sup>, et déforment l'ensemble d'images d'une façon directe (Zhang & Chen, 2004). D'autres méthodes utilisent la géométrie de l'image pour générer des vues cohérentes de la scène, le résultat de cette méthode est une reconstruction partielle (Blanc, 1998). Il existe aussi des méthodes hybrides qui utilisent les vues de la scène réelle combinées avec une modélisation classique pour améliorer les résultats et accélérer le rendu (Debevec, Taylor, & Malik, 1996).

<sup>9</sup> **Homme-Année** : unité correspondant au travail d'une personne pendant un an.

<sup>10</sup> **Interpolation** : Opération mathématique permettant de construire une courbe à partir des données d'un nombre fini de points, ou une fonction à partir de la donnée d'un nombre fini de valeurs.



Figure 7 : Modélisation basée images. (Florian Moreno - Avril 2009 - Portal of Architectural Image-Based Modeling)



Figure 8 : Rendu basé image. (Fujitaand, Masahiro, Takashi, Kanai - hardware assisted relief texture mapping)

Les méthodes *IBMR* deviennent très utilisées en raison de leurs avantages par rapport aux autres techniques. Parmi ces avantages (Neto, 2002):

- Les caméras et les webcams ne sont pas coûteuses. La possibilité de capturer des effets subtils des modèles 3D à partir de caméras ouvrirait la possibilité aux consommateurs de passer de la 2D vers la 3D.
- Les méthodes *IMBR* permettent d'éviter de nombreuses étapes difficiles suivies d'autres méthodes de modélisation 3D.
- Les performances des cartes graphiques ont augmenté de façon spectaculaire au cours de la dernière décennie. Néanmoins, le désir de réalisme et de grand détail limite souvent le taux de rendu de l'image au-dessus du temps réel. En appliquant des techniques *IBMR*, des images de certaines vues canoniques peuvent être pré-calculées, puis déformées en temps réel et interpolées dans la séquence finale souhaitée de la vitesse de la vidéo.

## 2. Quelques approches d'IBMR

De multiples méthodes de reconstruction 3D basées purement images sont existantes ; dans cette section nous nous intéressons aux approches les plus utilisées.

### 2.1 Forme à partir X (*Shape from X*)

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

L'approche shape from shading (Zhang, Tsai, Cryer, & Shah, 1999) 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 sous la forme d'une UDP voir (Berthold, 1989). Cette UDP appelée « L'équation d'irradiance », dont le principe est d'exploiter un ombre pour trouver le relief d'une scène à partir d'une seule photographie, puis d'estimer la forme d'un objet éclairé à partir des variations d'éclairage observé.

Pour cela, il est nécessaire d'avoir une image en niveau de gris et de connaître la direction exacte de la source lumineuse, dans ce cas il devient possible de déterminer la position 3D.

- Si la surface de la forme est décrite par l'information de la normale, cela revient à résoudre pour chaque pixel une équation linéaire à trois inconnus.
- Et si la surface est décrite en termes de gradient<sup>11</sup>, cela revient à résoudre une équation linéaire à deux inconnus.

Cette approche donne des modèles 3D robustes, mais un problème majeur qui se pose est que les algorithmes shape from shading donnent des résultats dans des temps loin du temps réel. La figure suivante représente le résultat d'une telle approche (Meyer, Briceño, & Bouakaz, 2007) :

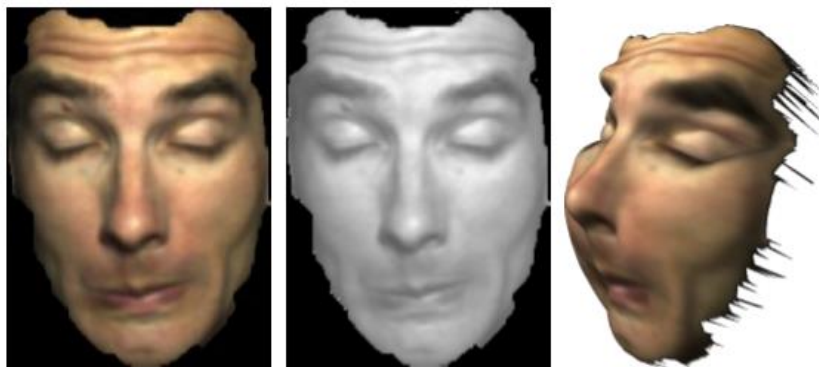


Figure 9 : Résultat de shape from shading. A gauche l'image source, au centre l'image en niveau de gris et le résultat dans l'image de droite.

#### 2.1.2 Forme à partir silhouette (*Shape from silhouette*)

Cette méthode utilise les silhouettes de l'objet pour obtenir la forme 3D (Gloria, 2012). Un ensemble d'images prises par différents points de vue qui seront traitées pour obtenir le fond de la forme de l'objet.

Les silhouettes sont généralement représentées par un masque binaire, toute erreur de l'extraction des silhouettes influe directement sur la qualité de la reconstruction 3D. Ces méthodes permettent

<sup>11</sup> **Gradient** : Est un vecteur représentant la variation d'une fonction par rapport à la variation de ses différents paramètres.



d'exprimer la forme 3D, mais elles sont limitées uniquement aux objets entièrement visibles pour toutes les caméras.

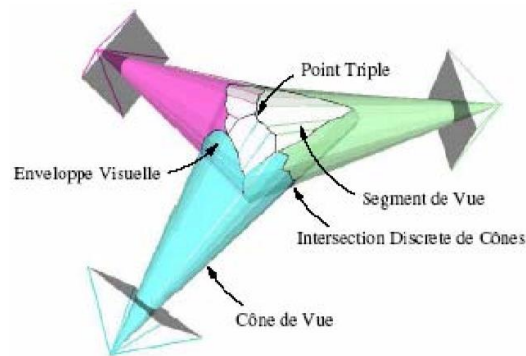


Figure 10 : Shape from silhouette à partir 3 vues différentes.

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

L'approche shape from texture (Loh & Hartley, 2005) permet d'estimer la forme 3D de l'objet en fonction de variations observées sur la texture, c'est-à-dire le changement des matériaux et des couleurs.

Une texture est une caractéristique matérielle d'une surface constituée d'un motif plus ou moins régulier (grain d'un bois, poli d'un métal ...etc.). Les variations de texture dans l'image peuvent donner une bonne indication du relief<sup>12</sup> mais seulement dans le cas où cette texture est homogène.

L'avantage de ces méthodes est la simplicité des calculs et le texturage immédiat du modèle 3D. Mais l'inconvénient réside dans la sensibilité aux changements d'éclairage et l'augmentation de la complexité en temps de calcul pour obtenir une bonne définition.

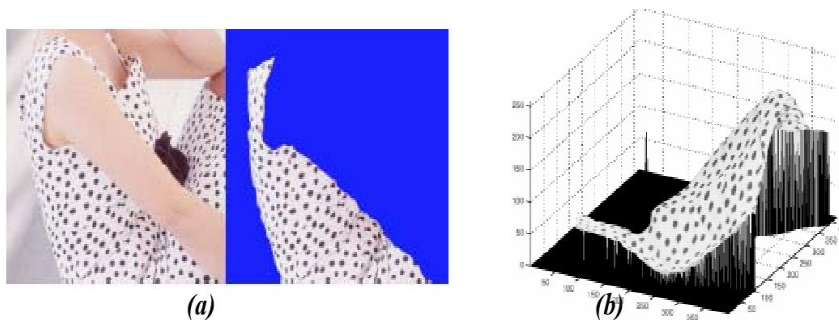


Figure 11 : (a) Texture extraite d'une robe, (b) surface reconstruite de la texture.

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

La reconstruction 3D à partir de mouvement (Bottino & Laurentini, 2001) consiste à capturer une séquence d'images d'un objet en mouvement. La forme 3D est calculée en estimant pour chaque image le mouvement visible d'un ensemble de points ou bien de tous les points de l'image, la vitesse calculée permet de retrouver la forme 3D. Ceci est effectué en supposant que le mouvement est très faible entre chaque image, et la scène est statique afin que le déplacement des points dans l'image ne soit dû qu'au mouvement de la caméra. Cette approche est caractérisée par la facilité de mettre en

<sup>12</sup> **Relief** : Désigne une image (dessin, film, etc.) donnant l'illusion d'une troisième dimension, visible en stéréoscopie.

correspondance les pixels entre les images par des méthodes de suivi visuel et ainsi d'exploiter les images. L'inconvénient réside dans la sensibilité au bruit par rapport à une méthode stéréo, dans ce cas le calcul de disparité est instable car le déplacement entre les deux images est très faible.

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

C'est une technique qui consiste à reconstruire la forme 3D à partir de plusieurs images prises par une caméra fixe, avec différentes valeurs de mise au point. Cette approche utilise le fait que l'image de la scène dépend de la radiance de la région au lieu de la forme de cette région. On trouve donc qu'il existe plusieurs approches qui permettent de résoudre ce problème en se basant sur l'hypothèse que la surface de la scène est localement plane et parallèle au plan de la caméra (Hypothèse de l'equifocal) (Asada, Fujiwara, & Matsuyama, 1998). Ces méthodes restent sensibles au manque de texture des objets à reconstruire et sont moins robustes au bruit.

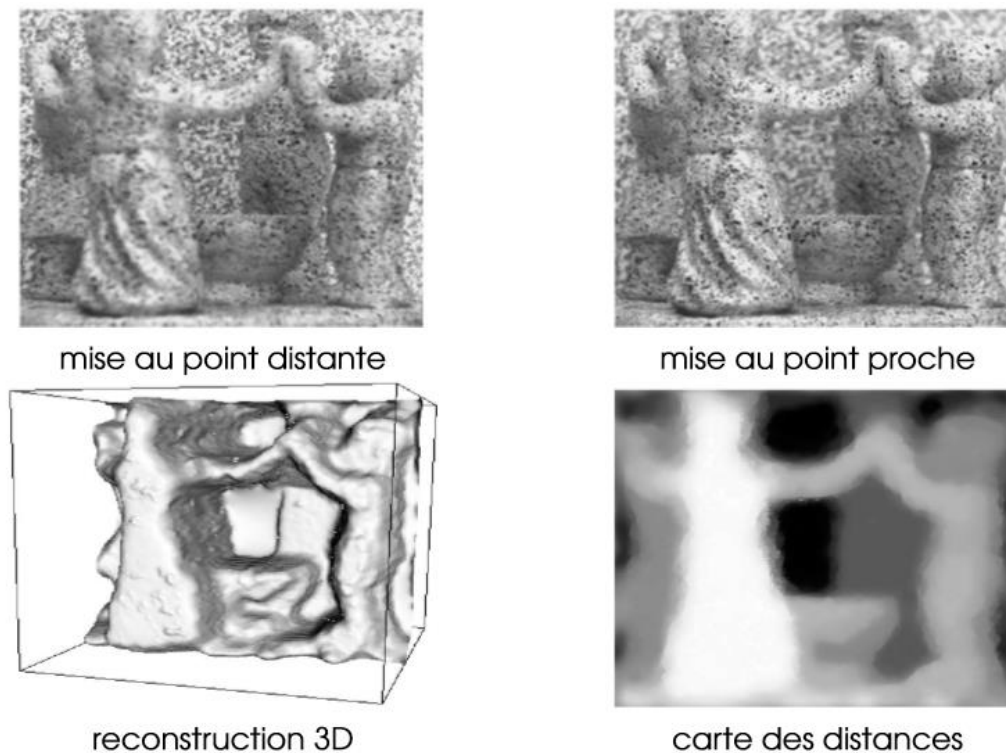


Figure 12 : Forme à partir de focus/defocus.

## 2.2 Les imposteurs

Les imposteurs (Damien, 2004) représentent une technique qui réalise un pré-rendu hors ligne d'une scène réelle, l'image créée est utilisée par la suite comme une texture plaquée sur un rectangle placé à l'intérieur de la scène (Précisément à la place de l'objet complexe de la scène). Le but de cette technique est d'obtenir une complexité visuelle à moindre coût.

Les imposteurs les plus connus sont les Panneau d'affichage (dits *billboards* en anglais) et les *sprites* (dits *Lutin*). Les *billboards* sont utilisés pour représenter des objets complexes et difficiles à modéliser, comme par exemple des arbres ou bien des nuages. Les *sprites* se différencient des *billboards*, un *sprite* est en mouvement et est appliqué directement à l'écran, il est généralement petit. On distingue deux types différents d'imposteurs : Les imposteurs statiques et les imposteurs dynamiques.

### 2.2.1 Imposteurs statiques

Les imposteurs statiques permettent une visualisation, ils visent à remplacer les objets complexes de la scène par des modèles simples est texturés dont le but est d'obtenir un taux d'affichage élevé et constant.

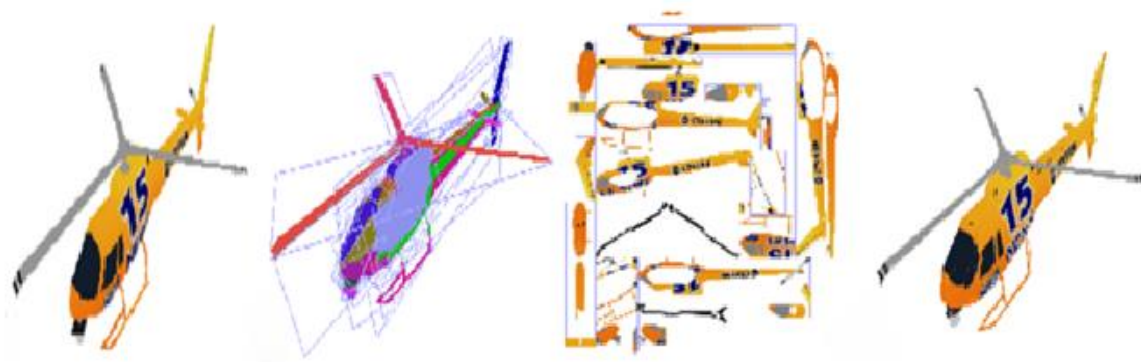


Figure 13 : De gauche en droite Modèle original, ensemble optimal de rectangles, plaquage de texture, Panneau d'affichage rendu.

### 2.2.2 Imposteurs dynamiques

Contrairement aux imposteurs statiques, les imposteurs dynamiques sont pré-générés. Dans cette méthode, l'imposteur est représenté par une image qui contient un polygone transparent où l'objet sera plaqué (Voir figure 14).

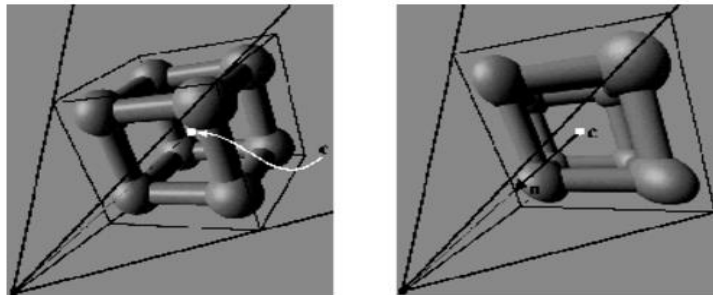


Figure 14 : Imposteur. A gauche, l'objet original, à droite, l'imposteur.

### **3. Classification multicritères des méthodes de reconstruction 3D**

On trouve qu'il y a de multiples méthodes de reconstruction 3D, mais elles sont souvent limitées à un domaine spécifique d'application. Dans cette section nous allons présenter une classification multicritère que nous avons proposée dans (Djaber & Chaouki, 2011) et qui regroupe les différentes approches existantes. Cette classification permet d'aider les personnes au choix de la méthode adéquate avec la scène étudiée ou à la qualité envisagée. Pour éviter toute confusion possible, le tableau suivant résume les différentes méthodes présentées dans cette classification selon des critères bien définis :

**Tableau 1: Critères de classification des méthodes de reconstruction 3D.**

<b>Critère de classification</b>	<b>Méthodes existants</b>
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

### 3.1 Niveau d'automatité

#### 3.1.1 Reconstruction automatique

Le but de ces approches est de reconstruire une scène tridimensionnelle purement avec machine (Alexandre, 2014). Le seul travail de l'utilisateur dans ce cas est de préparer la scène à reconstruire et les matériels à utiliser pour l'acquisition... etc.

Ce type d'approche est dédié pour la reconstruction d'objets simples. On trouve que la reconstruction purement automatique n'est pas encore efficace pour modéliser des scènes grandes et complexes, elle est aussi délicate pour le cas d'images éloignées. Par contre la reconstruction complètement automatique de scènes à partir des images proches est plus réussie, il devient possible d'avoir des formes 3D exacte.



Figure 15 : Exemple d'une voiture reconstruite automatiquement.

#### 3.1.2 Reconstruction semi-automatique

Ces méthodes nécessitent des interventions spécifiques des utilisateurs. Le rôle de l'utilisateur est de réaliser l'ensemble des tâches difficiles à automatiser, il peut par exemple désigner les points ou bien les droites sur une ou plusieurs images (Voir figure 16) pour approximer le modèle 3D à reconstruire, il peut même avoir des modèles génériques disponibles à augmenter jusqu'à obtenir une bonne estimation du modèle 3D. Le reste du travail sera effectué par la suite par le logiciel.



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

#### 3.1.3 Reconstruction manuelle

Cette technique est presque similaire à la reconstruction 3D classique, la différence est que dans ce cas les objets à modéliser sont des objets réels au contraire de la reconstruction classique (objets synthétisés). Ce genre de reconstruction est utilisé dans le cas d'indisponibilité des informations pertinentes à la reconstruction envisagée.

## 3.2 Niveau de complexité de la scène

### 3.2.1 Grandes scènes

La reconstruction 3D de scènes larges (comme par exemple la reconstruction des villes) nécessitent des techniques sophistiquées, elles ont des algorithmes spécifiques pour le plaquage de textures et pour la détection de formes, ces approches font intervenir les satellites, les données GPS, les avions ...etc.



Figure 17 : Exemple de reconstruction 3D d'une ville.

### 3.2.2 Scènes moyenne

La reconstruction de scènes d'ordres moyens (Des immeubles ou bien des bâtiments par exemple), utilise des techniques qui permettent l'acquisition des données nécessaires pour le processus de la reconstruction. Dans ce genre d'approches il est impossible par exemple de faire tourner l'objet considéré et de le photographier par une caméra fixe.

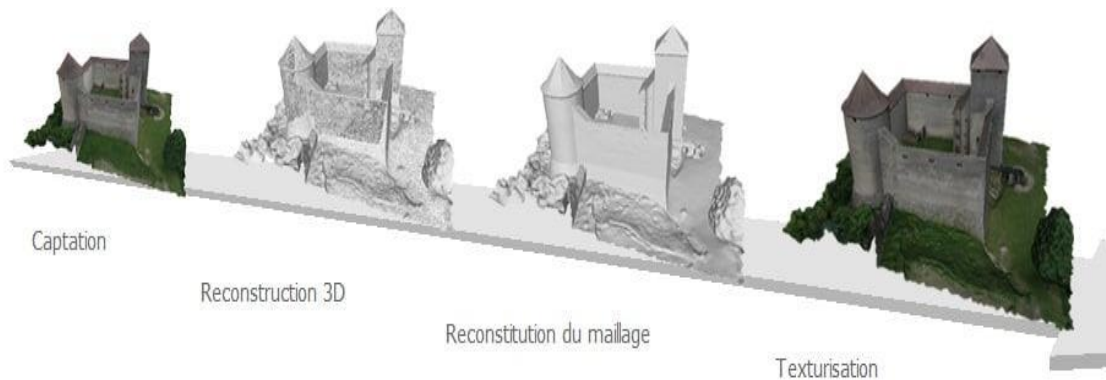


Figure 18 : Modèle 3D d'un bâtiment service publiques.

### 3.2.3 Petites scènes

Dans ce cas, l'acquisition de projections en 2D de l'objet à étudier n'est pas difficile et le temps requis pour collecter des données critiques pour la reconstruction est relativement petit.

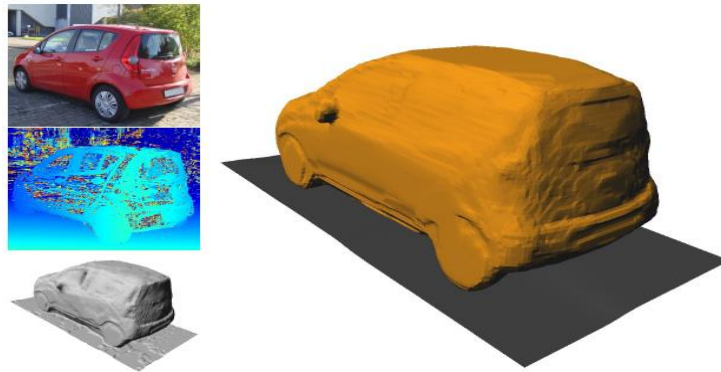


Figure 19 : Modèle de reconstruction 3D d'une petite scène.

### 3.3 Mouvement de la scène

#### 3.3.1 Scènes statiques

La reconstruction de scènes immobiles ne fait pas intervenir la notion de temps. Les études menées dans ce cadre visent essentiellement à améliorer les algorithmes ou d'obtenir des résultats plus réalistes que possible.

#### 3.3.2 Scènes dynamiques

Pour reconstruire des objets mobiles il est nécessaire d'estimer les paramètres du modèle en mouvement, dans ce cas le problème porte sur la détermination de la zone de recherche. Lorsque le modèle en mouvement est inconnu, il doit être nécessaire d'estimer le vecteur de vitesse de chaque point. Les algorithmes de ce type de reconstruction manquent de robustesse.

### 3.4 Types d'approches utilisées

#### 3.4.1 Approche ascendante

Elles sont connues par les approches (Bottom-up), dans ce cas la forme 3D est reconstruite à partir de primitives (Points 3D, courbes, segments...) sans prendre en compte d'éventuelles connaissances à priori sur les scènes que l'on cherche à reconstruire. Ces approches visent à reconstruire la surface par un maillage s'appuyant sur les primitives extraites, elles sont bien adaptées pour caractériser les structures au sein des objets et les limites saillantes de surface des objets.

Le manque d'informations qui concernent la scène génère parfois des résultats peu réalistes à cause d'une combinatoire importante, de plus l'extraction de primitives est sujette aux problèmes de sur-détection et de sous-détection qui influence le résultat final. Cette approche est utilisée par exemple pour la détection et la reconstruction des facettes 3D (Chehata, Jung, Pierrot-Deseilligny, & Stamon, 2004).

#### 3.4.2 Approche descendante

Dans les approches descendantes (Top-down), une connaissance concernant les scènes est exigée, on commence par un ensemble de modèles bien connus pour choisir celui qui est le plus cohérent avec l'objet à reconstruire. Ces approches sont limitées par la bibliothèque des modèles précédemment définis et elles sont plus robustes que les approches ascendantes.

#### 3.4.3 Approche hybride

Ces approches utilisent les approches (Bottom-up) et (Top-down) à la fois. Elles commencent par une formulation des hypothèses par une approche ascendante puis une vérification de ces

hypothèses est effectuée par une approche descendante et à la fin les hypothèses de modèles géométriques et morphologiques élémentaires sont ajustées. Cela offre une très bonne connaissance sur la scène ce qui permet de la décrire par une grammaire de formes.

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.

### 3.5 Passive ou active

#### 3.5.1 Les méthodes actives

Les méthodes actives (Levoy, et al., 2001) sont des méthodes qui utilisent des sources lumineuses avec des capteurs qui ont une action directe sur l'objet étudié. Si la position de la source lumineuse et le modèle de la caméra sont connus, il devient possible de calculer les coordonnées tridimensionnelles des points de la scène par une simple triangulation. Ces méthodes sont utilisables dans différentes applications, particulièrement dans le domaine de la robotique.

Parmi les méthodes actives, on trouve la méthode de télémétrie laser (Samain), ces méthodes utilisent en général le temps que met une impulsion à faire un aller-retour pour calculer la distance du point visé.

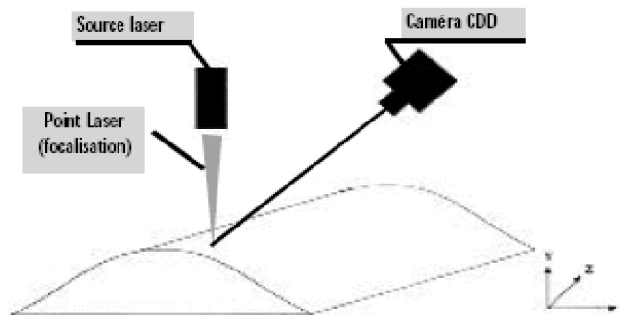


Figure 20 : Télémètre Laser.

Un autre type de méthodes actives est la reconstruction basée sur la lumière structurée (Scharstein & Szeliski, 2003). 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 cette 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 21 : Lumière structurée illuminant un visage.



### 3.5.2 Les méthodes passives

Le principe de la reconstruction 3D avec les méthodes passives (Nister, 2004) est d'imager une scène à reconstruire à partir de plusieurs points de vue différents, les capteurs utilisés dans ce cas sont généralement des caméras. Ces méthodes n'ont aucune action sur l'environnement, le terme « passive » vient du fait qu'on ne considère que la perception des radiations visibles dans la scène. La stéréovision est l'une des méthodes passives les plus connues et les plus utilisées, cette dernière se base sur l'utilisation d'au moins de deux caméras pour prendre les captures de la scène depuis des angles de vues différents. Cette méthode dépend de la calibration des caméras, de la mise en correspondance entre les différents points qui composent l'image qui seront utilisés par la suite dans une phase de triangulation qui permet de calculer les profondeurs de la scène. La phase la plus difficile dans la méthode de la stéréovision est la mise en correspondance, ce qui la rend inefficace pour reconstruire des scènes complexes.

Cette méthode fait partie intégrale de notre travail, et nous allons détailler son principe de fonctionnement dans le chapitre suivant.

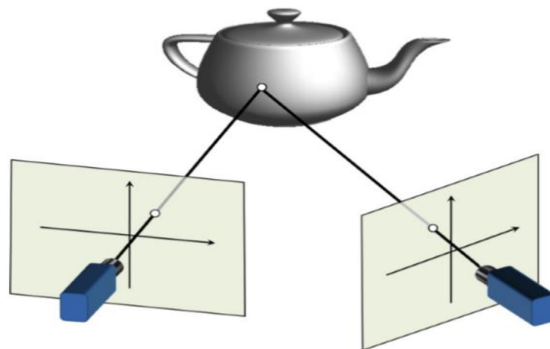


Figure 22 : Simple exemple du système de stéréovision.

## 3.6 Multiplicité des objets

### 3.6.1 Reconstruction mono-objet

Elle vise à traiter la scène comme un élément non décomposable, ceci est le cas d'un seul objet tout simplement ou bien de reconstruire la scène une fois pour toute si elle contient un ensemble d'objets indissociables. Cette méthode n'est pas fiable pour les objets complexes.

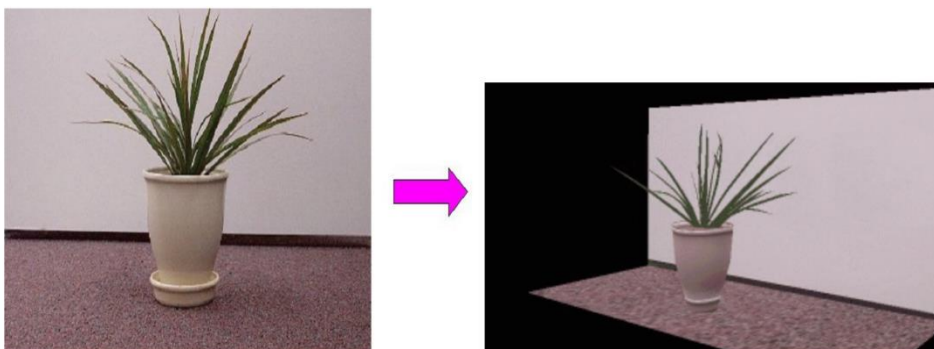


Figure 23 : Exemple de reconstruction mono-objet.

### 3.6.2 Reconstruction multi-objet

Pour aboutir à la reconstruction d'une scène complexe, cette dernière doit être décomposée en séparant les objets qui la compose et en procédant par la reconstruction de chaque objet indépendamment aux autres. Les modèles générés à la fin sont associés dans un seul modèle global.

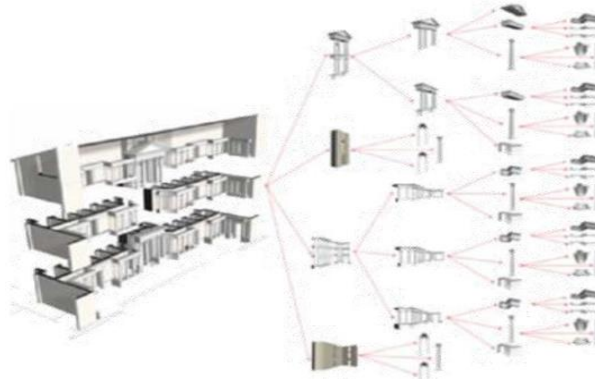


Figure 24 : Regroupement des éléments reconstruits.

## 3.7 Degré de précision

### 3.7.1 L'apparence visuelle

Ces méthodes se basent sur l'apparence visuelle des surfaces qui composent l'objet étudié. Cette apparence est estimée en étudiant les caractéristiques des matériaux et de la source lumineuse ainsi que la géométrie des objets.

### 3.7.2 Exigences de description

Cette approche consiste à regrouper un ensemble de points pertinents pour restituer les seules informations nécessaires à la description d'un aspect spécifique tels 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.

### 3.7.3 Exactitude du modèle géométrique

Ces méthodes s'appuient sur un balayage laser 3D pour réaliser la reconstruction 3D. Un problème lié à ces méthodes est que les surfaces composant l'objet entier sont difficilement calculées. La qualité des surfaces résultantes représente un autre défaut de ces approches, les modèles obtenus contiennent des artefacts de la géométrie (Des trous, une discontinuité... etc). De plus, le type de matériaux impactés par le faisceau laser détermine l'intensité du signal de retour, ce qui cause une perte d'informations, donc le maillage généré ne représente pas la forme réelle de l'objet.

### 3.7.4 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 par cet objet. Dans l'étape d'acquisition, la capture d'informations métriques et photométriques est prise en compte d'après les différents niveaux de complexité géométrique que l'objet examiné exhibe et par rapport aux contraintes distinctives du système de représentation choisi.

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

#### 3.8.1 Reconstruction projective

La reconstruction projective est réalisée à partir de plusieurs images d'une scène et des coordonnées des points correspondants dans les différentes images, ces coordonnées doivent être représentées dans un repère projectif. Cette reconstruction est réalisée sans aucune connaissance sur la structure de la scène ou sur l'orientation des caméras.

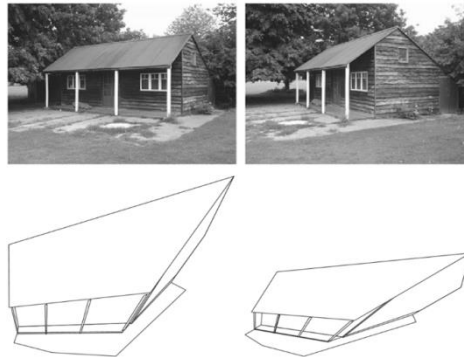


Figure 25 : Résultat d'une reconstruction projective. (Kassem Al Ismaeil - Structure from Motion & Camera Self-Calibration, 2011)

#### 3.8.2 Reconstruction euclidienne

Contrairement à la reconstruction projective, la reconstruction euclidienne nécessite une connaissance à priori de la calibration des caméras. Les paramètres de la caméra qui sont estimés à l'aide de la géométrie de la scène, permettent de donner une correspondance entre les dimensions réelles et les dimensions de la caméra. Ce type de reconstruction respecte les angles et les distances et est aussi la plus complète.

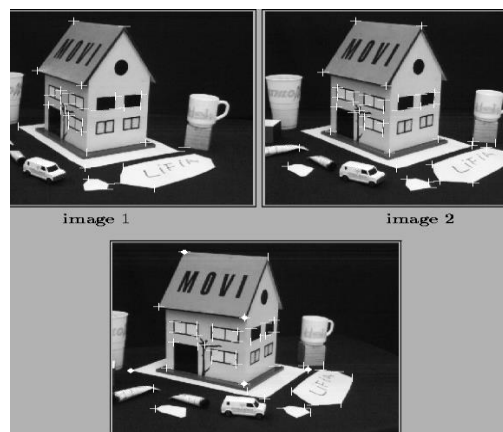


Figure 26 : Reconstruction euclidienne. (Bill Triggs, 1998)

#### 3.8.3 Reconstruction affine

Cette méthode peut être appliquée quand les points sont définis dans un repère affine. La difficulté réside dans l'estimation de 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.

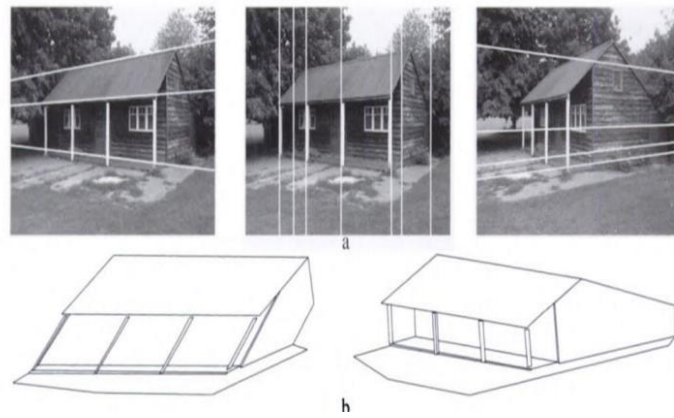


Figure 27 : Résultat d'une reconstruction affine. (Kassem Al Ismaeil - Structure from Motion & Camera Self-Calibration, 2011)

### 3.8.4 Reconstruction métrique

Cette approche est utilisée lorsque les coordonnées des points sont définies dans un repère métrique. La reconstruction est alors définie à une transformation métrique près. Une telle reconstruction contient des informations métriques (Changement d'échelle).

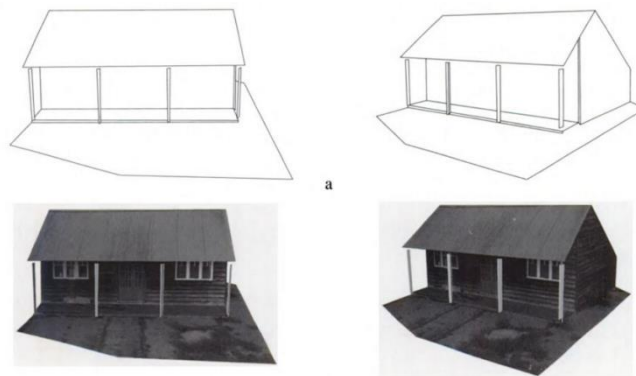


Figure 28 : Résultat d'une reconstruction métrique. (Kassem Al Ismaeil - Structure from Motion & Camera Self-Calibration, 2011)

## **Conclusion**

Dans ce chapitre nous avons passé en revue les méthodes de reconstruction 3D à partir d'images connues par le terme *IBMR*. Ce type de reconstruction utilise les informations présentes dans les photographies pour le calcul de la troisième dimension, nous avons présenté aussi comment ces approches sont assez utiles par rapport à une modélisation 3D classique.

Il existe plusieurs méthodes appartenant à la famille *IBMR* qui sont très connues dans la littérature. Parmi ces méthodes on trouve celles qui se basent sur une seule image comme le cas de la méthode shape from texture, ou bien sur plusieurs images comme le cas de shape from silhouettes. Certaines d'autres, utilisent d'autres principes totalement différents comme par exemple les imposteurs. De plus, il existe d'autres approches qui appliquent des sources lumineuses sur les images, la projection de cette source donne des informations supplémentaires qui aident au calcul de la profondeur. Pour toutes ces méthodes nous avons présenté une brève description sur le fonctionnement de chacune d'entre elles.

Nous avons aussi présenté au cours de ce chapitre une classification générale des différentes techniques de reconstruction 3D en se basant sur de multiples critères. Cette classification a pour but de simplifier le choix de la méthode à utiliser, ce qui permet d'éviter une mauvaise reconstruction 3D. Il n'est pas possible par exemple d'excepter de reconstruire une grande ville en temps réel, par contre cette méthode est utile pour la reconstruction d'un seul objet fixé.

Dans le prochain chapitre nous restons dans le cadre des méthodes *IBMR*, et nous allons détailler comment les deux méthodes enveloppe visuelle et stéréovision fonctionnent. Ces approches constituent la base de notre travail par la suite.

# Chapitre III

---

## *Enveloppe visuelle et stéréovision*

---

## Introduction

Nous avons vu au chapitre précédent que les méthodes *IBMR* utilisent les informations disponibles dans les photographies pour calculer la troisième dimension perdue au cours du processus de projection. Cette reconstruction est effectuée à base des points visibles dans l'ensemble d'images disposées. Ce problème est assez délicat pour le résoudre par une machine, et d'obtenir un résultat proche de la réalité. L'enveloppe visuelle et la stéréovision sont classées parmi les méthodes *IBMR* les plus connues.

Dans cette partie nous allons présenter en premier lieu la méthode d'enveloppe visuelle, en passant par son fonctionnement, arrivant aux approches existantes pour la réaliser ainsi que les problèmes liés à cette approche.

Par la suite nous allons exposer la méthode de stéréovision en expliquant son but et ces différentes étapes.

Ces méthodes constituent le vif de notre travail et seront utilisées par la suite dans l'implémentation de notre méthode.

### 1. L'enveloppe visuelle

L'enveloppe visuelle est une méthode assez populaire et utile pour la reconstruction 3D à cause de sa simplicité et de son efficacité. Elle est introduite en premier temps par *Baumgart* dans sa thèse de doctorat en 1979 (Baumgart & Bruce G, 1974). Le principe de base est de réaliser une représentation 3D d'un objet à partir d'un ensemble de silhouettes prises sous différents points de vue d'où vient le terme (Shape from silhouette).

L'enveloppe visuelle se définit généralement comme l'intersection de tous les cônes<sup>13</sup> de vues associées aux points de vue considérés, et constitue donc une région fermée de l'espace dont les points ne se projettent qu'à l'intérieur de tous les contours occultants. *Laurentini* (Laurentini, 1994) est considéré le premier qui a donné un sens à la notion de l'enveloppe visuelle, sa définition est comme suit :

« L'enveloppe visuelle  $VH(S, R)$  d'un objet  $S$  par rapport à une région d'observation  $R$  est une région de  $E^3$  telle que, pour chaque point  $P \in VH(S, R)$ , et chaque point de vue  $V \in R$ , la demi ligne partant de  $V$  et passant par  $P$  contient au moins un point de  $S$  ».

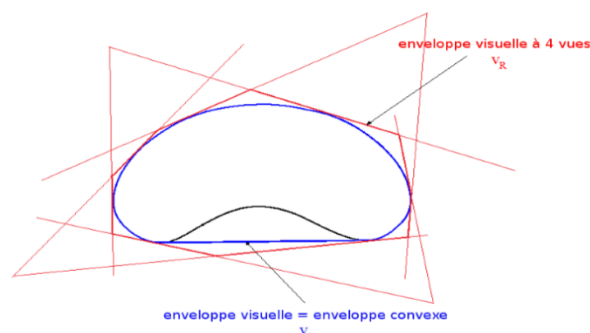


Figure 29 : Un objet 2D représenté avec une vue, sa silhouette et le cône correspondant.

<sup>13</sup> **Cône** : Un cône est défini de manière unique par son vertex et une section plane quelconque ne contenant pas le vertex.

Comme il est illustré dans la figure (30) on voit que pour chaque vue  $r$  l'objet est toujours contenu dans le cône  $V_r$ . Lorsque la taille de  $R$  tend vers l'infini et en même temps elle inclût toutes les vues possibles, on trouve que  $V_r$  converge vers la forme de l'enveloppe visuelle.

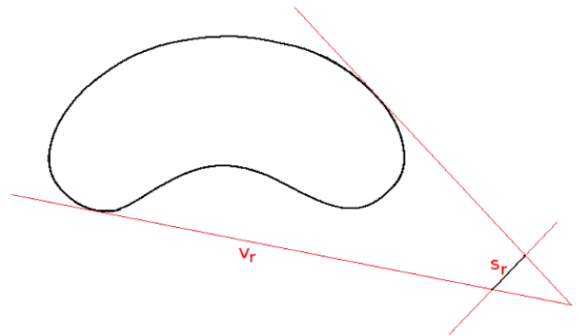


Figure 30 :  $v^\infty$  et  $V_r$  pour un objet 2D :  $v^\infty$  est l'enveloppe convexe de l'objet.

### 1.1 Problèmes liés à l'enveloppe visuelle

Il est impossible d'avoir une enveloppe visuelle complètement identique à la forme réelle de l'objet à reconstruire (Matsuyama, Nobuhara, Takai, & Tung, 2012).

- La méthode de l'enveloppe visuelle n'est pas utile dans le cas où l'objet contient des zones concaves (Voir figure 31). Ceci est dû au fait que l'enveloppe visuelle ne peut représenter que la forme de la surface 3D des pièces qui apparaissent sous forme de contours de silhouette 2D.
- L'enveloppe visuelle peut inclure certaines portions faussement positives appelées volumes fantômes comme illustré dans la (figure 34). Les deux zones B et C peuvent coexister avec les zones d'objet réel A et D en termes de consistance de silhouette.
- Les faux négatifs<sup>14</sup> dans les silhouettes 2D de n'importe quel point de vue peut gêner le résultat obtenu. Ce problème s'augmente si un grand nombre de caméras qui observent l'objet est utilisé.

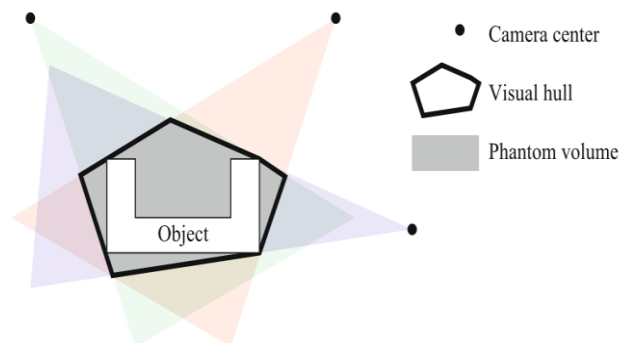


Figure 31 : Exemple d'objet avec une zone concave.

<sup>14</sup> **Faux négatifs** : ce sont des zones existantes dans l'objet réel mais ne sont pas disponibles dans la silhouette.



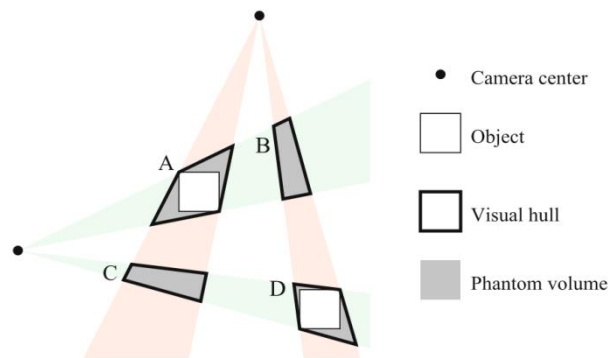


Figure 32 : Exemple d'objet avec fantôme.

## 1.2 Extraction des silhouettes

Comme il est mentionné ci-dessus, l'enveloppe visuelle peut être estimée à partir d'un ensemble de silhouettes prises par différents angles de vues. Dans chaque image il est nécessaire de déterminer l'ensemble des pixels correspondants à l'objet, ces pixels forment une silhouette. Les pixels qui n'appartiennent pas à la silhouette représentent la scène.

L'extraction de ces silhouettes doit être faite d'une façon exacte, pour ne pas avoir des conflits dans la forme 3D par la suite, elle représente un problème difficile qui continue à se développer et fait le sujet de plusieurs travaux de recherche. Il existe de nombreuses méthodes pour le calcul des silhouettes qui sont classées globalement comme suit :

- Méthodes basées sur la différence de pixels.
- Méthodes basées régions.
- Méthodes basées contours.

Les différentes méthodes d'estimation des silhouettes font partie de notre recherche, et nous citerons brièvement dans la section suivante le principe de chaque méthode.

### 1.2.1 Méthodes basées différence de pixels

Dans ce cas l'objet est filmé sur un fond de couleur unique afin de détecter les pixels qui séparent l'objet de son arrière-plan. Les couleurs les plus répandues sont le vert et le bleu.

La silhouette est estimée à partir des pixels qui donnent une différence par rapport à la couleur du fond. Il existe plusieurs méthodes (Wren, Azarbajani, Darrell, & Pentland, 1997) qui utilisent cette approche, car elle donne des résultats robustes mais elle est assez sensible au bruit issu par le capteur, ce qui nécessite l'utilisation des filtres pour l'amélioration des résultats.

### 1.2.2 Méthodes basées régions

Cette approche tente de traiter ces limitations à l'aide de remplissages des régions dans une image d'une façon heuristique. Ceci est fait par une propagation des critères locaux tels que la présence d'un faible gradient à l'intérieur de la région et d'un fort gradient aux bords de la région. Cette méthode est coûteuse en temps de calcul, surtout si un grand nombre d'images est utilisé.

### 1.2.3 Méthodes basées contours

Ce type d'approche consiste à traiter le problème d'extraction de silhouette comme un problème de minimisation d'une fonction d'énergie calculée sur un contour. Cette méthode prend en compte la présence d'un pic de gradient, il existe différentes méthodes de détection des contours. Le principal inconvénient lié aux méthodes basées contour réside dans la phase d'initialisation du système qui est généralement faite d'une façon manuelle, la mal initialisation donne des silhouettes manquantes.

### 1.3 Méthodes de calcul d'enveloppe visuelle

Les approches de la reconstruction 3D à partir de l'enveloppe visuelle sont classées en deux grandes catégories : L'approche surfacique (Benoît, 2009) et l'approche volumique (Dyer, 2001). La première catégorie est utilisée lorsqu'il s'agit de reconstruire la surface d'enveloppe visuelle en fournissant des points isolés ou un maillage. Mais on parle de la deuxième approche si on s'intéresse au volume délimité par l'enveloppe visuelle, et qu'on se base sur des discrétisations de l'espace.

#### 1.3.1 Les approches surfaciques

Ces méthodes se basent sur le calcul d'une représentation de la surface explicite de l'enveloppe visuelle. Les points ou bien les facettes, sont estimés par intersection des cônes de silhouettes. Ces approches sont particulièrement utilisées dans les contextes de reconstruction 3D de la géométrie pour le rendu depuis n'importe quel point de vue. Il devient possible de calculer la forme 3D à partir d'enveloppe visuelle surfacique en temps réel. Ces méthodes sont assez sensibles à l'extraction des silhouettes et aux imprécisions de calibrages des caméras, mais elles sont très utiles dans le contexte de modélisation pour le rendu.

#### 1.3.2 Les approches volumiques

Au contraire aux méthodes surfaciques, les approches volumiques estiment le volume de l'enveloppe visuelle par un ensemble de primitives élémentaires telles que : Les voxels<sup>15</sup>, les arbres octaux<sup>16</sup>. Les approches volumiques sont généralement plus rapides que les approches surfaciques, elles résistent aussi mieux que ces dernières aux silhouettes bruitées, ainsi que la reconstruction obtenue par les approches volumiques peuvent être raffinée d'une manière rapide.

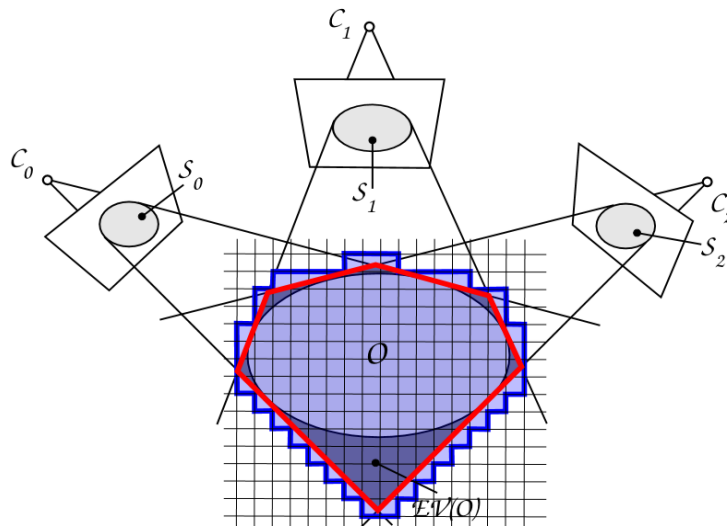


Figure 33 : Illustration d'une coupe de la géométrie estimée d'un objet  $O$ , en rouge enveloppe visuelle surfacique et bleue une enveloppe visuelle volumique.

<sup>15</sup> **Voxel** : C'est un pixel en 3D.

<sup>16</sup> **Arbre octal** : SDD de type arbre utilisé pour partitionner un espace 3D en le subdivisant sur 8 octants.

## 2. La stéréovision

La stéréovision vise à réaliser une tâche similaire à la vision humaine (Marr, 2010), cette méthode est apparue presque en même temps que la photographie, elle a connu un développement considérable au cours des dernières décennies.

Le but de la stéréovision est d'estimer la troisième dimension d'une scène à partir de deux images prises de deux positions différentes avec des caméras CCD<sup>17</sup> ou bien des caméras vidéo. Les caméras étant modélisées par une projection centrale, tous les points situés sur un rayon lumineux se projettent en même pixel  $p$  de l'image de gauche. Le centre de projection de la caméra de gauche dans le plan image de la caméra droite est appelé un épipôle. Cet épipôle avec un point de l'image forme une ligne épipolaire qui sera utilisée pour rechercher le correspondant  $p'$  de  $p$ , afin de pouvoir ensuite calculer la position de  $P$  dans l'espace par une triangulation en connaissant la position et l'orientation relatives des caméras.

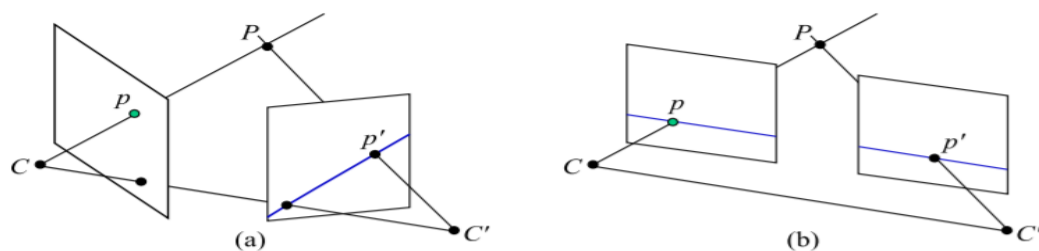


Figure 34 : (a) Géométrie épipolaire quelconque. (b) Géométrie épipolaire rectifiée.

On dit qu'une géométrie épipolaire est rectifiée lorsque tous les droites épipolaires sont parallèles et alignées.

La stéréovision s'effectue en trois étapes successives : Calibration, mise en correspondance et enfin la reconstruction 3D. Il existe un ensemble de problèmes liés à ce processus pendant leur automatisation (Monga & Horaud, 1993) :

- Problème d'estimation du modèle de projection et d'estimation des relations spatiales entre deux caméras.
- Problème de mise en correspondance qui vise à trouver pour chaque pixel d'une image, le pixel correspondant dans l'autre image.
- Le problème de la triangulation ou bien la reconstruction 3D (Hartley & Sturm, 1997).

### 2.1 Calibration

La calibration de la caméra est une tâche importante qui fait l'objet de nombreux travaux (Zhang Z. , 2000), c'est une étape cruciale pour la réussite du processus de la reconstruction tridimensionnelle. Elle a pour but d'estimer les valeurs des paramètres du modèle géométrique de la prise de vues, ces paramètres sont de deux types : Paramètres intrinsèques et paramètre extrinsèque.

<sup>17</sup> CCD : (Charge Coupled Device) : soit « Appareil à Transfert de Charges ». Désigne un capteur qui convertit la lumière en charges électriques.

### 2.1.1 Paramètres intrinsèques

Ces paramètres sont internes à la caméra, on note comme exemples à ces paramètres :

- $f$  : Distance focale.
- $u_0$  et  $v_0$  : Les coordonnées de la projection du centre optique de la caméra sur le plan image.
- $k_u$  et  $k_v$  : Les facteurs d'agrandissement de l'image.

### 2.1.2 Paramètres extrinsèques

Ce sont les paramètres géométriques d'orientation de position des caméras dans un repère attaché à la scène (le repère du monde). Parmi ces paramètres on a :

- $t_x, t_y, t_z$  : 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.
- $R_{3 \times 3}$  : Qui est la matrice de rotation permettant de passer du repère lié à l'espace de travail au repère lié à la caméra.

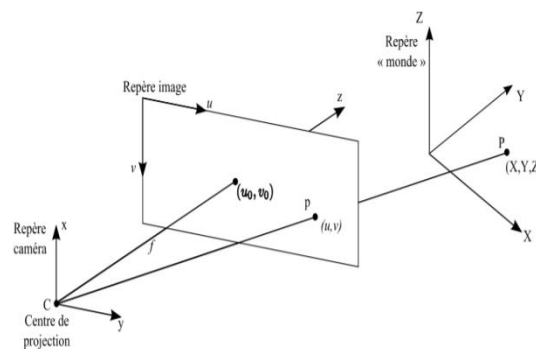


Figure 35 : Modèle géométrique de prise de vues.

Le passage entre les différents repères (repère monde et repère de la caméra) permet de retrouver les coordonnées de chaque point de la scène observée à partir de ces projections. Ce passage est caractérisé par deux matrices :

#### 2.1.3 Matrice essentielle

Cette matrice permet de décrire les paramètres de passage du repère objet vers le repère caméra à base des paramètres extrinsèques qui sont la position et l'orientation des capteurs l'un par rapport à l'autre selon le repère de référence.

#### 2.1.4 Matrice fondamentale

Cette matrice décrit les paramètres qui permettent le passage des coordonnées des points exprimés selon le repère caméra vers ceux exprimés selon le repère image, cela dépend des paramètres intrinsèques qui sont reliés par les caractéristiques de chaque caméra.

Avec la calibration, il devient possible d'estimer l'angle obtenu par l'intersection des deux rayons lumineux émis par le point correspondant, ceci est à partir de deux points homologues, il devient par conséquent possible d'estimer la position de ce point selon le repère objet (repère de référence). La calibration est classée en deux catégories globales.

##### 2.1.4.1 Calibrage classique (hors ligne)

Cette calibration est faite par l'analyse d'images d'une mire dont la géométrie est connue précisément. Le principe de base est de photographier cette mire dont on connaît la forme, ces formes sont détectées par la suite sur les images afin de calculer les paramètres intrinsèques.

### 2.1.4.2 Auto-calibration (en ligne)

C'est une calibration automatique qui ne nécessite pas une connaissance des composants de l'image. Ce type de calibration est surtout utilisé dans les recherches en robotique. 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 (Debevec, Taylor, & Malik, 1996).

## 2.2 La mise en correspondance

La mise en correspondance ou bien connue aussi par l'appariement (Gales, 2011) est un problème largement étudié et toujours posé en vision stéréoscopique, elle représente l'étape la plus délicate dans la reconstruction 3D. De nombreuses techniques ont été proposées pour résoudre ce problème, seront détaillé ci-dessous.

La mise en correspondance de deux images stéréoscopiques consiste à établir une relation point à point entre celle-ci, elle s'agit de retrouver sur deux images les paires de pixels qui correspondent à la projection d'un même élément de la scène.

### 2.2.1 Primitives stéréoscopiques

Avant de passer à l'étape de la mise en correspondance, il faut d'abord connaître ce que l'on doit chercher à apparier, on s'intéresse 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 notion de la disparité est bien introduite dans le domaine de la vision humaine, elle représente la différence de position de la projection d'un même objet sur la rétine de l'œil. La disparité dans le domaine de la mise en correspondance est définie par :

La disparité notée par *disp* est la différence de position entre la projection d'un même point de la scène observée sur les plans images des caméras de gauche et de droite.

$$d: N \rightarrow R^+$$

$$I_g(u, v), I_d(u, v') \mapsto disp(I_g(u, v)) = |v - v'|.$$

#### 2.2.2.1 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 deux pixels correspondants entre deux images gauche *g* et droite *d*. On a donc :  $d_{i,j}^{g \rightarrow d} = d_{i,j}$

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

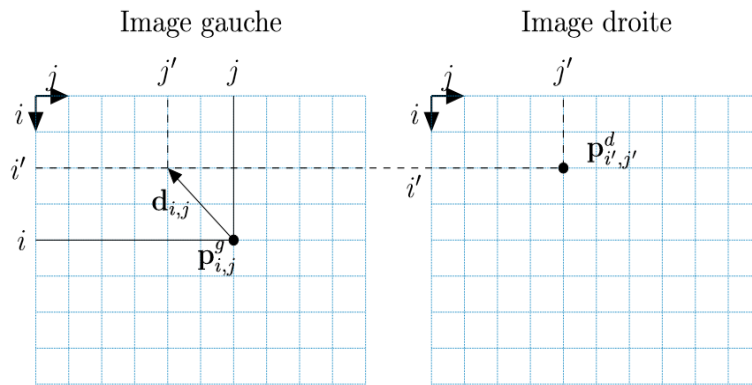


Figure 36 : Vecteur de disparité.

### 2.2.2.2 Carte de disparité

Le résultat final de la mise en correspondance est visualisé par une image dont le niveau de gris de chaque pixel représente sa disparité dans l'image de référence, cette image est appelée une carte de disparité, cette carte contient seulement les disparités des zones qui sont visibles dans les deux vues.



image de gauche de *Teddy*

image de droite de *Teddy*

carte de disparité  
de référence

Figure 37 : Exemple de carte de disparité. (D. Scharstein and R. Szeliski - Middlebury Stereo Vision dataset).

### 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 (Gales, 2011):

- **Information manquante** : Des fois on trouve que l'information nécessaire pour la mise en correspondance n'est pas disponible dans l'image, c'est le cas qu'on peut trouver dans :
  - **Les zones occultées** : Il s'agit des zones où un élément de la scène visible dans une image et n'apparaît pas dans l'autre image. (Figure 38)
  - **Les zones de raccourcissement** : Différents points de la scène se projettent en plusieurs pixels d'une image mais sur un même pixel de l'autre image. Ce cas peut être détecté dans les surfaces très inclinées par rapport au plan image.
  - **Les zones de discontinuité de profondeur** : Cela peut se traduire par des voisinages de deux correspondants qui ne se ressemblent pas nécessairement (figure 39).

- **Informations ambiguës** : Une information ambiguë est une information qui est présente dans une image mais il n'est pas possible de distinguer les correspondants.
  - **Régions homogènes** : Il est difficile de distinguer les points correspondants puisque tous les pixels de cette zone se ressemblent (figure 40).
  - **Régions avec textures répétitives** : Il existe une ambiguïté lorsqu'un même motif se répète plusieurs fois (voir figure 41).

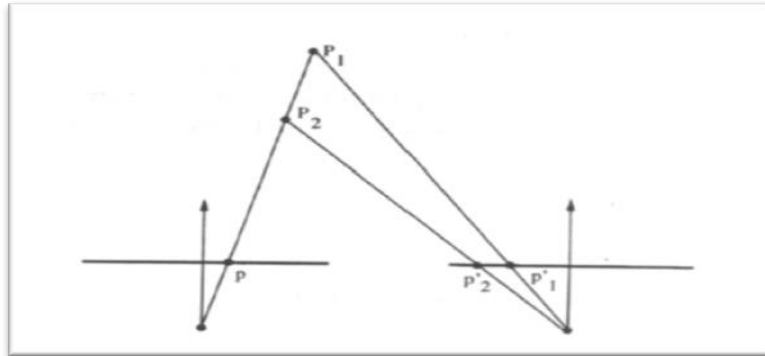


Figure 38 : Zone occultée – Le pixel marqué d'un point rouge a son correspondant occulté dans l'image de droite-

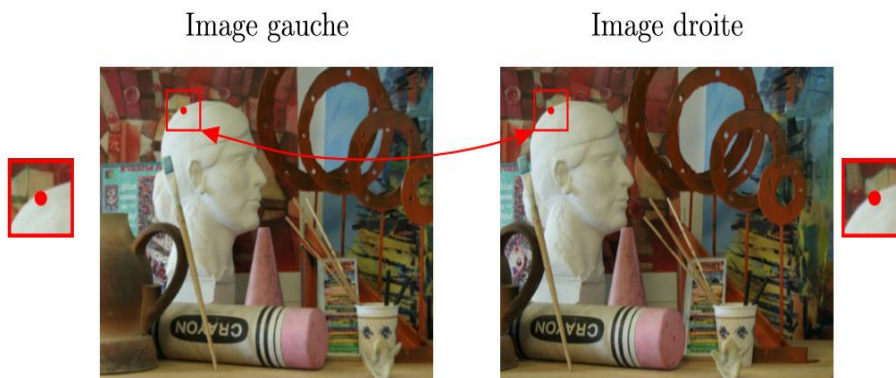


Figure 39 : Zone de discontinuité de profondeur – Les deux pixels d'un point rouge se correspondent mais leurs voisinages différent car ils sont proches d'une discontinuité de profondeur.

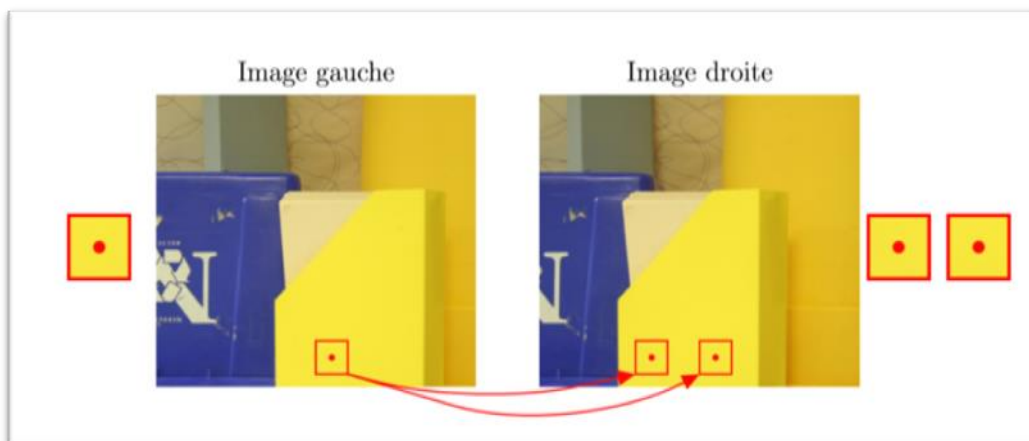


Figure 40 : Zone homogène –Manque d'information qui caractérise le pixel-

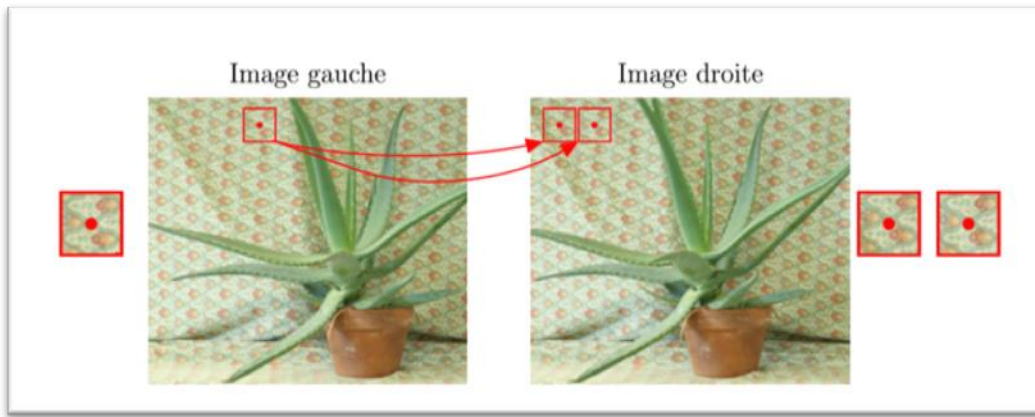


Figure 41 : Zone avec textures répétitives –Pour un pixel il existe plusieurs candidats dans l'autre image. (D. Scharstein and R. Szeliski - Middlebury Stereo Vision dataset)

### 2.2.4 Les contraintes de l'appariement

Un autre aspect assez important est la notion des contraintes de la mise en correspondance, on distingue deux types différents de ces contraintes : Contraintes géométriques et contraintes figurales. Ces contraintes sont utilisées dans le but de minimiser le champ de recherche de correspondants, ou bien pour éliminer certaines correspondances déjà établies.

#### 2.2.4.1 Contraintes géométriques

##### Contrainte épipolaire

Il est possible de réduire la zone de recherche à une seule droite épipolaire en utilisant des propriétés de la géométrie épipolaire comme illustré dans la figure 42.

$P$  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.

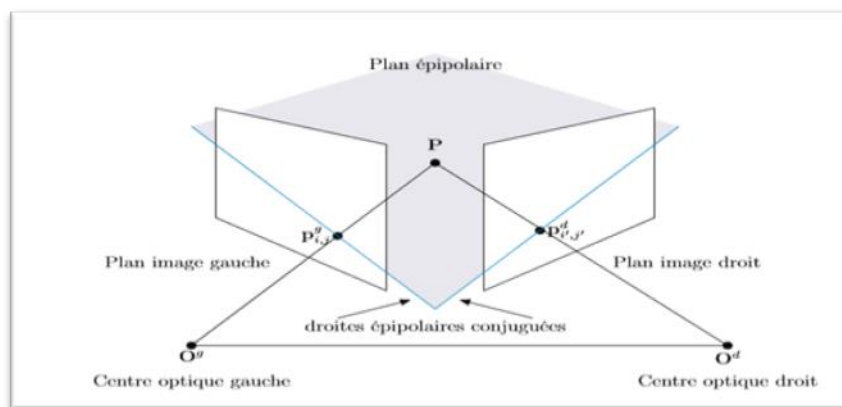


Figure 42 : Géométrie épipolaire.



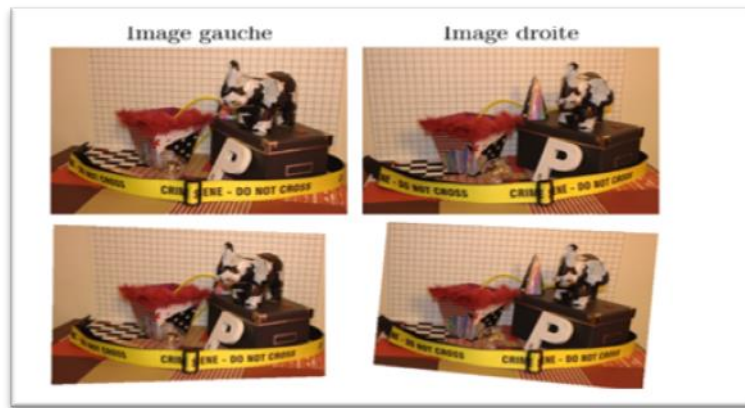


Figure 43 : Exemple de la rectification épipolaire. (G.Galles-Pixel matching for binocular stereovision by propagation of feature points matches and region-based randomized voting schema)

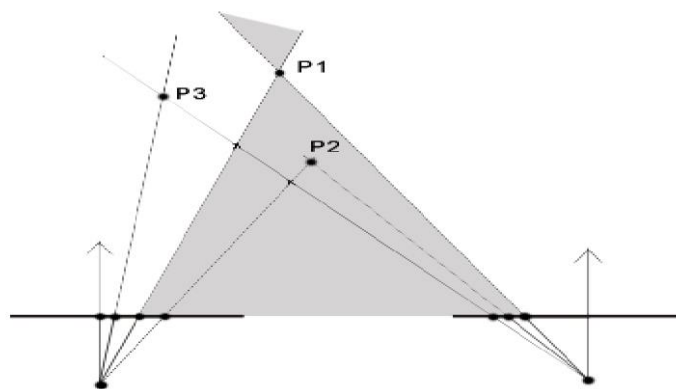


Figure 44 : Un point image correspondant à un point objet est un seul.

**Contrainte d’unicité**

Le principe de cette contrainte est qu’un pixel d’une image gauche ne peut avoir qu’un seul pixel correspondant sur l’image de droite. S’il existe pour un pixel de l’image de droite, plus d’un pixel homologue dans l’image gauche, cela peut être dû à la présence des régions occultées, c’est le cas d’un pixel de l’image gauche qui n’a pas d’homologue dans l’image de droite.

Cette contrainte peut être formulée par :

$$\forall I_g(u, j) \in P_d, \exists! I_g(u, v) \in P_g / (disp(I_g(u, v)) = |v - j|$$

Si cette contrainte n’est pas respectée, la formulation devient :

$$\begin{aligned} \forall I_g(u, j) \in P_d, \exists! I_g(u, v), I_g(u, v') \in P_g / (disp(I_g(u, v)) = |v - j|; \\ (disp(I_g(u, v')) = |v' - j| \text{ et } v \neq v' \end{aligned}$$

**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. Maintenant le principe de la contrainte d'ordre est le suivant :

L'ordre des pixels le long de la droite épipolaire doit être le même que celui de leurs correspondants le long de la droite épipolaire de l'autre image. Voir la figure 45

Cette contrainte peut être formulée par :

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

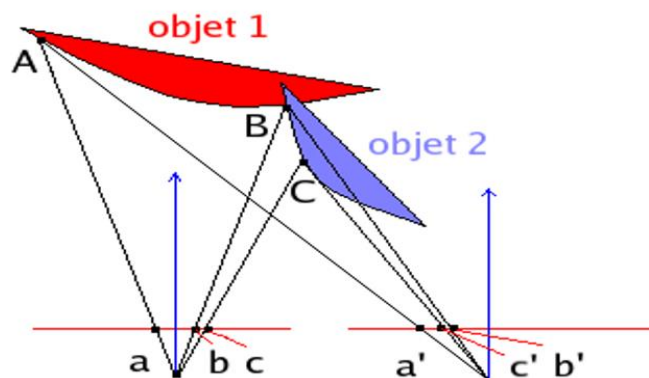


Figure 45 : Contrainte d'ordre.

#### 2.2.4.2 Contraintes figurales

Les contraintes figurales peuvent imposer une contrainte d'ordre géométrique, dans cette catégorie on trouve toutes les contraintes spécifiques à une primitive particulière.

#### Disparité locale constante

Les contraintes d'ordre et d'unicité ne fonctionnent pas dans les cas où la scène contient des objets avec des surfaces fortement inclinées. Ce problème peut être évité si on impose une limite au gradient de disparité entre deux paires de points appariés consécutifs. Ce faisant on limite, en fait, l'inclinaison des surfaces des objets présents dans la scène, et de fait, on limite par là même la variété des objets à reconstruire.

#### Continuité figurale

Cette contrainte permet de limiter les variations de disparité le long des contours. Elle permet aussi d'éviter que les points de contours de l'image gauche soient appariés avec des points de plusieurs contours dans l'image de droite. Cela suppose aussi qu'un contour peut appartenir à plusieurs surfaces mais ne doit pas traverser le bord d'un objet. Ce qui peut arriver malgré tout.

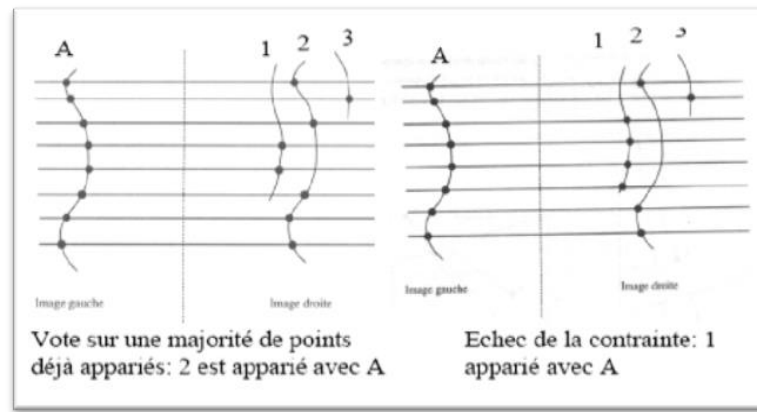


Figure 46 : Contrainte de continuité figurale.

### 2.2.5 Techniques de mise en correspondance

Il existe une diversité de méthodes (Fakhfakh, Détection et localisation tridimensionnelle par stéréovision d'objets en mouvement dans des environnements complexes : application aux passages à niveau, 2011) de mise en correspondance, les problèmes liés à cette dernière, vus ci-dessus, provoquent une instabilité de ce processus. 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 éparses. Une autre classification qui est largement étudiée et qui classe l'appariement en méthodes locales et méthodes globales. Ces méthodes seront détaillées dans la section suivante.

#### 2.2.5.1 Méthodes d'appariement denses

Ce type de méthode prend en considération l'ensemble des pixels composants l'image. La mise en correspondance exploite l'information contenue dans le voisinage d'un pixel. La vraisemblance de la mise en correspondance de deux pixels est mesurée avec la corrélation<sup>18</sup> existante entre les voisins de chacun de ces deux pixels. Cette vraisemblance peut être estimée par l'écart des valeurs photométriques ou colorimétriques des pixels voisinages correspondants. Cette technique suppose la continuité des surfaces dans les scènes réelles, traduite par une continuité de disparité.

#### 2.2.5.2 Méthodes d'appariement éparses

Le changement d'échelle dans les images causées par les transformations géométriques provoque une invariance dans les points d'intérêts<sup>19</sup>. L'inconvénient majeur de ce type de méthode réside dans l'impossibilité de reconstruire une scène 3D complète. L'appariement de certains points d'intérêts considérés comme non ambigus, est adopté dans plusieurs algorithmes de mise en correspondance. Ces points correspondent aux primitives qui peuvent avoir un certain nombre d'attributs. Plus les primitives sont discriminantes, plus l'appariement est précis, robuste et fiable.

#### 2.2.5.3 Méthodes d'appariement locales

Les méthodes locales mesurent la similarité des voisinages de deux pixels, Ce voisinage est appelé zone d'agrégation ou bien fenêtre de corrélation, pour plus de simplicité il est supposé que cette fenêtre carrée de corrélation soit obtenue par la formule :

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

<sup>18</sup>**Corrélation** : Utilisée pour calculer la similarité entre deux images.

<sup>19</sup>**Points d'intérêts** : Ensemble de points spécifiques tels que : Des coins, des courbes, des points de contours... qui sont utilisés comme des échantillons dans la phase de la mise en correspondance.

Cette fenêtre est centrée sur le pixel considéré. Il est noté aussi que  $a$  est le vecteur de dimension  $n^2$  qui contient les valeurs des pixels voisins du pixel considéré sur l'image de référence, et  $b$  celui de 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 47 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.

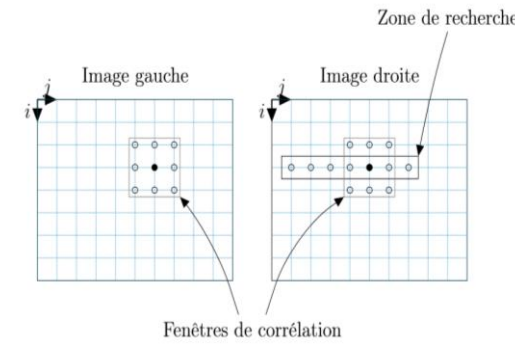


Figure 47 : Mise en correspondance locale.

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. 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 :

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

Tels que :  $I_1$  et  $I_2$  sont les deux images de la mise en correspondance,  $\underline{I_1}$  et  $\underline{I_2}$  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.

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

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

$$SAD(u, v, d) = \sum_{i,j} |I_1(u + i, v + j) - I_2(u + d, v + j)|$$

**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.

### 2.2.5.3 Méthodes d'appariement globales

Ces méthodes prennent en compte la totalité de l'image et cherchent à minimiser une fonction de coût qui exprime les erreurs de mise en correspondance. Le principe est d'estimer une fonction de disparité  $d$  permettant d'optimiser une énergie globale  $E$ , les deux étapes majeures de cette technique sont la définition d'une fonction d'énergie et l'application d'une technique d'optimisation. Le grand inconvénient de ce type de méthodes est la grande complexité des calculs.

Les recherches effectuées sur ces méthodes se basent sur la minimisation d'une fonction de coût globale. Cette technique regroupe les méthodes suivantes :

#### Programmation dynamique

Cette méthode consiste en la recherche d'une solution optimale par la minimisation de la complexité, en décomposant le problème général en plusieurs sous problèmes de complexité similaires. La programmation dynamique est équivalente à la recherche d'un chemin optimal dans un graphe. La première étape effectuée dans ce cas est la même dans les méthodes locales, la différence se situe au niveau de contraintes : Dans ce cas une contrainte de continuité de disparité est imposée, avec cette contrainte, on peut favoriser (défavoriser) la présence de discontinuités des profondeurs dans les appariements retenus. De cette façon l'approche assure une valeur de disparité en tout point de l'image de référence.

#### Relaxation

L'implémentation de cette méthode est très simple et facile, elle se base sur un algorithme itératif et localement parallèle et qui améliore une solution en minimisant localement l'énergie (RA & SW, 1983).

#### Flot de graphe (Graph-cut)

Le but de cette approche est de trouver une coupure de graphe qui possède un coût minimum parmi toutes les coupures possibles en se basant sur la minimisation d'une fonction d'énergie. Cette approche est abordée par Roy et Cox (Roy & Cox, 1998). Les terminaux du graphe représentent les disparités pouvant être associées à un pixel de l'image, et les différents sommets correspondent aux pixels de l'image. Cette méthode permet d'obtenir une carte de disparité avec des zones homogènes en gardant les contours et son inconvénient fondamental est l'aplatissement des objets.

## 2.3 Reconstruction 3D

La reconstruction tridimensionnelle à partir d'images stéréo est basée sur les correspondants entre ces deux images. En lançant pour chaque caméra un rayon passant par le centre de projection et par le plan image à la position pixel donnée. Il suffit de trouver le point d'intersection de ces deux rayons qui correspond à la position 3D du point observé, cette position signifie que la profondeur est calculée à base d'une simple triangulation. La profondeur donc n'est qu'une fonction de disparité : Pour chaque point de l'image de référence si on peut trouver son point homologue dans la deuxième image, il devient possible de retrouver sa profondeur, si la correspondance entre tous les pixels des deux images est connue on peut avoir une reconstruction 3D complète de la scène photographiée.

On peut distinguer différents cas au cours du processus de reconstruction tridimensionnelle (Monga & Horaud, 1993) :

- Les caméras sont calibrées donc les paramètres internes sont connus ainsi que la transformation rigide entre les deux caméras : Il s'agit donc d'une reconstruction euclidienne dans le repère de calibrage.
- Les paramètres internes sont connus mais la transformation rigide n'est pas connue : Dans ce cas il est nécessaire d'estimer la matrice essentielle qui permet d'extraire la transformation rigide des deux caméras à un facteur d'échelle pré. C'est une reconstruction euclidienne dans le repère de l'une des deux caméras.

Si les capteurs ne sont pas calibrés, c'est-à-dire ni les paramètres internes ni la transformation rigide sont connus. On a donc besoin d'estimer la matrice fondamentale, à partir de cette dernière on peut obtenir une reconstruction projective tridimensionnelle.

### 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 connus.

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  $\vec{P}$  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  $\vec{N}$  au plan du capteur. Et soient  $(\vec{u}, \vec{v})$  les vecteurs des directions principales du capteur.

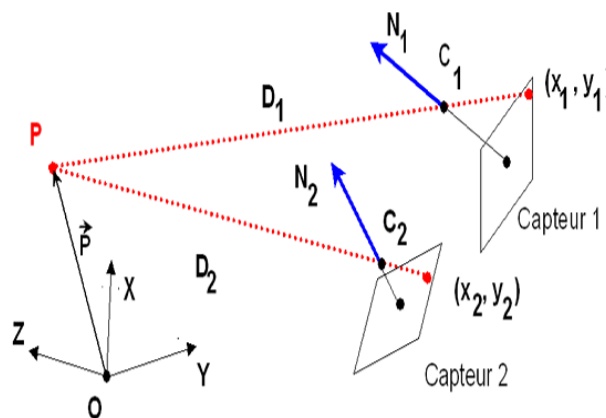


Figure 48 : Système de projection des scène –Représentation de triangle formé par le système.-.

La projection de  $P$  dans le capteur1 est donnée par :

$$x_1 = f \frac{\vec{D}_1 \cdot \vec{u}_1}{\vec{D}_1 \cdot \vec{N}_1} \quad y_1 = f \frac{\vec{D}_1 \cdot \vec{v}_1}{\vec{D}_1 \cdot \vec{N}_1} \quad \text{avec : } \vec{D}_1 = \vec{P} - \vec{C}_1$$

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

$$x_2 = f \frac{\vec{D}_2 \cdot \vec{u}_2}{\vec{D}_2 \cdot \vec{N}_2} \quad y_2 = f \frac{\vec{D}_2 \cdot \vec{v}_2}{\vec{D}_2 \cdot \vec{N}_2} \quad \text{avec : } \vec{D}_2 = \vec{P} - \vec{C}_2$$

Afin de déterminer les trois composantes  $(x, y, z)$  du point  $P$ , il faut résoudre cet ensemble d'équations au sens des moindres-carrés<sup>20</sup>.

<sup>20</sup> **Moindres-carrés** : C'est-à-dire résoudre l'ensemble d'équations d'une façon expérimentale

## **Conclusion**

Dans ce chapitre nous avons traité deux méthodes différentes de reconstruction 3D basées images : L'enveloppe visuelle et la stéréovision. Nous avons détaillé la démarche de chaque méthode, son utilité et les problèmes rencontrés au cours de leurs automatisations.

L'enveloppe visuelle permet de générer la forme tridimensionnelle par une intersection des cônes des silhouettes. La forme produite fournit un volume englobant l'objet d'intérêt.

Par contre la stéréovision permet, quant-à-elle, l'extraction de la structure 3D d'une scène à partir de deux images prises sous différents points de vue.

Comme on a vu dans ce qui précède, chacune de ces deux méthodes possède des limitations qui influent sur le résultat final de la reconstruction.

Si on combine les résultats de ces deux méthodes, on peut avoir des formes 3D plus robustes et très proches de la forme réelle de l'objet. Ceci sera le sujet du chapitre suivant, où on va voir quelques travaux qui se basent sur cette combinaison.

# Chapitre IV

---

## *Travaux connexes*

---



## Introduction

Dans les chapitres précédents nous avons vu deux méthodes différentes de reconstruction 3D à partir d'images : L'enveloppe visuelle et la stéréovision et, nous avons expliqué leur principe de fonctionnement, les avantages et les difficultés liés à chaque méthode.

L'enveloppe visuelle ne peut pas traiter les zones concaves même si un grand nombre de silhouettes est utilisé (Crowley, 2010). La stéréovision est non utilisable pour les cas des objets faiblement texturés, pour les régions occluses, aussi pour le cas où une partie de l'objet est sphérique ou curvilignes (Brandou, 2008). Ces difficultés causent un appariement difficile (voir impossible), et donnent des résultats non fiables et instables.

Par contre on trouve que ces deux méthodes sont bien complémentaires en nature et peuvent être combinées pour surmonter leurs inconvénients dans le but d'optimiser la qualité de la reconstruction 3D.

L'idée de l'hybridation de la méthode d'enveloppe visuelle avec la stéréovision est déjà abordée. Dans ce chapitre nous allons présenter quelques travaux de combinaison, et nous détaillerons la démarche de ces travaux, les objectifs, les inventions et les limitations des résultats de chaque travail.

## 2. Travaux de Eric Nowark et al

Nowark et al (Nowak, 2003) ont proposé une méthode hybride pour la reconstruction 3D d'un joueur de Golf. L'utilisation de la méthode d'enveloppe visuelle seulement induit des erreurs de modélisations des zones concaves, ces erreurs sont réduites par la technique de la stéréovision qui permet de positionner précisément des points dans l'espace.

Les joueurs sont filmés par un ensemble de caméras synchronisées ; qui capturent les mouvements 3D. Il s'agit en premier lieu de modéliser l'enveloppe visuelle du joueur à partir des silhouettes. La méthode de création d'enveloppe visuelle choisie est la méthode volumique. Le processus de sa création est le suivant :

- **Calibration :**

Une grille de calibration est utilisée avec une géométrie bien connue et un emplacement visible par toutes les caméras, il est supposé que le centre optique se situe au centre de la caméra. L'orientation des caméras, les positions et leurs distances focales sont donc retrouvées grâce à des outils de calibration avec correspondance 3D-2D, ces outils permettent de fournir :

- Distance focale en cm :  $f_c$ .
- Position de la caméra en cm dans le repère monde :  $T_c$ .
- Matrice de rotation permettant de passer du repère monde au repère caméra :  $R$ .

- **Segmentation :**

Cette étape est réalisée manuellement dans le but de traiter la suite du problème d'une façon rapide.

- **Création de l'enveloppe visuelle :**

Pour induire l'enveloppe visuelle, une création par arbres octaux est utilisée, il s'agit d'une méthode volumique. Cette dernière fait appel à des théories issues de divers domaines, le principe est de tester si un cube se projette intégralement à l'intérieur de toutes les silhouettes de l'objet à modéliser.

Si cette condition est vérifiée le résultat d'enveloppe sera correct, si ce n'est pas le cas, ce résultat sera manquant. Cette méthode n'a d'intérêt que si la position d'un cube par rapport à la silhouette peut être déterminée rapidement, sinon il est plus simple de travailler directement avec les voxels (cubes élémentaires).

Cette phase est suivie par un calcul de la résolution maximale de l'enveloppe visuelle à partir de la formule suivante :

$$\delta_{theorique} = \left\{ \frac{L_c d_i}{L_p f_i} \right\}$$

La résolution maximale de l'enveloppe visuelle est la distance  $\delta = \delta_{theorique}$  tels que :

$L_c$  est la largeur en cm du plan film théorique,  $L_p$  la largeur en pixels des images,  $C_i$  le centre optique de la caméra  $i$ ,  $d_i$  le centre du plan focal physique de cette dernière et  $f_i$  sa distance focale. Enfin  $d_i$  le point 3D de l'objet à modéliser le plus éloigné de la caméra  $i$ .

**Remarque :** Cette valeur théorique est calculée à l'aide d'une configuration de théorème de Thalès.<sup>21</sup>

- **Transformation de l'enveloppe visuelle en maillage :**

L'algorithme de transformation en maillage parcourt tous les voxels composant le volume, et à chaque fois qu'un voisin est absent, une face carrée est créée, identique à la face frontière entre le voxel et l'emplacement de son voisin absent. Cette face est ensuite ajoutée au maillage.

A cette étape, la phase de création d'enveloppe visuelle est terminée, le résultat à ce moment est combiné avec la méthode stéréovision. Les étapes de ce processus sont comme suit :

- **Détection des points caractéristiques :**

Pour effectuer la mise en correspondance, des points caractéristiques sont automatiquement extraits des images. Ces points doivent se différencier suffisamment de leurs voisins pour faciliter les appariements. Un pixel  $(i, j)$  est défini comme un point caractéristique si et seulement si :

$$D_x = \max \{ |NG(i, j) - NG(i-1, j)|, |NG(i, j) - NG(i+1, j)| \} \geq f$$

Différentiation sur l'axe des abscisses.

Et:

$$D_y = \max \{ |NG(i, j) - NG(i, j-1)|, |NG(i, j) - NG(i, j+1)| \} \geq f$$

Différentiation sur l'axe des ordonnées.

Tels que  $NG(i, j)$  est le niveau de gris des pixels  $(i, j)$  qui sont compris entre 0 et 1. Le seuil de différenciation minimal est  $f$  choisi arbitrairement tel que  $0 \leq f \leq 1$ .

Enfin pour ne pas avoir plusieurs points caractéristiques, un pixel qui a la plus grande valeur de différenciation est choisi :  $\min(D_x, D_y)$ .

- **Calcul de la position 3D avec stéréovision :**

Le but de l'algorithme est de calculer la position 3D des points caractéristiques vus précédemment, le principe est de calculer la disparité entre ces points à partir de la méthode SAD (sum of absolute difference) et SSD (sum of square difference). On constate que l'application de cet algorithme permet de placer tous les points à la bonne profondeur sur l'objet 3D.

<sup>21</sup> **Théorème de Thalès :** théorème de la géométrie qui consiste à calculer la longueur des droites.

- **Déformation du maillage.**

Pour obtenir la structure 3D, deux informations différentes sont disponibles. Premièrement, l'enveloppe visuelle définit un nuage de points très dense, dont les positions sont correctes dans les zones convexes, et incorrectes dans les zones concaves, et deuxièmement l'algorithme de stéréovision qui permet de positionner précisément des points dans l'espace. Le maillage est donc déformé selon cet ensemble de points corrects.

### Avantages et inconvénients

Ce travail est appliqué sur un type particulier d'objet « Golfeur ». En ce qui concerne l'enveloppe visuelle, si le point de vue du joueur est en face de la caméra, la forme finale de cette enveloppe est parfaitement présentée. Mais si ce n'est pas le cas, si le point de vue créé par une seconde caméra de synthèse où les concavités sont vues de profil on trouve que l'objet 3D résultant est mal modélisé.

La combinaison du résultat d'enveloppe visuelle avec la stéréovision permet d'améliorer quelques parties concaves, mais pas d'une façon précise. Après la déformation du maillage certaines parties deviennent plus visibles, mais on trouve que quelques parties du corps sont mal construites ; comme par exemple la représentation d'un bras plié.

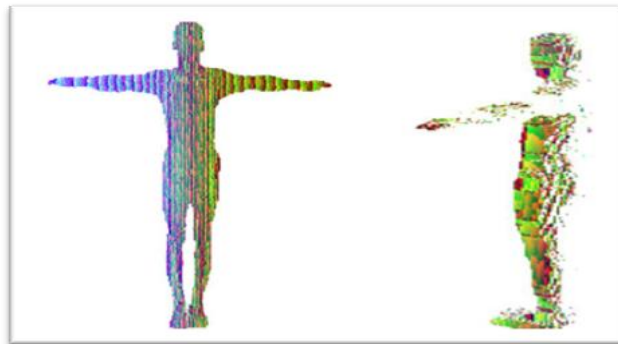


Figure 49 : Exemple des résultats obtenus par l'approche de Nowark et al.

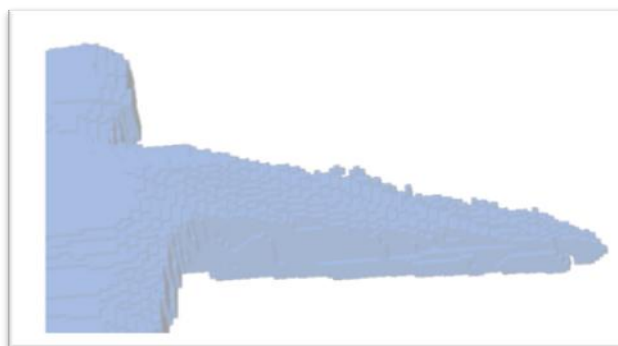


Figure 50 : Exemple de modélisation des zones concaves.

### 3. Travaux de Ming Li et al

Li et al (Li, Schirmacher, H., & Siedel, 2002) ont adopté le principe de combinaison de la méthode de l'enveloppe visuelle avec celle de la stéréovision pour l'acquisition de scènes réelles dynamiques aux taux interactifs. Tout d'abord ils ont utilisé des silhouettes de plusieurs vues pour construire une enveloppe visuelle qui représente une estimation initiale de l'objet dans la scène. Cette dernière est utilisée par la suite pour limiter la plage de disparité au cours du calcul de la profondeur à partir de la stéréovision, ce qui permet d'améliorer la qualité de la reconstruction stéréo ainsi que la vitesse. L'un des inconvénients de la méthode d'enveloppe visuelle est l'impossibilité de reconstruction des zones concaves et les détails de la surface, c'est pour cette raison que le travail consiste à la combiner avec la stéréoscopie.

Le système implémenté se compose de plusieurs caméras observant la même scène à partir de différentes directions de visualisation, chaque paire de caméras est connectée à un ordinateur client qui communique avec un autre ordinateur serveur via un réseau TCP/IP standard, à savoir que toutes les caméras sont calibrées à l'avance comme il est présenté par :



Figure 51 : Trois paires de caméras sont disposées le long d'un arc.

L'initialisation du système comprend l'enregistrement d'une image d'arrière-plan pour chaque caméra et l'envoi d'informations d'étalonnage de chaque client vers le serveur, ensuite le système commence le cycle de traitement qui est défini comme le temps de traitement d'un ensemble d'images synchronisées collectées par toutes les caméras. Selon la direction du transfert de réseau, le processus de ce travail se compose de trois étapes :

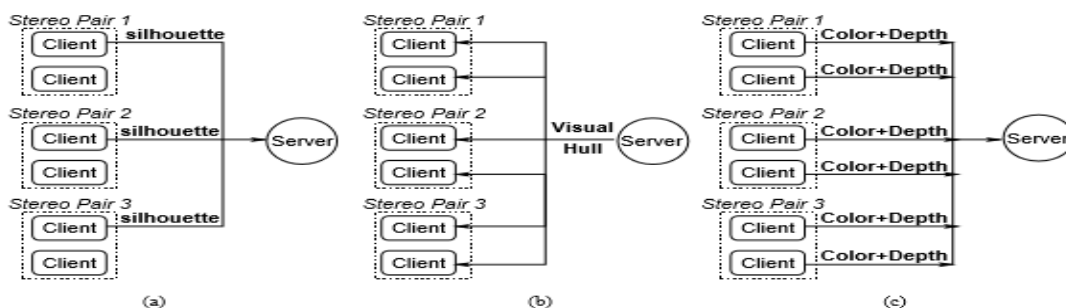


Figure 52 : Démarche de l'approche proposée. (a) : Extraction des silhouettes. (b) : Création d'enveloppe visuelle. (c) Estimation de profondeur à partir de la stéréovision.

- **Extraction des silhouettes :**

Tout d'abord, les paires d'images stéréo sont rectifiées individuellement pour s'aligner le long des lignes de balayage. Pour chaque image, l'objet en mouvement du premier plan est segmenté du fond acquis précédemment. Comme les caméras stéréo sont très proches l'une

de l'autre, l'utilisation des deux silhouettes n'améliore pas de façon significative la reconstruction de l'enveloppe visuelle. Par conséquent une silhouette est extraite pour une caméra de la paire stéréo, cette silhouette est envoyée par la suite au serveur.

- **Calcul d'enveloppe visuelle :**

Lorsque toutes les silhouettes sont disponibles, le serveur commence le calcul de l'enveloppe visuelle à base d'une méthode surfacique car elle requiert moins de mémoire et convient pour un rendu direct rapide, cette enveloppe est diffusée par la suite à toutes les machines clientes.

- **Calcul de profondeur à partir de la stéréovision :**

Tous les clients utilisent l'enveloppe visuelle pour guider le calcul de la profondeur. A savoir que les informations sur les silhouettes sont disponibles, le calcul de la stéréoscopie ne doit être effectué que sur le masque d'objet au premier plan au lieu que sur l'image entière. Tant que les images sont rectifiées au début, alors le serveur commence directement la phase de la mise en correspondance, à ce moment la méthode SAD est choisie pour le calcul des correspondants.

Une fois que la correspondance entre les pixels est établie, il devient possible de générer la carte de disparité qui à partir de laquelle la carte de profondeur peut être obtenue d'une manière simple à l'aide de l'information de caméra, les cartes de profondeur<sup>22</sup>, ainsi que les images en couleur obtenues, sont ensuite renvoyées au serveur pour le rendu.

## Avantages et inconvénients

L'architecture du système a été conçue pour distribuer la charge de calcul entre le serveur et les clients. On trouve que le temps nécessaire pour l'acquisition des images, la rectification, la récupération des silhouettes et la reconstruction 3D est indépendant du nombre de de paires stéréo, ce qui offre une bonne évolutivité du travail dans des petites durées de traitement.

Le résultat obtenu dans ce travail est présenté dans la figure suivante :



Figure 53 : Résultat de modélisation de la méthode de Li et al (Exemple 1).

<sup>22</sup> **Carte de profondeur** : Image ou un canal d'image qui contient des informations relatives à la distance entre les surfaces des objets de scène de caméra.

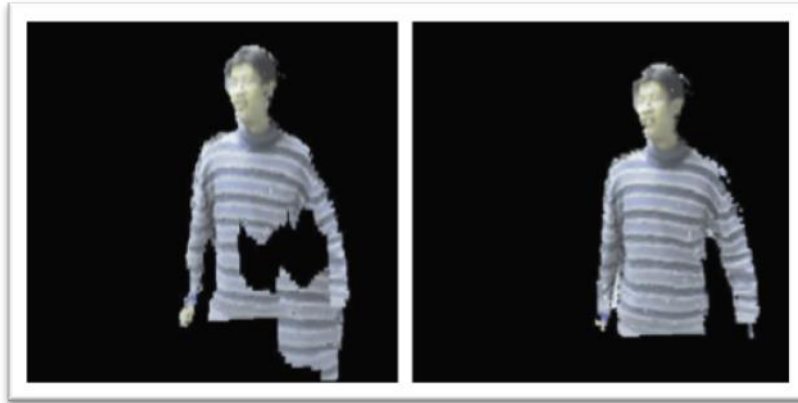


Figure 54 : Résultat de modélisation de la méthode de Li et al (Exemple 2).

#### **4. Travaux de Matsumoto et al**

Un autre algorithme de reconstruction tridimensionnelle proposé par Matsumoto et al (Fakhfakh, Debled-Rennesson, Domenjoud, Kerautret, & Even, 1999), qui se base aussi sur la combinaison des deux méthodes vues précédemment. Dans cet algorithme l'enveloppe visuelle permet d'estimer une forme approximative de l'objet cible, puis la stéréovision est appliquée dans la zone estimée pour redéfinir la forme. Ce dernier remédie aux lacunes de chaque méthode et fournit une forme plus précise et un traitement rapide.

##### **Formalisation du problème :**

- Une série d'images de différentes vues tout autour d'un objet cible est en entrée.
- Les positions de la caméra sont connues pour toutes les images, ainsi que les paramètres intrinsèques et extrinsèques.
- La zone probable de la présence d'objet peut être définie dans un espace 3D.
- La description est basée sur les notations suivantes :
  - $V$  : Ensemble de voxels.
  - $V_s$  : Un ensemble de voxels qui correspondent à la surface de l'objet cible dans l'espace voxel ;  $V_s \subseteq V$  (Surface d'objet).
  - $I^i$  ;  $i^{\text{ème}}$  plan d'image d'entrée,  $I_o^i$  et  $I_s^i$  signifie les  $i$ -ème images d'entrée originales (niveau de gris ou bien couleur) et sa silhouette (B/W), de plus  $I_{ss}^i$  et  $I_{sb}^i$  sont la silhouette et la zone de l'arrière-plan respectivement.
  - $P(v, i)$  : Est la projection du point  $v$  dans  $I^i$ .

Les images originales et les images de silhouette donnent des indices et des contraintes importantes concernant la zone de présence de l'objet dans l'espace de voxel. Un voxel de  $V_s$  est projeté sur la partie silhouette des images d'entrée (contrainte de silhouette). D'autre part, les pixels d'images d'entrée sur lesquels un voxel de  $V_s$  est projeté devraient avoir la même couleur ou intensité (contrainte de couleur). En supposant :

- 1) Environnement d'illumination homogène.
- 2) Images sans bruit.
- 3) Surface Lambertienne<sup>23</sup>.

<sup>23</sup> **Surface Lambertienne** : C'est une surface qui reflète la lumière également dans toutes les directions. Les surfaces Lambertiennes sont également appelées réflecteurs isotropes / Uniformes / Parfaitement diffus.

**Contrainte de silhouette :**

La contrainte de silhouette locale pour  $v \in V_s$  dans  $I^i$  est définie par :

$$C_s(v, i) = \begin{cases} 1 & \text{si } p(v, i) \in I_{ss}^i \\ 0 & \text{sinon} \end{cases}$$

La contrainte totale de silhouette est :

$$\prod_i \prod_{v \in V_a} C_s(v, i) = 1. \quad (1)$$

**Contrainte de couleur :**

$$Co(v, i) = \begin{cases} 1 & \text{si } v \text{ est projeté sur } I^i \text{ sans occlusion} \\ 0 & \text{sinon} \end{cases}$$

$$\sum_{v \in V_a} \sum_{i \in O^v} D(v, i) = 0, \quad (2)$$

Où  $D(v, i)$  représente une évaluation de dispersion de couleur pour  $v$ . un exemple typique pour  $D(v, i)$ :

$$\sum_{j=i-m}^{i+m} \sum_{blockara} \{v(p(v, i)) - v(p(v, j))\}^2$$

Où  $v(p(v, j))$  est la valeur de couleur dans  $p(v, j)$ .

Le problème cible est de déterminer l'ensemble  $V_s$  en utilisant ces contraintes, le but principal de l'approche proposée est de déterminer l'ensemble de voxels localisés  $V_s$  à partir de  $V$ .

L'enveloppe visuelle permet de produire un ensemble de voxels  $V_b$  qui donne une estimation approximative de la zone de présence d'objet en fonction de la condition de contrainte de silhouette.

La stéréovision basée sur les voxels permet de localiser  $V_s (\subseteq V_b)$  avec la contrainte de couleur.

**Déroulement du processus de l'approche :**

- **Enveloppe visuelle basée sur vote :**

Comme il est mentionné ci-dessus, la silhouette d'un objet cible donne des contraintes importantes de la forme de l'objet. Plus précisément, un point dans l'espace 3D projeté dans l'intérieur d'une silhouette représentant l'objet, alors qu'un point qui est projeté sur l'extérieur de la silhouette est certainement libre. L'espace 3D occupé par l'objet peut être estimé plus précisément en utilisant plusieurs images de silhouette prises à différents points d'observation.

La méthode d'enveloppe visuelle est simple et conventionnelle, elle considère les silhouettes comme des contraintes fortes. Cela signifie que la méthode donne des silhouettes avec des erreurs. Par contre dans la méthode utilisée dans ce cas, les erreurs de silhouette sont sérieusement résolues où elles sont traitées comme des contraintes faibles. Dans l'algorithme étendu, l'estimation de forme est réalisée avec un système de localisation de vote (Figure 55). Dans ce cas seulement les voxels pleinement marqués sont jugés dans la zone de l'objet.

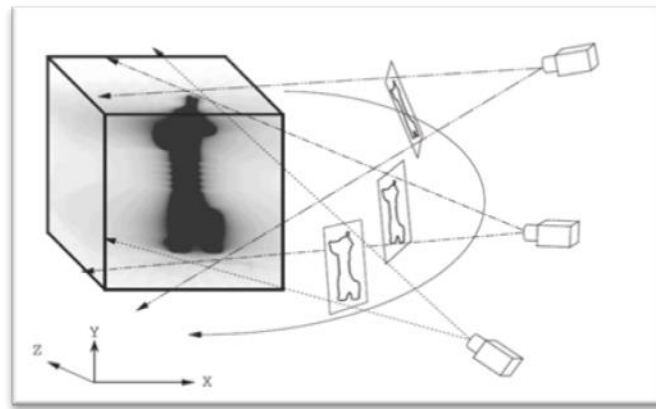


Figure 55 : Shape from silhouette basé vote.

```

For all voxels:           — (initialization)
    score(v) := 0.
    V_b := φ.
For all voxels:           — (voting)
    For all silhouettes:
        If C_s(v, i) ≠ 0,
            Then score(v) := score(v) + 1.
For all voxels:           — (localization)
    If score(v) ≥ Threshold,
        Then V_b := V_b ∪ {v}.
    
```

Figure 56 : Vote sur l'espace voxel : la densité montre le score de chaque voxel.

▪ **Stéréovision basée vote :**

Si l'objet cible contient une partie concave, l'enveloppe visuelle ne reconstruit pas correctement la forme 3D. L'utilisation de la stéréovision permet de récupérer la connectivité ainsi que la surface de l'objet et des informations de textures. Cette étape comprend les sous étapes suivantes :

○ **Estimation partielle de surface :**

L'idée de base dans cette phase est : Si un voxel  $v$  correspond à la surface de l'objet, la couleur de son point de projection sur chaque plan d'image est similaire à l'autre. Les erreurs de ces couleurs sont accumulées sur chaque voxel (erreur de vote). Ensuite, les voxels avec l'erreur minimale accumulée sont choisis, cet ensemble de voxels peut être déterminé pour chaque point d'observation et est considéré comme des surfaces partielles pour le point d'observation. A ce moment le  $i$ -ème point d'observation est référencié à la surface partielle comme  $V_{SS}^i$ .

Il est supposé aussi que l'ensemble d'images observables  $O^v$  inclut au moins  $n$  membres consécutifs pour tout  $v \in V_{sp}^i$ , ceci est pour tout point d'observation  $i$ . Cela signifie que le problème d'occlusion peut être évité pour  $n$  images d'entrée consécutives pour tous les voxels de toutes les surfaces partielles.

Soit  $V_{rv}^i$  un ensemble de voxels dans le long d'un rayon du  $i$ -ème point d'observation à travers  $v$ . Pour le  $i$ -ième point d'observation, la surface partielle  $V_{sp}^i$  est identifiée avec la procédure représentée dans (a) sur la figure (IV.4), les étapes sont appliquées de façon répétée à tous les points d'observation.

A la fin, un ensemble de surfaces partielles est obtenu comme des images originales.



o **Génération totale de modèle :**

Ce niveau génère un modèle 3D total à partir des surfaces partielles. Bien qu'une surface partielle soit calculée à chaque point d'observation, chacune peut avoir des erreurs. En intégrant les surfaces partielles avec le programme de localisation de vote, les erreurs sont arrondies pour obtenir le modèle total de l'objet avec une plus grande précision.

Soit  $L(v, i)$  la distance euclidienne<sup>24</sup> entre  $v$  et la  $i$ ème point d'observation.  $V_s$  est un ensemble de voxels qui donne la représentation volumique de l'objet. Le processus d'intégration est détaillé dans la partie (b) de la figure (57). Cette procédure ne fournit pas la solution exacte qui satisfait à la constante de couleur détendue mais elle donne une bonne approximation.

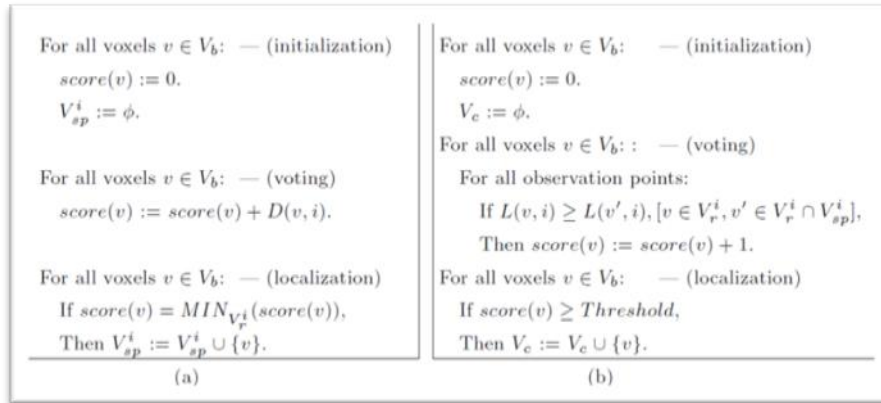


Figure 57 : (a) Estimation de surface partielle. (b) Génération totale de modèle.

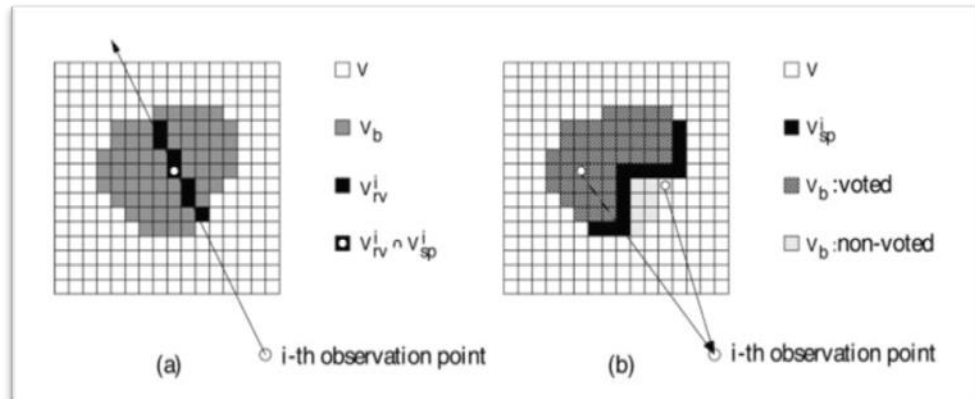


Figure 58 : (a) Estimation de surface partielle (b) Génération totale du modèle.

▪ **Combinaison d'enveloppe visuelle avec stéréovision :**

Pour tester l'utilité de l'algorithme adopté, un numériseur 3D est construit. Le numériseur comprend un plateau tournant, une caméra monoculaire et un ordinateur personnel. Ce numériseur est capable de générer automatiquement le modèle 3D texturé de l'objet cible.

Ce processus se compose de cinq étapes majeures :

<sup>24</sup> **Distance Euclidienne :** C'est une distance ordinaire c'est-à-dire entre deux points dans l'espace euclidienne.

- **Calibration :**  
 Dans ce processus, la position de la caméra par rapport à la table tournante est automatiquement calculée en utilisant un objet de forme connue placé sur le plateau tournant, l'objet est capturé à partir de divers points d'observation (images originales). Dans la mise en œuvre actuelle, un panneau d'étalonnage avec motifs de cercle est utilisé. Après détection de chaque position de cercle, la transformation de Hough (Séverinie, 2008) est appliquée pour calculer la position de caméra pour renforcer la robustesse.
  
- **Capture d'images :**  
 Dans cette étape, une série d'images est capturée pour couvrir toutes les vues de l'objet. En outre, l'image d'arrière-plan est capturée aussi. Cette dernière est presque identique aux images d'origine, sauf qu'elle exclut l'objet cible et est utilisée pour extraire des silhouettes.
  
- **Extraction des silhouettes :**  
 Dans cette étape, l'extraction de silhouette naïve, ou l'extraction au niveau de pixel, est effectuée en premier. Ensuite, extraction de silhouette supplémentaire, extraction de région-niveau, est appliquée, ce qui améliore la précision de la silhouette obtenue. Dans l'extraction au niveau de la région, la moyenne de la soustraction absolue à chaque pixel dans une région est calculée pour juger si la région correspond à l'objet ou non.
  
- **Modélisation de formes :**  
 L'étape de modélisation de la forme se compose de deux processus : la modélisation volumétrique et la conversion de la représentation de surface. Dans la conversion de représentation de surface, de nombreux polygones initiaux sont générés d'abord en connectant des polygones de surface adjacents. Ensuite, une opération de réduction de polygone fusionne des polygones pour réduire le nombre total de polygones.
  
- **Acquisition de texture :**  
 La texture d'une pièce peut être dérivée à partir de plusieurs images d'origine, car une pièce peut être vue à partir de divers points de vue. Par conséquent, une image originale doit être identifiée pour chaque polygone pour en dériver sa texture. Afin de déterminer l'image source de manière appropriée, les points suivants doivent être considérés :
  - 1) Une plus grande quantité d'informations de texture dans l'image source conduit à une meilleure dérivation de la texture.
  - 2) Le lissage de texture sur la limite des pièces est nécessaire pour masquer les bords.



Figure 59 : Présentation du système.

L’algorithme est testé sur des objets en céramique (un château miniature et un écureuil). Dans les expériences, l’étape de rotation de la table était de dix degrés, donc 36 images originales ont été prises. La taille de l’image était de 640 à 480. La résolution de l’espace de voxel est de 320×320×240. Les figures ci-dessous montrent des exemples d’images originales de chaque objet, des modèles en fil de fer<sup>25</sup> 3D obtenus et des images 3D texturées en résultats. Les résultats de modélisation montrent que cette approche reconstruit la forme avec une précision pratique. Le temps d’exécution de chaque étape de cette approche est présenté dans le tableau 2.

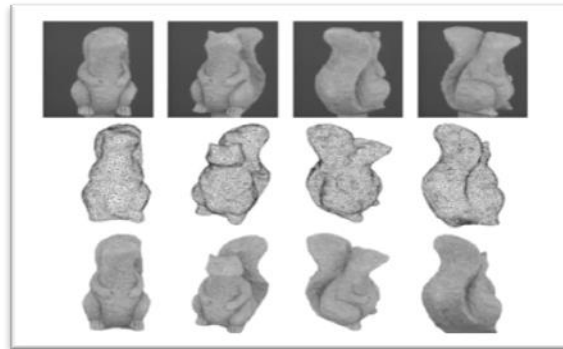


Figure 60 : Exemples d’images originales et de résultats de modélisation (Ecureuil)

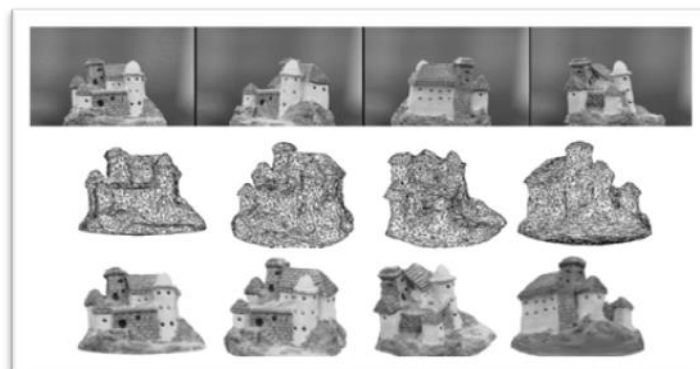


Figure 61 : Exemples d’images originales et de résultats de modélisation (Château).

**Tableau [2] : Analyse de temps d’exécution de chaque étape du processus**

Étape		Temps d’exécution (sec)	
		Château	Ecureuil
Extraction des silhouettes		53	31
Shape Modeling	Shape-from-silhouette	93	36
	Voting-based stereo	1230	1870
	Surface representation	106	38
Acquisition de texture		43	43

<sup>25</sup> **Image à fil de fer** : C’est un modèle de modélisation, où un objet est représenté par un ensemble de segments, de courbes ou droites.

## Avantages et inconvénients :

- Les principaux avantages de cet algorithme sont les suivants :
  - Traitement plus rapide que la simple reconstruction à partir de stéréovision.
  - Reconstruction de forme plus précise que la forme simple de l'enveloppe visuelle.
  - Meilleure réduction du bruit que la simple stéréovision.
- L'un des problèmes actuels de l'algorithme est que parfois la surface d'une partie mince peut être manquante en raison d'erreurs dans les surfaces partielles. Cette erreur peut être détectée en vérifiant la cohérence entre deux silhouettes : une générée à partir du modèle électronique et l'autre de la zone d'objet probable estimée par Shape-from-Silhouette.

## 5. Conclusion

Dans ce chapitre nous avons présenté quelques travaux qui utilisent l'idée de l'hybridation de la méthode de la stéréovision avec l'enveloppe visuelle. Pour chaque travail nous avons expliqué la démarche suivie pour réaliser chaque méthode et le principe de combinaison, nous avons présenté aussi les résultats obtenus par chaque travail ainsi que les avantages et les lacunes de chacun.

Pour le travail d'Éric Nowark et al, nous avons constaté que malgré que l'approche soit limitée à type particulier d'objet, mais nous constatons qu'il manque d'optimisation et ne donne pas des formes 3D robustes. De plus, la méthode d'extraction des silhouettes utilisée dans cette approche est assez compliquée, ainsi que la forme de silhouette obtenue qui contient d'énormes trous et de bruit, ce qui aboutit par la suite à une enveloppe visuelle non complète. Plusieurs parties sont mal présentées comme par exemple la partie du bras plié, ceci est dû à la mauvaise méthode choisie pour le calcul des silhouettes, ce qui se reflète directement sur le résultat final obtenu. Un autre problème est signalé ; c'est la déformation du maillage d'objet 3D à partir des résultats obtenus par les deux méthodes. Donc la technique utilisée pour la combinaison des deux méthodes reste encore moins efficace.

Pour le travail de *Ming Li et al* une architecture client/serveur est utilisée dont le but est de diviser les différents calculs sur un ensemble de machines, ceci permet d'améliorer le temps d'exécution nécessaire pour la réalisation de travail. Mais comme un inconvénient de cette approche, malgré l'utilisation des bibliothèques optimisées et spécifiées pour le traitement d'images, on constate que les résultats obtenus sont inexacts et elles semblent un peu loin de la réalité. En plus on constate que le coût de traitement est très élevé malgré la diffusion de travail sur plusieurs machines clients.

Enfin l'approche proposée par *Matsumoto et al* est appliquée seulement sur des images synthétisées. L'algorithme n'est pas vérifié sur des images réelles qui sont généralement plus complexes et s'influencent directement par le monde extérieur. Mais d'une façon générale les structures 3D résultantes sont presque identiques à l'objet d'intérêt.

Dans le chapitre suivant nous allons présenter notre proposition qui traite la même idée de l'hybridation enveloppe visuelle/Stéréovision, mais d'une façon complètement différente à celles présentées dans ce chapitre.

# Chapitre V

---

## ***Réalité augmentée***

---

## **Introduction :**

La réalité augmentée (AR) est l'un des domaines technologiques les plus à jour et les plus avancés. Le domaine de l'application de la Réalité augmentée a commencé dans les années 1990 (A.Rauschnabel, Felix, Hinsch, Shahab, & Alt, 2022) et son importance n'a cessé de croître depuis, parallèlement aux progrès technologiques et au développement des technologies de l'information et de la communication (TIC). Alors que la réalité virtuelle (VR) se réfère à la création d'un monde numérique qui reproduit un environnement qui existe déjà dans ou issu d'imagination, La réalité augmentée fournit un cadre pour ajouter des informations à l'environnement du monde réel. La technologie de la réalité augmentée permet aux utilisateurs de mieux percevoir le monde ; elle améliore l'expérience des utilisateurs en superposant de l'information générée par ordinateur, y compris des graphiques, des sons et parfois des commentaires tactiles, sur l'environnement réel (Aggarwal, 2019). Ainsi, de nouvelles métaphores de l'interaction avec le monde physique et l'information synthétique sur l'environnement doivent être inventées, testées et appliquées.

Les techniques de réalité augmentée peuvent être classées en deux grandes catégories : La réalité augmentée fondée sur la vision et la réalité augmentée fondée sur la localisation [ (Pang, Yuan, Nee, Ong, & Youcef-Toumi, 2006), (Sato, Fukuda, Yabuki, Michikawa, & Motamedi, 2016)]. La réalité augmentée basée sur la localisation utilise la capacité d'un dispositif spécifique pour déterminer sa position dans le monde, par exemple, GPS, puis la récupération de l'information pertinente à cet endroit. Ensuite, cette information est superposée à la sortie de la caméra de leur appareil pour permettre une présentation des données plus naturelle par rapport à l'utilisation de la carte seule. La réalité augmentée basée sur la vision dépend en particulier du traitement des données extraites des images ou des images vidéo prises par l'appareil. Ce type de réalité augmentée comprend un certain nombre de techniques qui se prêtent de façon significative de la vision par ordinateur à la portée, où les progrès de la recherche en réalité augmentée reposent sur les progrès de ces dernières (Gao, Wan, Tang, & Chen, 2017).

La réalité augmentée a été appliquée dans divers domaines, de la médecine à l'éducation, l'industrie automobile, la santé ou le tourisme (Arena, Collotta, Pau, & Termine, 2022). En particulier, il s'est avéré utile pour la gestion et la préservation du patrimoine culturel (HC). La réalité augmentée est utilisée dans l'HC à certaines fins principales : améliorer l'expérience des visiteurs, la reconstruction et l'exploration (Okanovic, et al., 2022), ainsi que la conservation et la préservation (Merchán, Merchán, & Pérez, 2021) et donner vie aux événements passés (Boboc, et al., 2019). La technologie la plus souvent utilisée dans tous ces cas est la numérisation. Il est utilisé comme méthode complémentaire pour préserver les actifs de l'héritage culturel et améliorer les procédures de conservation traditionnelles (Siti Zuliana & Abd Razak, 2022).

Une fois numérisé, le patrimoine culturel est prêt à être consulté par tous les bénéficiaires, scientifiques et non scientifiques. En commençant par les non-scientifiques, le grand public est impatient de consommer cette information disponible sous forme numérique. Afin de maximiser ce flux du patrimoine culturel vers le consommateur ordinaire, il est nécessaire de maximiser l'absorption des connaissances patrimoniales par les utilisateurs grâce à une telle visualisation.

Cela signifie que les visualisations doivent être adaptables, interactives et orientées vers l'utilisateur. Les scientifiques doivent comprendre profondément les informations précieuses qu'ils obtiennent en utilisant des techniques de numérisation à plusieurs niveaux. Ils doivent ainsi s'efforcer pour enregistrer tout ce qui concerne les artefacts. Toutes ces métadonnées doivent être enregistrées afin d'être facilement utilisées par la technologie qui sera disponible à l'avenir (Poulopoulos & Wallace, 2022). Nous considérons que toucher interactivement le passé à travers les artefacts numériques, tout en écoutant son récit et en observant sa réaction aux actions des utilisateurs, est l'une des formes ultimes d'immersion culturelle qui peut être offerte à l'utilisateur.

D'autre part, la numérisation est une chance de survie pour le patrimoine culturel mondial sous une double menace, celle des cataclysmes naturels et celle induite par des actions criminelles humaines directes. Une technique de numérisation idéale permettra de reproduire parfaitement un artefact numérisé. De nos jours, les résultats de la numérisation sont largement visualisés à l'aide de différentes technologies, comme la réalité virtuelle et la réalité augmentée, et sont manipulés par l'utilisateur dans des configurations supervisées à l'aide de métaphores gestuelles, tactiles ou sans contact. À l'aide des premières étapes de la technologie d'impression 3D, des maquettes à grande échelle d'artefacts numérisés sont actuellement accessibles au grand public pour manipulation gratuite et directe. Sans la crainte de la destruction et en offrant une expérience individuelle personnalisée et des connaissances sur ce que l'objet était et comment il a été utilisé (Chatzigrigoriou, Nikolakopoulou, Vakkas, Vosinakis, & Koutsabasis, 2021).

Les outils basés sur l'AR/VR sont devenus des éléments de notre vie quotidienne et sont là pour rester pour la prochaine décennie. La réalité augmentée est en fait un domaine de recherche dynamique, et il est nécessaire de souligner l'état de la recherche et de fournir les dernières tendances qui permettront le développement d'études futures. Les outils de Réalité augmentée facilitent l'accès au patrimoine culturel de manière interactive et engageante (Van Ruymbeke, Nofal, & Billen, 2022). Le présent chapitre vise à mener une étude concluante sur les efforts déployés au cours des dix dernières années pour le développement des applications de Réalité augmentée utilisées pour la préservation et la promotion de l'héritage culturel. Nous analysons comment l'AR est impliqué dans les applications du patrimoine culturel en menant des recherches bibliométriques sur les documents publiés sur les applications de l'AR en HC au cours des dix dernières années.

Nous examinons la documentation pertinente afin d'examiner les questions de recherche suivantes : Q1. Quelles sont les tendances de l'utilisation de la Réalité augmentée pour les applications du patrimoine culturel ? Q2. Quelle est la structure de la dernière décennie de la littérature ? Q3. Quels sont les principaux résultats de la recherche ?

Dans ce chapitre, nous avons présenté une enquête sur l'AR. Cette enquête est basée sur l'état de l'art lié au patrimoine culturel en AR. La section 1 présente l'AR et la section 2 est attribuée à ces types d'applications, tandis que la section 3 en explique le principe de fonctionnement. Dans la section 4, les composantes communes de l'AR sont présentées. Quelques avantages sont mentionnés dans la section 5. La section 6 énumère quelques applications de l'AR et la section 7 présente les notions de patrimoine culturel / héritage culturel. Les recherches liées au patrimoine culturel en réalité augmentée sont regroupés dans la section 8. Le travail est conclu par la section 9 qui est la conclusion.

## 1. Réalité augmentée :

La réalité augmentée est considérée comme l'une des technologies modernes qui mélange les objets virtuels dans le monde réel. La réalité augmentée (AR) peut simplement être définie comme une vision directe ou indirecte d'un environnement physique et réel. Des éléments sont complétés en temps réel par du contenu virtuel généré par ordinateur, par exemple, des données sonores, vidéo, graphiques ou GPS [ (Azuma, et al., 2001), (Shetty, Lasrado, & Mohammed, 2015)]. La réalité augmentée est considérée comme un domaine dans lequel les objets virtuels 3D sont complètement intégrés dans un environnement réel 3D en temps réel. Un environnement Réalité augmentée complète le monde réel avec des objets virtuels, qui sont générés en utilisant un ordinateur et qui semblent coexister dans le même espace que le monde réel (Sandip & Bharat, 2014). En d'autres termes, la réalité augmentée peut être définie comme l'interactivité des humains avec des objets virtuels situés dans l'environnement réel afin d'aider l'utilisateur à exécuter une tâche dans un environnement physique. La réalité augmentée est l'une des formes significatives de la réalité mixte (MR), dans lequel les objets réels et virtuels sont mélangés et montrés en un seul affichage en même temps et au même endroit comme le montre la figure 63. La réalité augmentée semble être de la fiction parce qu'elle crée des interfaces interactives qui créent l'illusion que les mondes physiques et virtuels sont connectés ensemble et que les utilisateurs peuvent physiquement passer de l'un à l'autre (Koleva, Schnädelbach, Benford, & Greenhalgh, 2000).

En termes simples, la réalité augmentée est une technologie qui augmente le monde réel en y apposant des couches d'éléments numériques. Ces éléments comprennent des graphiques générés par ordinateur, des effets sonores ou vidéo, la rétroaction haptique ou les projets sensoriels. L'intention derrière l'ajout de cette information numérique est de fournir une expérience client engageante et dynamique qui est activée avec l'entrée reçue de matériel varié comme les lunettes intelligentes, les lentilles intelligentes et les smartphones.

La réalité augmentée (AR) est souvent confondue avec la réalité virtuelle (VR). La principale différence entre les deux est que si la réalité virtuelle remplace tout l'environnement réel par un environnement artificiel, la réalité augmentée est appliquée dans une vue directe d'un environnement réel existant et y ajoute des éléments comme des sons, des vidéos ou des graphiques.



Figure 62 : Positionnement de l'AR et la VR

Le terme de réalité augmentée a été inventé en 1990 par le chercheur chez Boeing Tom Codell et l'une des premières utilisations commerciales de cette technologie était dans la télévision et l'armée. Au fur et à mesure que le monde évoluait vers une approche davantage axée sur la technologie, la Réalité augmentée est devenue de plus en plus importante dans de multiples domaines, déployant sa deuxième vague et établissant son lien avec le concept interactif. À

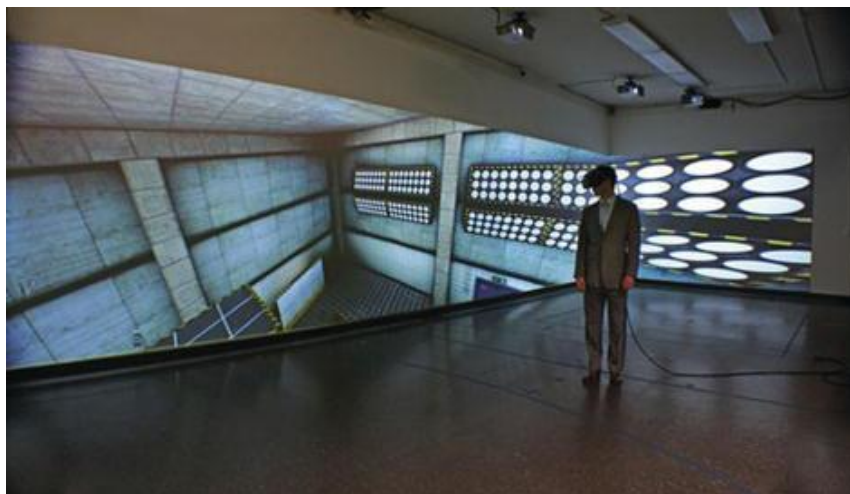


mesure que la technologie se développe, plusieurs nouvelles tendances émergent en réalité augmentée.

La réalité augmentée ne remplace pas la réalité virtuelle (VR), mais elle complète l'environnement réel par des informations numériques, des graphiques virtuels et/ou des objets virtuels générés par ordinateur, comme le montre la figure 62. Toutefois, les utilisateurs naviguent dans la réalité virtuelle en utilisant un environnement simulé ou imaginaire appelé environnement virtuel. Dans cet environnement, tous les sens des utilisateurs sont contrôlés à l'aide d'un ordinateur et immergés dans un environnement simulé (Sites & Elvet, 2013) comme le montre la figure 63.



*Figure 63 : Une vue d'environnement réelle est complétée par des informations numériques.*



*Figure 64 : Un utilisateur navigant dans un environnement virtuel VR.*

## **2. Technologie derrière le fonctionnement de la réalité augmentée :**

En premier lieu, des images d'environnements réels doivent être acquises. Deuxièmement, il faut utiliser une technologie qui permet de superposer des images 3D sur des images d'objets du monde réel. La troisième étape consiste à l'utilisation de la technologie pour permettre aux utilisateurs d'interagir avec les environnements simulés.

La réalité augmentée peut être affichée sur des écrans, des lunettes, des appareils portatifs, des téléphones mobiles et des écrans montés sur la tête.

Elle utilise la technologie de localisation et cartographie simultanées S.L.A.M. (Simultaneous Localisation and Mapping), et la technologie de suivi de profondeurs pour calculer la distance à l'objet à l'aide de données issues de capteurs, en plus d'autres technologies.

La technologie de réalité augmentée permet une augmentation en temps réel dans le contexte de l'environnement. Des animations, des images, des vidéos et des modèles 3D peuvent être utilisés et les utilisateurs peuvent voir des objets en lumière naturelle ou synthétique.

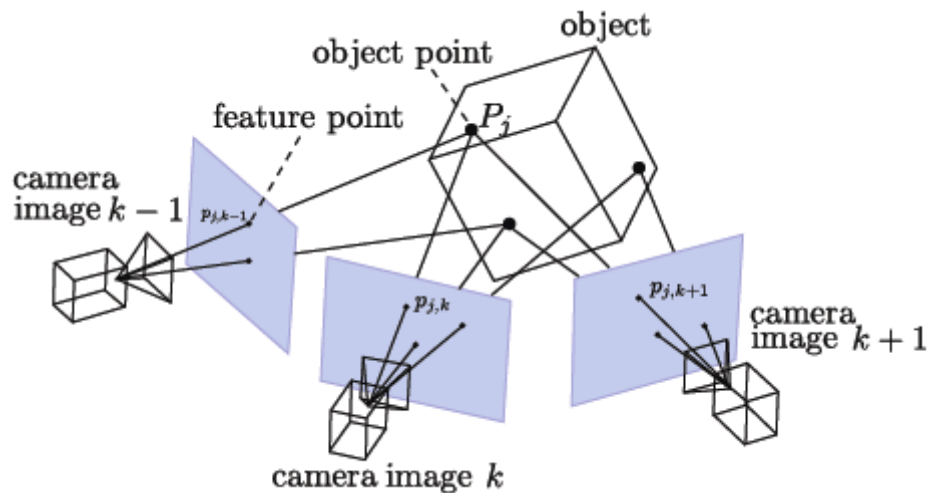


Figure 65 : Principe de la réalité augmentée.

## S.L.A.M

Est un ensemble d'algorithmes qui résolvent des problèmes de localisation et de cartographie simultanée, il utilise des points de fonctionnalité pour aider les utilisateurs à comprendre le monde physique. La technologie permet aux applications de manipuler les objets et les scènes 3D. Il permet aussi de suivre le monde physique instantanément tout en permettant également la superposition de simulations numériques.

SLAM utilise un capteur mobile comme pour détecter l'environnement, puis créer une carte virtuelle ; trace sa position, sa direction et son chemin sur cette carte. Outre la AR, il est utilisé sur les drones, les véhicules aériens, les véhicules sans pilote et les robots nettoyeurs. L'algorithme utilise l'intelligence artificielle et l'apprentissage automatique pour comprendre les emplacements.

Dans AR, SLAM aide à insérer et à fusionner l'objet virtuel en un objet réel. La détection et l'appariement des caractéristiques sont effectués à l'aide de caméras et de capteurs qui recueillent des points d'intérêt de divers points de vue. La technique de triangulation déduit ensuite l'emplacement en trois dimensions de l'objet.

### **3. Types de réalité augmentée :**

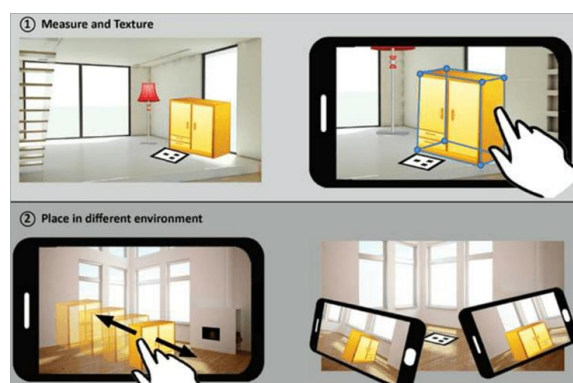
La réalité augmentée est de six types : sans marqueurs, basée sur marqueurs, basée sur projection, basée sur la superposition, fondée sur l'emplacement et celle fondée sur la technologie de suivi de profondeur.

#### **3.1 AR à base de marqueurs (reconnaissance)**

Un marqueur, qui est un objet visuel spécial comme un signe spécial ou quoi que ce soit, et une caméra sont utilisés pour lancer les animations numériques 3D. Le système calculera l'orientation et la position du marqueur pour positionner efficacement le contenu.

Ce type de réalité augmentée utilise des marqueurs (un écran prédéfini bidimensionnel) qui sont placés dans la scène et dans le champ de vision de la caméra afin d'aider à guider le processus d'estimation de la pose de la caméra (Shetty, Lasrado, & Mohammed, 2015). Les marqueurs sont souvent indiqués comme fiduciaires parce que leur position et leur orientation par rapport au paysage sont stables. Les marqueurs sont toujours planaires et ont généralement des caractéristiques puissantes, par exemple, les bords longs, ainsi que les coins entre les régions noires et blanches. Dans cette technique, l'AR met un accent puissant sur la conception du marqueur. L'un des types les plus courants de conception de marqueur est carré parce que la caractéristique du carré permettra la localisation précise des marqueurs en utilisant ses quatre points d'angle (Gherghina, Olteanu, & Tapus, 2013). La réalité augmentée basée sur les marqueurs utilise des techniques de vision par ordinateur afin de calculer la position et l'orientation de la caméra par rapport au marqueur. Les objets 3D virtuels peuvent être superposés avec précision sur les marqueurs comme le montre la figure 66. Il a un principe de fonctionnement principal : capturer l'entrée vidéo de la caméra, ajouter des graphiques 3D à la scène et montrer les cadres augmentés sous forme de flux vidéo (Patirupanusara, 2012).

Ce type d'algorithmes, utilise une caméra pour identifier les marqueurs de sorte qu'une superposition est possible si un marqueur est détecté. L'appareil détecte et calcule la position et l'orientation du marqueur et remplace le marqueur du monde réel par sa version 3D. Il calcule ensuite la position et l'orientation des autres. La rotation du marqueur fait tourner l'objet entier.



*Figure 66 : Une application mobile d'ameublement Réalité augmentée basée sur marqueurs.*

#### **3.2 AR sans marqueurs**

L'exemple ci-dessous montre qu'une réalité augmentée sans marqueur n'a pas besoin de marqueurs physiques pour placer des objets dans un espace réel.



Figure 67 : Exemple de réalité augmentée sans marqueurs.

### 3.3 AR par projection

Ce type utilise de la lumière synthétique pour la projetée sur les surfaces physiques pour détecter l'interaction de l'utilisateur avec les surfaces. Un exemple commun est celui des hologrammes.



Figure 68 : Un système d'affichage par projection (hologramme).

### 3.4 AR basée sur la superposition

Dans ce cas, l'article original est remplacé - entièrement ou partiellement - par une augmentation. L'exemple ci-dessous permet aux utilisateurs de placer un meuble virtuel sur une image de pièce avec une échelle.

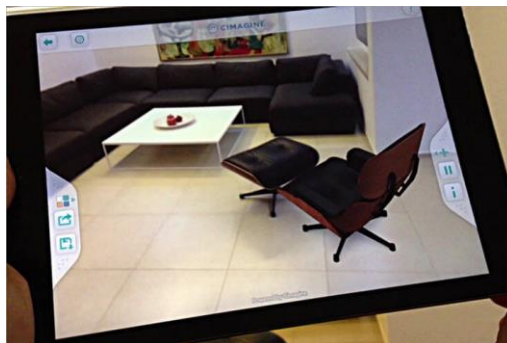


Figure 69 : Exemple de réalité augmentée basée sur la superposition.

### 3.5 Approche fondée sur l'emplacement :

Ici, les simulations ou visualisations sont générées à partir de données recueillies par GPS, boussoles numériques, accéléromètres et vélocimètres. Elle est très commune dans les smartphones.

### 3.6 Technologie de suivi de profondeur :

Les caméras de suivi de profondeur telles que Microsoft Kinect génèrent une carte de profondeur en temps réel en utilisant différentes technologies pour calculer la distance en temps réel des objets dans la zone de suivi de la caméra. Les technologies isolent un objet de la carte de profondeur générale et l'analyse.

## 4. Composantes communes de la AR

Il y a plusieurs composants qui travaillent ensemble pour offrir une expérience de réalité augmentée immersive.

- **Les caméras et les capteurs** sont utilisés pour capturer l'environnement physique ou réel. Cela peut également inclure des capteurs de profondeur et de lumière, des accéléromètres et des gyroscopes. Ils sont utilisés pour mesurer la distance, la direction, la vitesse, l'angle, etc.
- **Logiciel** comme une application pour smartphone. L'information recueillie par les caméras et/ou les capteurs est ensuite transmise au logiciel pour traitement. Ce logiciel rend également la Réalité augmentée accessible à l'utilisateur.
- **Un système d'exploitation** pour alimenter le processus AR.
- **Un jumeau numérique** peut être utilisé pour créer une version numérique 3D de l'objet dans l'espace et pour combler le fossé entre les mondes réel et virtuel.
- **Objectif ou plate-forme de visualisation** pour que vous puissiez voir le contenu augmenté étendu sur votre environnement physique.
- **L'intelligence artificielle (IA)** peut aussi être utilisée pour offrir une expérience plus immersive avec la réalité augmentée - permettant aux utilisateurs de manipuler des objets par la voix ou le toucher, par exemple.

## 5. Avantages de la réalité augmentée

Bien sûr, comprendre la Réalité augmentée (de la façon de l'utiliser à la façon de la développer) commence par l'éducation. La technologie immersive peut être utilisée partout, des environnements élémentaires aux environnements supérieurs. Les simulations immersives de la vie réelle offrent des avantages pédagogiques aux apprenants. Les simulations d'apprentissage et de formation basées sur le jeu apportent des avantages psychologiques et augmentent l'empathie chez les apprenants, comme le montrent les chercheurs. Une foule agréable pour l'ensemble élémentaire est l'application Dinosaur 4D+. Les élèves de tous âges peuvent en apprendre davantage sur les animaux préhistoriques grâce à une vue à 360 degrés et les cartes-éclair Dinosours4D+ qui les accompagnent. L'apprentissage à distance ne disparaîtra pas, et les éducateurs peuvent utiliser l'apprentissage à distance pour améliorer l'expérience d'apprentissage de leurs élèves, peu importe où ils se trouvent. Avec CoSpaces, les éducateurs peuvent créer du contenu multimédia interactif pour les plans de leçon existants

Les vendeurs peuvent produire et publier du contenu de réalité augmentée de marque intéressante et y insérer des annonces afin que les gens puissent apprendre à connaître leurs produits lorsqu'ils regardent le contenu. La réalité augmentée améliore l'engagement.

Dans la fabrication, les marqueurs Réalité augmentée sur les images des équipements de fabrication aident les chefs de projet à surveiller le travail à distance. Il réduit la nécessité d'utiliser des cartes numériques. Par exemple, un appareil ou une machine peut être pointé sur l'emplacement pour déterminer s'il s'adaptera à la position.

Les étudiants en médecine peuvent utiliser les simulations de réalité augmentée et de réalité virtuelle pour essayer virtuellement des chirurgies difficiles sans gros budgets ou blessures inutiles aux patients, le tout avec une immersion et des expériences presque réelles.

La réalité augmentée permet de mettre en évidence le tourisme virtuel. Les applications, par exemple, peuvent fournir des indications sur les destinations souhaitables, traduire les panneaux dans la rue et fournir des informations sur la visite. Un bon exemple est une application de navigation GPS.

L'une des principales industries pour l'avenir de la Réalité augmentée est le tourisme et les voyages. Le contenu de la Réalité augmentée permet la production de nouvelles expériences culturelles, par exemple, où une réalité supplémentaire est ajoutée aux musées. Les hôtels et les entreprises touristiques donnent aux voyageurs un avant-goût de ce à quoi pourrait ressembler leur expérience grâce à des éléments Réalité augmentée interactifs.

Grâce à la AR, les futurs astronautes peuvent essayer virtuellement leur première ou leur prochaine mission spatiale.

Le développement d'applications de marque propres est l'un des moyens les plus courants que les entreprises utilisent pour s'engager avec la technologie de AR. Les entreprises peuvent toujours placer des annonces sur des plateformes et du contenu Réalité augmentée tiers, acheter des licences sur des logiciels développés, ou louer des espaces pour leur contenu de AR.

Les développeurs peuvent utiliser des plates-formes de développement de la réalité augmentée comme webAR, ARKit et ARCore pour développer des applications et intégrer la réalité augmentée dans des applications commerciales.

## **6. Exemples d'applications de la réalité augmentée :**

L'exemple le plus connu est celui de Pokémon GO, mais il y a beaucoup d'applications de Réalité augmentée qui valent d'être mentionnées :

- Snapchat : Si vous n'avez jamais essayé cette application de médias sociaux - ou cela fait un moment que vous ne lui avez pas donné un spin - vous manquez ses fonctionnalités Réalité augmentée et IA. Vous pouvez tout faire, de mettre des lunettes en forme d'œil de chat à essayer les dernières baskets Puma.
- HoloLens : Pour l'expérience Réalité augmentée d'entreprise, HoloLens est là où il est. Le dispositif holographique de Microsoft est intégré aux applications d'entreprise afin que vous puissiez accomplir des tâches, collaborer et innover avec plus de précision.

- Just a Line : Vous pouvez faire des dessins simples sur un environnement physique avec juste votre doigt et ensuite partager avec des amis à travers la vidéo. Vous pouvez également collaborer avec d'autres gribouilleurs en temps réel.
- Mission vers Mars : L'application Mission vers Mars est vraiment hors de ce monde. Grâce à la AR, vous pouvez explorer et en apprendre davantage sur la planète rouge, observer un lancement de fusée et marcher sur la quatrième planète depuis le Soleil.
- IKEA Place : IKEA pourrait vous demander d'assembler vos propres meubles, mais ils vous aideront à imaginer à quoi ressemblent leurs pièces dans votre propre maison. Grâce à ce que le détaillant de forfaits plats appelle des « expériences axées sur la caméra », vous pouvez meubler vos espaces avec un simple clic.

## **7 Patrimoine virtuel, Héritage culturel virtuel**

Le patrimoine virtuel en Réalité augmentée peut être défini comme une technologie informatique interactive, qui peut être utilisée pour réaliser une reconstruction visuelle, aider les érudits et les éducateurs des entités traditionnelles, par exemple, les bâtiments, les artefacts et la culture (Frank, Bernhard, & Benjamin, 2000). Cette technologie est utilisée - par exemple - pour maintenir les bâtiments historiques fragiles contre les catastrophes naturelles et le sabotage (Amparo, Felipe Buill, Joaquín Regot, & Andrés de Mesa, 2012). Afin de créer un héritage virtuel, il y a sept principaux principes de conception, qui doivent être pris en compte tels que : la haute précision géométrique, le haut niveau de capture automatique pour tous les détails, le faible coût, le photoréalisme, la flexibilité, la portabilité et l'efficacité de la taille du modèle (El-Hakim, Beraldin, Picard, & Godin, 2004). Les couches du patrimoine culturel sont proposées pour visualiser des supports historiques tels que des dessins, des peintures et des photographies de bâtiments et de scènes historiques superposés en toute transparence sur l'environnement réel à travers la vidéo voir à travers la 3D (Zoellner, Keil, Drevensek, & Wuest, 2009). Ces couches locales sont des contextes qui racontent l'histoire du lieu et donnent l'impression d'un voyage dans le temps virtuel.

Bres et al. ont proposé une réalité augmentée pour le tourisme historique utilisant des appareils mobiles (Stéphane & Bruno, 2009). Le cœur du système proposé est lié à une solution de réalité augmentée extérieure sans marqueur. Cette technique est basée sur les fonctions de transformation des fonctions invariantes d'échelle (SIFT) pour la localisation et l'intégration des modèles 3D en vidéo. Ces caractéristiques sont utilisées pour projeter un modèle numérique des façades du bâtiment de la place afin d'obtenir des coordonnées 3D pour chaque point caractéristique. Les algorithmes exécutés sont chargés de calculer la pose de la caméra pour le cadre d'une vidéo à partir de correspondances de points 3D-2D parmi les caractéristiques extraites du cadre vidéo actuel et les points dans l'ensemble de données de référence. Les algorithmes ont été évalués avec succès sur des films vidéo de places de ville. Bien qu'ils ne fonctionnent pas encore en temps réel, ils sont en mesure de corriger l'estimation de la pose et la projection de données artificielles dans la scène. La recherche a montré que la possibilité des fonctionnalités SIFT sont purement utilisées pour des applications de réalité augmentée extérieure sans marqueur (Stéphane & Bruno, 2009). Cette recherche présentait une application mobile simple qui servait à augmenter une petite image 3D. HeladivaAR (Galmangoda, Gajanayake, Indika, Jayaweera, & Raj, 2016) a proposé de reconstruire le patrimoine historique et culturel du Sri Lanka. HeladivaAR est une application de téléphone mobile qui servait à montrer un modèle 3D reconstruit de ces ruines anciennes comme elles étaient dans leur état

initial. En plus de l'utilisation de la technologie AR, l'application a utilisé la caméra du téléphone mobile pour déterminer et suivre les ruines restantes du lieu historique et reconstruire le modèle 3D sur elle.

Cette application a utilisé différents aspects pour reconstruire le bâtiment du patrimoine culturel tels que le traitement d'images, la modélisation 3D, l'identification par marqueur à l'aide de la plate-forme Android, des livres historiques et la reconstruction de sites en ruine. En utilisant l'AR, la scène réelle est renforcée par des informations multimédias interactives afin d'augmenter l'expérience de l'utilisateur, qui peut récupérer ces informations par une interface utilisateur-facile à travers leur téléphone mobile. Dans le domaine de l'éducation, le patrimoine virtuel devient une plateforme d'apprentissage, de motivation et de compréhension d'événements particuliers et d'éléments historiques pour les étudiants et les chercheurs. Cette recherche permet de mieux comprendre le patrimoine culturel sri-lankais et d'acquérir des connaissances interactives sur les faits archéologiques des anciens royaumes (Galmangoda, Gajanayake, Indika, Jayaweera, & Raj, 2016). Cependant, cette recherche a plusieurs limites. La première est que l'application ne peut s'appliquer qu'aux appareils de réalité augmentée basés sur Android ; elle ne peut pas s'appliquer aux appareils du système d'exploitation basés sur IOS. La deuxième limite est la qualité de l'application basée sur l'appareil mobile, car il ne s'agit pas d'une application de bureau. La dernière limitation est l'application développée pour Android 3.0 ou supérieur. Les versions ci-dessous peuvent rencontrer des problèmes de rendu lors de l'exécution.

Indrawan a développé la réalité augmentée sans marqueur en utilisant le gyroscope pour démontrer la position de Dewata Nawa Sanga (Indrawan, a Bayupati, & Putri, 2018). Cette application est conçue pour apprendre, comprendre et reconnaître les propriétés de Dewata Nao Sanga en utilisant un gyroscope. Le capteur travaille pour atteindre l'objet des divinités dans les coordonnées à identifier, ainsi, elle est réalisée pour fournir des informations sur Dewata Nawa Sanga par des animations 3D informatives. Cette recherche évalue l'utilité, la fonctionnalité de l'application, en plus de l'impact de l'application Réalité augmentée Dewata Nawa Sanga qui peut motiver ses utilisateurs. Le résultat de la convivialité et de la valeur du questionnaire de satisfaction montre que la moyenne en pourcentage est de 84,8 %. Il illustre que l'application est très utile pour les participants d'avoir des connaissances sur Dewata Nawa Sanga ainsi que très satisfait d'utiliser (Indrawan, a Bayupati, & Putri, 2018).

Kolivand et El Rhalibi ont présenté une nouvelle technique pour augmenter un bâtiment virtuel réaliste dans des environnements réels à observer en direct à travers une caméra Réalité augmentée (Hoshang, Abdennour, Mohd, & Tanzila, 2018). Il y a quelques composants extérieurs quand un bâtiment réaliste est augmenté, par exemple, la position du soleil, les ombres, l'illumination du ciel et les personnages animés traditionnels virtuels. L'augmentation est réalisée sur des environnements réels à la place de bâtiments historiques réels, ou des emplacements souhaitables, à différents moments de la journée et différents jours de l'année (Hoshang & Mohd, 2013). Les auteurs ont présenté de nouvelles idées dans le cas du patrimoine virtuel. Tout d'abord, il s'agit de modéliser le modèle 3D du Malacca portugais. Un système structuré en temps réel est fourni pour tracer la position du soleil, en utilisant la datation Julian, et le modèle de ciel Perez est utilisé pour modéliser la couleur du ciel, les auteurs ont présenté une nouvelle technique pour créer l'éclairage extérieur. Un algorithme d'ombre semi-douce a été mis en œuvre pour soutenir le réalisme des systèmes extérieurs de réalité augmentée. Un système de configuration de caméra simple a été utilisé pour présenter



Réalité augmentée sans marqueur. Le système final peut être installé sur un écran monté sur la tête (HMD) ou dans le dispositif proposé appelé ReVitAge pour montrer les bâtiments patrimoniaux virtuels reconstitués réalistes, en tenant compte des principaux composants de l'éclairage extérieur (Hoshang, Abdenmour, Mohd, & Tanzila, 2018).

## **7. Travaux connexes :**

Cette section est au cœur de notre examen, mettant l'accent sur les principales applications de la réalité augmentée (AR) dans le contexte du patrimoine culturel. Plusieurs recherches intéressantes ont été établies [ (Augello, Infantino, Pilato, & Vitale, 2021), (Lee, et al., 2021), (Parfenov, Igoshin, Masaylo, Orlov, & Kuliashou, 2022)], mais nous trouvons que la plus établie et diversifiée est la recherche (Boboc, Elena, Florin, Norina, & Dorin-Mircea, 2022). En fait, cette étude couvre un large éventail d'application de réalité augmentée dans le contexte de l'héritage culturel étalées sur les 10 dernières années, ayant fait l'analyse de plus de 1200 articles universitaires et qui utilise comme piliers principaux huit des mots-clés / thèmes de recherche les plus utilisés dans la production scientifique relatives au domaine, à savoir « **la reconstruction 3D** », « **le patrimoine numérique** », « **le musée virtuel** », « **l'expérience utilisateur** », « **l'éducation** », « **le tourisme** », « **les jeux** » et « **le patrimoine culturel immatériel** ».

Au cours de la dernière décennie, les technologies de Réalité augmentée ont connu une croissance importante. Il y a eu des progrès constants dans le développement d'applications de Réalité augmentée pour l'héritage culturel (HC), ce qui a mené au développement d'applications attrayantes qui ont un plus grand impact sur plusieurs sujets.

### **7.1 Reconstruction tridimensionnelle**

Les anciens artefacts contiennent des renseignements inestimables sur la culture et la civilisation d'une société. Cependant, ils sont souvent mal conservés ou inaccessibles en raison de leur taille et de leur fragilité. On estime qu'au cours des dix mille dernières années, plus de 90% du patrimoine culturel de l'humanité a été perdu (Cultural heritage at risk: United States, 2022). La fusion de la reconstruction 3D et de la Réalité augmentée pour le patrimoine culturel est en train de devenir une technologie émergente (Panou, Ragia, & Dimelli, 2018).

Les technologies de reconstruction tridimensionnelle présentent un grand potentiel pour la conservation numérique des artefacts d'héritage culturel [ (Campi, di Luggo, Palomba, & Palomba, 2019), (Machidon, Postelnicu, & Girbacia, 2016), (Noh, Sunar, & Pan, 2009)]. La recherche dans ce domaine ayant débuté avant 2010, une revue publiée en 2009 répertorie plusieurs projets liés à la Réalité augmentée et au patrimoine virtuel (Noh, Sunar, & Pan, 2009). Le processus de reconstruction 3D consiste en une séquence d'étapes (Gomes., Bellon., & Silva, 2014). La première étape fondamentale consiste à recueillir les données initiales à l'aide de plusieurs technologies, dont celles qui sont pertinentes pour la reconstruction des artefacts HC sont : la modélisation 3D photogrammétrique, la triangulation et le balayage laser en temps de vol, la lumière structurée, la stéréoscopie optique et le Lidar.

La photogrammétrie est l'une des méthodes les plus utilisées dans la modélisation 3D des artefacts HC [ (Portalés, Lerma, & Pérez, 2009), (Putra, Wahyudi, & Dumingan, 2016)]. Cependant, elle repose sur un nombre considérable d'images et un temps de calcul accru pour une reconstruction efficace. Il faut donc plus de temps pour converger. D'autre part, l'approche

de reconstruction basée sur la géométrie 3D dépend d'une caméra de détection de profondeur qui est souvent trop coûteuse pour de nombreuses applications. Des travaux récents proposent des méthodes pour obtenir des informations sur la profondeur à l'aide de lasers ou de lumière structurée (Fritsch & Klein, 2018).

Un scanner laser est une solution 3D de haute qualité qui permet de créer des modèles en balayant n'importe quelle surface - de la pierre au bois, en passant par le métal et bien d'autres - mais l'équipement matériel est très coûteux et un logiciel complexe est nécessaire pour traiter les données [ (Barrile, Bilotta, Meduri, De Carlo, & Nunnari, 2017), (Barrile, Nunnari, & Ponterio, Laser scanner for the Architectural and Cultural Heritage and Applications for the Dissemination of the 3D Model, 2016), (Scianna, Gaglio, Grima, & La Guardia, 2020)]. Le développement récent de la technologie Lidar et l'intégration avec une plateforme mobile permet une vitesse et une précision adéquates pour les grands sites CH (Rodríguez-González, Jimenez Fernandez-Palacios, Muñoz-Nieto, Arias-Sanchez, & Gonzalez-Aguilera, 2017).

Une image ne suffit pas, dans la plupart des cas, pour une représentation complète d'un modèle 3D (Gomes., Bellon., & Silva, 2014). Les images d'artefacts peuvent avoir plusieurs points de vue différents. L'étape suivante de la reconstruction 3D est l'alignement des informations de couleur et de profondeur. L'alignement des images de vue peut être réalisé en utilisant l'approche itérative du point le plus proche (Noh, Sunar, & Pan, 2009). Lorsque l'alignement est effectué, toutes les vues 3D générées sont combinées en un seul maillage. Dans la dernière étape, le modèle 3D est texturé pour une représentation réaliste de l'artefact CH numérique. La sélection de la solution optimale pour la reconstruction 3D est une tâche difficile.

Une méthode qui combine plusieurs techniques de restauration d'objets 3D pour les applications de Réalité augmentée est présentée dans (Van Nguyen, Le, Tran, & Tran, 2022). La solution comprend l'obtention du nuage de points 3D de l'artefact HC à l'aide d'un scanner 3D, le traitement et la reconstruction du modèle 3D à l'aide de Meshlab, et la construction de l'application de Réalité augmentée pour la visualisation de l'artefact HC restauré.

La réalité augmentée fournit une méthode qui permet d'expérimenter une reconstruction 3D d'un artefact HC (voir Figure 70) avec un niveau de détail élevé et une dimension à grande échelle de la perception de l'espace, sans exiger de l'utilisateur des connaissances spécialisées [ (Shih, Diao, & Chen, 2019), (Blanco-Pons, Carrión-Ruiz, Lerma, & Villaverde, 2019)]. Les modèles tridimensionnels combinés à la réalité augmentée permettent au grand public de mieux comprendre et expérimenter l'histoire par des moyens plus intuitifs et interactifs [ (Cruz, Sevilla, San Gabriel, Cruz, & Caselis, 2018), (Pietroni, 2012), (Sebastiani, 2021)]. De plus, la reconstruction 3D des artefacts HC peut être utilisée pour les répliques 3D qui impliquent la combinaison d'écrans tactiles numériques, d'applications de Réalité augmentée et d'impression 3D de haute qualité. Les sujets qui ont évalué la 3D en Réalité augmentée ont considéré l'utilisation de modèles ou d'impressions 3D physiques plus intuitives, principalement pour la compréhension contextuelle et les expériences intangibles (Malik, Tissen, & Vermeeren, 2021).



Figure 70 : Exemple d'application Réalité augmentée mobile pour l'exploration de la reconstruction 3D

La réalité augmentée est utilisée pour la restauration d'objets religieux détériorés (Girbacia, Butnariu, Orman, & Postelnicu, 2013) et peut aider à faire revivre ces objets importants. Il peut être utilisé pour identifier l'approche de restauration la plus appropriée pour produire une réplique réaliste, ce qui réduit les coûts et accélère le processus. En utilisant les technologies de reconstruction 3D, un modèle numérique de l'objet endommagé peut être créé (Parfenov, Igoshin, Masaylo, Orlov, & Kuliashou, 2022), qui est ensuite utilisé pour générer une application de Réalité augmentée (Abate, et al., 2018) qui aide à déterminer la meilleure approche de restauration (Blanco-Pons, Carrión-Ruiz, Lerma, & Villaverde, 2019). Le principal inconvénient de la technologie Réalité augmentée est lié à l'absence de lignes directrices appropriées sur la façon d'utiliser cette technologie. L'expérience limitée avec les applications de Réalité augmentée est un autre problème important qui limite l'adoption de cette technologie dans le processus de reconstruction.

## 7.2 Patrimoine numérique

Selon la définition du patrimoine numérique de l'UNESCO, « le patrimoine numérique est constitué de matériaux informatiques de valeur durable qui doivent être conservés pour les générations futures. Le patrimoine numérique émane de différentes communautés, industries, secteurs et régions. Ce ne sont pas tous les documents numériques qui ont une valeur durable, mais ceux qui le sont nécessitent des approches de préservation active si l'on veut maintenir la continuité du patrimoine numérique. » (UNESCO, 2022). À ce titre, nous nous intéressons aux applications qui englobent le patrimoine numérique et les technologies telles que la réalité augmentée, la réalité virtuelle ou la réalité mixte (RM). En ce qui suit, nous passons en revue certains des travaux influents à cet égard dans la suite.

Les sites patrimoniaux numériques bénéficient largement de l'avènement de plateformes et d'applications de médias sociaux qui intègrent les commentaires de ces plateformes (Balduini, et al., 2012). Balduini et al. ont analysé le trafic sur les réseaux sociaux (plus précisément les messages Twitter) à Séoul, en 2011, et découvert des messages lourds concernant divers points d'intérêt dans la ville, tels que les restaurants ou les cafés. À partir de là, ils ont proposé un système complet basé sur les technologies sémantiques qui comprend une application de Réalité augmentée qui recommande les emplacements, qui étaient auparavant classés en analysant les opinions des utilisateurs sur une période de trois ans (Balduini, et al., 2012). Une autre application qui démontre la valeur du patrimoine numérique dans le tourisme est présentée dans (Tan & Lim, 2017). Ici, une application de Réalité augmentée exploite le concept de jeux pour faciliter le processus d'apprentissage de l'histoire.

L'immersion du visiteur dans la représentation théâtrale est facilitée par la réalité augmentée en (Jacquemin, Caye, Luca, & Favre-Brun, 2014). Les auteurs apportent leur contribution sous la forme de modèles 3D d'une église historique, ainsi que dans le développement d'un système multimédia qui permet aux participants au spectacle d'être immergés dans l'espace virtuel. Le patrimoine numérique capitalisé au moyen d'applications de AR, appuyées par des récits sonores (Comunità, Gerino, Lim, & Picinali, 2021), aide à créer une expérience crédible (Jacquemin, Caye, Luca, & Favre-Brun, 2014).

Dans une autre étude pertinente au patrimoine numérique, la numérisation des biens culturels dans les musées en Chine est inspectée (Ch'ng, Cai, Leow, & Zhang, 2019). L'étude a porté sur 22 sites sélectionnés comme bancs d'essai pour des applications patrimoniales numériques, comme les systèmes AR, VR ou multitouches. Les auteurs ont étudié les facteurs qui influencent l'utilisation réussie des expositions numériques et les modèles de comportement des visiteurs afin d'obtenir une bonne compréhension de la façon dont l'utilisation de la culture numérique peut être améliorée en Chine.

### 7.3 Musées virtuels

Les musées qui veulent aller au-delà de leur expérience traditionnelle de fournir de l'information fondée sur des guides audio et des catalogues papier commencent à innover au niveau de l'expérience utilisateur (voir la figure 71), ce qui la rend plus intuitive (Laudazi & Boccaccini, 2014). Dans le même temps, ils adoptent des systèmes innovants capables de fournir le contenu de leurs informations tout en améliorant l'expérience du visiteur. L'expérience d'information du musée en Réalité augmentée permet aux visiteurs de découvrir les collections d'art [ (Partarakis, Zidianakis, Antona, & Stephanidis, 2015), (Teneketzis, 2020)], les sites archéologiques (Silva & Teixeira, 2020), ou les anciens bâtiments [ (Balduini, et al., 2012), (Kallergis, et al., 2020)] d'une manière simple et intuitive (Kyriakou & Hermon, 2019). Un autre avantage de l'utilisation des technologies de Réalité augmentée dans le musée est la capacité d'ajouter de façon dynamique une composante d'interactivité numérique à une exposition du musée. La composante interactive numérique peut être un jeu de réalité augmentée (Nisi, Cesario, & Nunes, 2019) qui engage un joueur comme un type d'interactivité numérique, ou il peut s'agir d'une fonctionnalité qui fournit une éducation améliorée [ (Lee, et al., 2021), (May, Schmidt, & Vlachopoulos, 2020)] ou un autre type de contenu lié à l'exposition physique (Partarakis, Antona, Zidianakis, & Stephanidis, 2016). Cela a pour effet d'accroître le niveau d'engagement du visiteur envers les artefacts et leur signification.

La nécessité d'élaborer des méthodes novatrices pour numériser les artefacts de l'habitat essentiel en contenu 3D facilement déployable devient une priorité à mesure que la Réalité augmentée gagne en importance dans le secteur de l'habitat essentiel. En outre, les flux de travail nécessaires pour transformer les artefacts CH en contenu 3D immersif sont à peine pris en charge. Dans (Rahaman, Champion, & Bekele, 2019), un flux de travail pratique pouvant être utilisé pour la reconstruction 3D à la visualisation Réalité augmentée des actifs CH a été développé, et il identifie et combine une variété d'outils et de services d'accès libre et personnalisés.

En utilisant les technologies de AR, les galeries, les bibliothèques, les archives et les musées peuvent ajouter de multiples expériences sensorielles qui peuvent intéresser les visiteurs avec du contenu en CH. Un tel exemple est le récit de la réalité alternée (AltR) qui a été introduit dans (Hoang & Cox, 2018). C'est un entrelacement de deux récits apparemment disparates, un

réel et un virtuel. Au cours de cette expérience en temps réel, l'utilisateur visualise et interagit avec le récit virtuel. La table basse du Musée (Teneketzis, 2020) est une surface physique qui a été augmentée afin que des objets réels puissent être utilisés pour accéder aux données sur les artistes et leurs œuvres. En utilisant ce système, toute la famille peut se réunir autour de la table, prendre un café et améliorer leur visite du musée en apprenant de nouvelles informations et en s'amusant.

Une étude sur la façon dont le contenu narratif numérique est utilisé pour communiquer de l'information sur les sentiers patrimoniaux au moyen d'applications mobiles est présentée dans (Basaraba, Conlan, Edmond, & Arnds, 2019). En créant une liste des fonctionnalités les plus fréquemment utilisées dans les applications mobiles réussies, il vise à rendre les applications de sentiers patrimoniaux plus accessibles et utilisables.

Les problèmes d'accessibilité dans les musées et les expositions exigent une attention particulière. Les solutions actuelles du guide du musée ne facilitent pas l'expérience du contenu en HC pour les personnes ayant une perte auditive ou d'autres handicaps. La technologie de la réalité augmentée offre aux personnes ayant une perte auditive une façon immersive de participer aux expositions muséales (Jaffar, Abubakar, & Zulkifli, 2018). Dans (Partarakis, et al., 2022), une application de Réalité augmentée qui utilise les humains virtuels comme conteurs en langage des signes pour les expériences sur place de l'habitat essentiel est proposée.

Le catalogage est une partie importante du processus muséal qui permet la classification systématique de l'information sur les artefacts, comme le matériel, la forme, la fonction et l'histoire. Les technologies de Réalité augmentée peuvent être utilisées pour cataloguer l'information qui décrit les artefacts des musées dans une description dynamique et augmentée (Patti, 2020).

La réalité augmentée peut aider à impliquer un nouveau public dans les musées grâce à des environnements multidimensionnels qui transposent les visiteurs des expériences physiques aux expériences numériques et permettent l'exploration de nouvelles vues [ (Münzer, 2020), (Popovici, Iordache, Comes, Neamțu, & Băutu, 2022)]. De nombreuses études (Rizvić, et al., 2021) estiment que, dans un proche avenir, la Réalité augmentée devrait être utilisée dans une grande variété d'expositions muséales qui accroîtront la connaissance et l'interactivité des visiteurs. Cependant, certains problèmes limitent encore l'utilisation des technologies de AR, principalement en ce qui concerne l'expérience utilisateur (UX) et les connaissances techniques requises pour développer des applications de Réalité augmentée (Münzer, 2020).

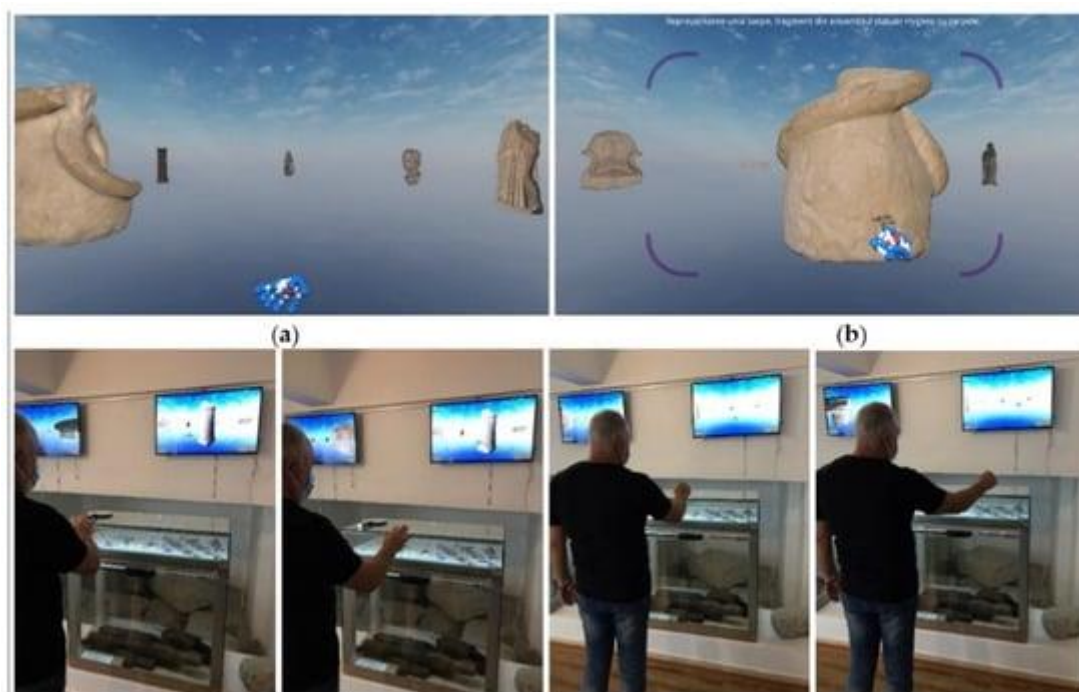


Figure 71 : Visualisation et interaction avec l'ensemble de la scène 3D (a) et un artefact virtuel (b). Avec l'aimable autorisation de Popovici et al. (Popovici, Iordache, Comes, Neamțu, & Băutu, 2022).

## 7.4 Expérience utilisateur

Une évaluation du potentiel technologique de la Réalité augmentée du point de vue des intérêts des parties prenantes est effectuée dans (Tom Dieck & Jung, Value of augmented reality at cultural heritage sites: A stakeholder approach, 2017). Les intervenants sont considérés comme internes, comme les membres du personnel d'un musée, ou externes, comme les visiteurs du musée ou les représentants de la collectivité. Les auteurs explorent les implications de l'utilisation de la réalité augmentée sur diverses dimensions, telles que l'impact économique, les effets sociaux ou l'impact éducatif, dans une étude qualitative basée sur l'utilisateur réalisée dans un petit musée au Royaume-Uni. L'étude des utilisateurs a permis de recueillir les impressions des intervenants sur les éléments liés à chacune des dimensions où la valeur de la Réalité augmentée était évaluée (économique, expérientielle, sociale, épistémique, historique et culturelle, éducative). Il a été prouvé que le patrimoine culturel profite grandement de l'adoption de la AR, du point de vue des intervenants internes et externes.

La volonté des utilisateurs d'adopter des gadgets spécifiques à la réalité augmentée, tels que les lunettes intelligentes de la réalité augmentée, et le potentiel derrière ces dispositifs du point de vue de l'utilisateur sont étudiés dans [ (Han, Tom Dieck, & Jung, 2019), (Litvak & Kuflik, 2020)]. L'étude (Han, Tom Dieck, & Jung, 2019) a utilisé un échantillon de 28 visiteurs dans une petite galerie d'art du Royaume-Uni, et elle s'est concentrée sur l'identification des facteurs qui influencent les préférences des utilisateurs pour les lunettes intelligentes dans un contexte culturel. Bien que l'innovation ait été perçue comme un atout, l'interaction a été difficile pour certains utilisateurs, même s'ils ont assimilé les lunettes intelligentes aux lunettes régulières qu'ils portaient. Le plaisir perçu était positif, tout comme l'utilité perçue.

En (Litvak & Kuflik, 2020), des lunettes intelligentes ont été utilisées dans un cadre extérieur, améliorant avec des aides audio et visuelles les expositions d'un musée. Dans l'étude

expérimentale, un guide touristique mis en œuvre comme application mobile est comparé à la solution de guide de lunettes intelligentes AR, en ce qui concerne la convivialité et l'acceptation par l'utilisateur, suivant une méthodologie statistique approfondie. Bien que la nouveauté de l'utilisation de lunettes intelligentes ait été la principale attraction, l'étude a également révélé certains problèmes liés à l'environnement extérieur spécifique, tels que l'éclairage, qui doivent être abordés à l'avenir.

Les lignes directrices pour la conception d'applications utilisant des dispositifs portables et en particulier des lunettes intelligentes sont dérivées d'une expérience au MIT Museum en (Mason, 2016). Des données d'observation ont été recueillies auprès de 12 participants recrutés parmi les étudiants du MIT, interagissant librement avec un prototype de verrerie dans la galerie de robotique, suivie d'une étude qualitative qui a extrait les tendances émergentes du texte, transcrit à partir d'entrevues audio.

L'influence de la réalité mixte sur l'expérience muséale est modélisée en (Trunfio & Campana, A visitors' experience model for mixed reality in the museum, 2020). Le contexte du musée est fourni par l'exposition « l'Ara tel qu'il était » du Musée Ara Pacis. Dans la conception de la recherche, l'expérience du visiteur est évaluée selon sept dimensions (information muséale, personnalisation, format, convivialité, économie d'information, interaction, expérience) avec 23 variables individuelles. L'échantillon d'utilisateurs relativement vaste et diversifié (586 visiteurs locaux et internationaux) donne de la valeur aux conclusions de l'étude, à savoir que les utilisateurs sont modérément ou très satisfaits de l'intégration de la réalité mixte au Musée. Cependant, l'étude n'a été réalisée que dans un seul musée, et elle laisse aux futures questions de travail sur l'influence que divers facteurs, tels que les antécédents culturels ou les expériences antérieures de réalité mixte, ont sur l'expérience globale de réalité mixte.

Des avantages des applications de Réalité augmentée touristiques relatives au patrimoine culturel ont été explorées pour améliorer l'attrait des sites touristiques culturels locaux (voir la figure 71). L'expérience de l'utilisateur avec l'utilisation d'applications de Réalité augmentée mobiles dans le contexte du tourisme du patrimoine culturel est modélisée en (Han, Tom Dieck, & Jung, 2019), et elle est confirmée empiriquement dans une mise en œuvre particulière pour le tourisme urbain à Dublin, en Irlande. L'échantillon d'utilisateurs était limité, et l'étude qualitative n'a pas été suivie d'une étude quantitative pour tester son pouvoir de généralisation.

## 7.5 Education

Alors qu'on traverse la troisième décennie du XXI<sup>e</sup> siècle, les innovations en éducation sont plus nécessaires que jamais, car de nombreuses distractions se font concurrence pour détourner l'attention des élèves de l'acte éducatif. Les technologies numériques, auxquelles les jeunes utilisateurs sont fortement attirés, peuvent être capitalisées pour transformer la façon dont les élèves apprennent.

La réalité virtuelle et la réalité augmentée peuvent jouer un rôle crucial, permettant aux utilisateurs d'apprendre et de répéter des expériences avec des coûts minimes autant de fois qu'ils le souhaitent (Besbes, Collette, Tamaazousti, Bourgeois, & Gay-Bellile, 2012). En plus des avantages de la réalité virtuelle, la réalité augmentée est moins coûteuse et peut fournir des moyens de communication plus interactifs et coopératifs entre les étudiants (Liu, Liu, Yang, Guo, & Cai, 2018).

L'équipement de Réalité augmentée portable - en particulier Google Glass - est utilisé en (Tom Dieck, Jung, & Tom Dieck, Enhancing art gallery visitors: learning experience using wearable augmented reality: Generic learning outcomes perspective, 2018) dans une expérience contrôlée visant à évaluer l'effet des vêtements sur l'expérience d'apprentissage des utilisateurs, dans le contexte d'une galerie de peintures britanniques du XIXe siècle. Une évaluation du niveau de plaisir des utilisateurs, leur niveau de compréhension et de connaissance et l'amélioration de leurs compétences, et leur appréciation de l'expérience d'apprentissage a complété l'expérience contrôlée (Tom Dieck, Jung, & Tom Dieck, Enhancing art gallery visitors: learning experience using wearable augmented reality: Generic learning outcomes perspective, 2018). Certains inconvénients de l'utilisation de Google Glass ont également été identifiés, comme la diminution de l'interaction sociale entre les apprenants utilisant les dispositifs portables.

La réalité augmentée combinée à un double affichage et à un écran multitactile a été trouvée pour aider les élèves à mieux comprendre le contexte spatial des monuments historiques dans (Novotný, Lacko, & Samuelčík, 2013).

L'apprentissage mobile au moyen d'une application de Réalité augmentée sur une tablette utilisée dans l'apprentissage réel sur le terrain est comparé à l'apprentissage en ligne traditionnel qui se déroule dans une salle de classe avec un ordinateur de bureau ordinaire en (Joo-Nagata, Abad, Giner, & García-Peñalvo, 2017). Le contexte d'apprentissage est fourni par des éléments patrimoniaux de la ville de Santiago du Chili. L'étude a révélé que le processus d'apprentissage mobile améliore considérablement les résultats scolaires. L'acquisition de connaissances historiques est au cœur de l'application de jeu mobile multi-utilisateurs présentée dans (Angelopoulou, et al., 2011), utilisée à la fois en intérieur et en extérieur. Des fragments d'histoire sont découverts de manière ludique par les utilisateurs, les engageant de manière transparente dans le processus d'apprentissage (Etxeberria, Asensio, Vicent, & Cuenca, 2012). Les jeux sérieux, c'est-à-dire les « jeux à des fins éducatives », sont un outil pédagogique précieux dans le domaine des sciences humaines, permettant aux utilisateurs de s'engager et d'être immergés dans l'environnement virtuel créé (Mortara, et al., 2014). Différents genres de jeux (comme l'aventure, la stratégie ou les jeux de puzzle) et différents objectifs d'apprentissage aident à distinguer les nombreux jeux sérieux examinés dans (Damala, Hornecker, Van der Vaart, van Dijk, & Ruthven, 2016) : les jeux qui ciblent la sensibilisation.

Bien que l'utilisation de nouvelles technologies soit un déclencheur de la motivation des élèves à apprendre (González Vargas, Fabregat, Carrillo-Ramos, & Jové, 2020), il est intéressant d'évaluer les impressions de ceux qui dispensent habituellement des cours. Les points de vue des enseignants sur l'utilité des outils de Réalité augmentée dans l'apprentissage, ainsi que leur volonté de s'apprendre continuellement et de participer à la création de ces outils, sont présentés dans (Tzima, Styliaras, & Bassounas, Augmented reality applications in education: Teachers point of view, 2019). Les enseignants se sont montrés désireux de se préparer en permanence, en participant à des séminaires et à des sessions de formation sur les nouvelles technologies. Bien que la majorité des personnes examinées n'aient aucune expérience de la modélisation 3D, elles ont toutes manifesté de l'intérêt pour l'apprentissage des techniques de modélisation 3D et envisageraient même de participer à l'élaboration de nouveaux outils d'apprentissage de la réalité augmentée.



L'art et les technologies multimodales s'entrelacent dans les applications de l'éducation ou du patrimoine culturel (Camurri & Volpe, 2016) qui décrivent, entre autres, un environnement interactif créé pour arbitrer la communication non verbale et la collaboration sociale, en utilisant une interaction audio complexe. La collaboration et la multimodalité sont les caractéristiques distinctives des applications patrimoniales VR/AR qui permettent des relations appropriées entre les utilisateurs et le contexte culturel (Bekele & Champion, 2019). Dans (Pietroni, 2012), la Réalité augmentée est utilisée pour soutenir l'expérience du visiteur du paysage culturel de Matera grâce à une application mobile.

Au cours de la dernière décennie, la réalité augmentée et la réalité virtuelle ont touché le processus éducatif dans divers domaines, comme les processus industriels (Besbes, Collette, Tamaazousti, Bourgeois, & Gay-Bellile, 2012), les musées [ (Novotný, Lacko, & Samuelčík, 2013), (González Vargas, Fabregat, Carrillo-Ramos, & Jové, 2020)], l'art (Liu, Liu, Yang, Guo, & Cai, 2018), l'histoire [ (Angelopoulou, et al., 2011), (Etxeberria, Asensio, Vicent, & Cuenca, 2012), (Mortara, et al., 2014)], la géographie (Angelopoulou, et al., 2011), et même la littérature (Fenu & Pittarello, 2018). Nous nous attendons à ce que de telles applications soient florissantes à l'avenir, car le pipeline technologique est maintenant à maturité et le développement des applications Réalité augmentée devient plus facile.

## 7.6 Tourisme

Le potentiel de la technologie réalité mixte dans le tourisme a été de plus en plus exploité au cours de la dernière décennie, en raison de l'accessibilité croissante de la dernière technologie pour la création de contenu patrimonial 3D, la visualisation et la diffusion interactive, dans les milieux d'éducation publique et dans les unités universitaires/de recherche.

Avec des prédécesseurs comme Rome renaît [ (Rome Reborn, s.d.), (Frischer, Abernathy, Giuliani, Scott, & Ziemssen, 2006)] et Milet antique (Gaitatzes, Christopoulos, Voulgari, & Roussou, 2000), Archeoguide (Vlahakis, et al., 2002) est l'un des premiers projets de Réalité augmentée de patrimoine culturel qui ont mis en évidence l'influence que la technologie Réalité augmentée peut instiller dans l'autonomisation du tourisme en Grèce au début du 21<sup>ème</sup> siècle et n'a pas cessé d'inspirer depuis.

Quelle que soit la solution de visualisation mise en œuvre, CAVE (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992) avec la promesse de visites virtuelles orientées groupe ou HMD (Marasco, Buonincontri, van Niekerk, Orłowski, & Okumus, 2018) avec une qualité croissante d'immersion utilisateur ou de grands affichages visuels (Czernuszenko, et al., 1997) installés dans un musée, les défis auxquels la technologie a dû faire face en ce qui concerne la satisfaction des touristes à utiliser le système étaient fortement liés à l'expérience culturelle (Mason, 2016).

En étudiant l'impact que les technologies de réalité virtuelle peuvent avoir sur les activités touristiques, Marasco et al. (Marasco, Buonincontri, van Niekerk, Orłowski, & Okumus, 2018) ont prouvé que même si une forte implication émotionnelle des utilisateurs est déclenchée par une expérience basée sur la réalité virtuelle avec un grand attrait visuel perçu, l'impact sur les intentions comportementales de visiter un site du patrimoine culturel n'était pas aussi positif que prévu (Marasco, Buonincontri, van Niekerk, Orłowski, & Okumus, 2018). Les auteurs avertissent que leur étude a été réalisée sur un seul site de tourisme virtuel, de sorte qu'une généralisation de leurs résultats ne peut pas être faite. Cependant, ils apprécient le potentiel de

marketing que représente l'expérience de VR, en raison des implications émotionnelles de l'expérience des utilisateurs, qui se traduit par la propagation d'impressions positives à d'autres utilisateurs potentiels.

Le plus grand pas de CAVE ou de grands systèmes d'affichage public VR a été fait une fois que les appareils mobiles sont devenus assez puissants pour être plus que simples appareils d'affichage (Etxeberria, Asensio, Vicent, & Cuenca, 2012). De plus, les solutions axées sur la réalité virtuelle ont commencé à être remplacées par des solutions de réalité augmentée comme nouveau moyen d'améliorer la satisfaction des utilisateurs tout en préservant le patrimoine culturel et en étant en mesure d'offrir une expérience d'apprentissage adaptée aux utilisateurs (Tom Dieck & Jung, Value of augmented reality at cultural heritage sites: A stakeholder approach, 2017).

Dans (Jung, Lee, Chung, & Tom Dieck, 2018), l'effet des influences culturelles sur l'acceptation ou l'adoption des demandes de Réalité augmentée est évalué dans une étude impliquant des participants de la République de Corée du Sud et de la République d'Irlande. L'esthétique de la Réalité augmentée a une forte influence positive sur l'utilité perçue, la facilité d'utilisation perçue et la jouissance perçue de la AR, dans les deux pays.

Une application multimodale VR/AR dédiée au monument du patrimoine culturel de l'île de Jeju en Corée du Sud est décrite en détail dans (Jung, Nguyen, Piscarac, & Yoo, 2020). Un assistant virtuel aide l'utilisateur à naviguer sur le site patrimonial, tandis que l'immersion complète est rendue possible par l'utilisation du casque Google Cardboard VR, et la visualisation Réalité augmentée et la manipulation virtuelle des artefacts est également rendue possible par l'application. Plusieurs hypothèses sont exprimées, dans le cadre fourni par le modèle d'acceptation de la technologie, et vérifiées dans une étude auprès de 251 utilisateurs qui ont constaté que les caractéristiques hédoniques de l'application l'emportent sur d'autres caractéristiques liées à l'utilité.

La valorisation du patrimoine culturel au moyen d'un projet de Réalité augmentée portable dans un contexte de musée d'art fait l'objet d'une étude en (Tom Dieck, Jung, & Tom Dieck, 2018). Comme équipement portable, le Google Glass est considéré, et l'étude effectue une comparaison de deux échantillons, les porteurs de Google Glass et un groupe de contrôle, en ce qui concerne les résultats d'apprentissage, le plaisir et l'influence sur le comportement futur. Le groupe témoin a montré un gain de connaissances légèrement supérieur à celui des participants à l'application d'IA portable, qui étaient davantage distraits par la nouveauté de l'application et qui ne se concentraient pas aussi bien sur l'expérience d'apprentissage. D'autre part, les utilisateurs de Google Glass avaient un accès meilleur et plus rapide à une plus grande quantité d'informations, de manière transparente. L'isolement social a été signalé comme un inconvénient de l'utilisation de la solution AR.

Les lunettes intelligentes sont utilisées dans une autre application de Réalité augmentée dans un musée d'art (Han, Tom Dieck, & Jung, 2019). Les participants à l'étude ont perçu positivement le caractère novateur de l'application, mais ils ont apprécié l'interaction comme difficile, n'étant pas habitués à l'équipement portable auparavant. La majorité des participants ont trouvé l'application agréable et utile du point de vue du tourisme culturel.

Quel que soit le niveau de mélange entre l'environnement réel et le contenu généré par ordinateur, les environnements virtuels du patrimoine culturel sont inspirés par des faits

historiques et ont pour fonction principale la visualisation et l'interprétation de répliques numériques. Une fois que l'utilisateur veut atteindre le niveau suivant de l'expérience culturelle virtuelle, l'interaction avec les artefacts virtuels doit se faire sur la base de scénarios alternatifs qui introduisent dans des configurations hypothétiques le patrimoine historique connu et contesté (Bec, et al., 2019).

La navigation assistée annotée sur place permet une profonde exploration multimodale des sites du patrimoine culturel [ (Garau, 2014), (Garau & Ilardi, 2014)]. L'expérience des utilisateurs, tant à l'intérieur qu'à l'extérieur des musées, est améliorée par l'utilisation d'applications mobiles qui facilitent le partage de contenu entre les utilisateurs et la création d'une mémoire culturelle (Girbacia, Butnariu, Orman, & Postelnicu, 2013), en tirant profit du patrimoine culturel matériel et immatériel.

L'intégration d'éléments VR/AR/RM dans les activités touristiques, comme une visite de musée, offre des opportunités dans un projet de réalité mixte développé à Rome, dans le musée Ara Pacis (Trunfio, Campana, & Magnelli, Measuring the impact of functional and experiential mixed reality elements on a museum visit, 2020), mais présente également certaines difficultés inhérentes à un tel projet. L'immersion des utilisateurs dans un espace muséal repensé pour s'adapter à des éléments IRM, tels que l'audio, la vidéo ou le toucher, améliore l'expérience à partir de divers points de vue, y compris l'éducation, la socialisation et le divertissement. L'expérience des utilisateurs participant à une visite virtuelle augmentée à l'aide de lunettes intelligentes est évaluée en (González-Rodríguez, Díaz-Fernández, & Pino-Mejías, 2020) à travers l'analyse du texte des commentaires publiés sur la plateforme TripAdvisor. Deux applications mobiles Réalité augmentée pour la capitalisation du patrimoine urbain sont décrites dans (Han, tom Dieck, & Jung, 2018). Ils sont analysés dans une étude d'utilisateurs qui a évalué leur utilité, leur esthétique et leur intérêt, et les deux applications enrichissent l'expérience touristique, tant en extérieur qu'à l'intérieur.

## 7.7 Jeux

Ces dernières années, il y a eu des développements très intéressants dans l'espace de gamification qui peuvent aider à changer la perception des types d'applications et d'utilisations de la Réalité augmentée dans l'HC. Le concept de gamification a été défini comme étant « l'utilisation d'éléments de conception de jeux dans des contextes autres que celui des jeux » (Deterding, Dixon, Khaled, & Nacke, 2011) ou « la science du jeu qui explore les diverses techniques de conception et les préoccupations connexes, qui peuvent être utilisées pour ajouter des éléments de jeu aux processus réels existants » (Landers, Auer, Collmus, & Armstrong, 2018). Plus récemment, une autre définition précise que « la gamification consiste à prendre quelque chose qui n'est pas un jeu et à appliquer des mécanismes de jeu pour accroître l'engagement, le bonheur et la loyauté des utilisateurs » (The Ultimate Definition of Gamification (with 6 Real World Examples), s.d.). Les avantages offerts par la gamification mis en évidence par la définition ci-dessus ont également été exploités pour améliorer le domaine de l'héritage culturel en le rendant plus intéressant et en offrant des moyens uniques et divertissants pour les gens d'apprendre et de vivre ces espaces. Souvent appelée « jeux omniprésents » ou « apprentissage par le jeu » (Xu, Buhalis, & Weber, 2017), la gamification a été utilisée dans ce domaine principalement sous deux formes : les jeux sérieux (SGs) et la raconte d'histoire (storytelling).

Les SG ont été créés dans le but de produire des résultats constructifs ou éducatifs en plus de la valeur du divertissement. Leur objectif est donc d'aider l'utilisateur à atteindre ses objectifs d'apprentissage grâce à une expérience amusante (Ibañez-Etxeberria, Gómez-Carrasco, Fontal, & García-Ceballos, 2020). En HC, les SG sont considérés comme l'une des façons les plus novatrices de comprendre la culture et l'histoire en utilisant les technologies interactives (Ferdani, Fanini, Piccioli, Carboni, & Vigliarolo, 2020). Les SG sont caractérisés par les caractéristiques suivantes : (a) le comportement de l'utilisateur, l'apprentissage et les objectifs de réalisation sont plus importants que le divertissement ; (b) la conception du jeu est spécifiquement créée pour améliorer les connaissances, les compétences et les comportements des utilisateurs cibles.

Il y a plusieurs façons d'utiliser les SG pour un plus grand engagement avec l'HC. Par exemple, les concepts SG pourraient être utilisés pour gamifier les interactions avec les artefacts en utilisant des interfaces multimodales pour créer des expériences immersives pour les utilisateurs (Liarokapis, Petridis, Andrews, & De Freitas, 2017). Un système VR haptique immersif pour recréer l'expérience de tir à l'arc a été présenté dans (Butnariu, et al., 2018). La rétroaction haptique avait le potentiel d'améliorer l'expérience utilisateur, comme le montre (Ceccacci, Generosi, Leopardi, Mengoni, & Mandorli, 2021). En outre, le SG et la gamification ont été utilisés pour dans le contexte des espaces sociaux, offrant aux visiteurs la possibilité d'explorer les espaces HC d'une manière amusante et engageante en même temps. En particulier, la gamification dans le contexte des musées a été un sujet de recherche fructueux avec quelques exemples intéressants de systèmes Réalité augmentée utilisés pour l'interaction de type jeu [ (Augello, Infantino, Pilato, & Vitale, 2021), (Hammady, Ma, & Temple, 2016)]. Les expériences améliorées de Réalité augmentée dans les musées peuvent maximiser la satisfaction des utilisateurs et les résultats d'apprentissage (Paliokas, et al., 2020). Malgré les possibilités offertes par les environnements intérieurs, de nombreuses expériences de réalité augmentée récentes ont été conçues en utilisant des téléphones intelligents et des tablettes comme matériel de soutien pour les environnements extérieurs [ (Slavec, Sajincic, & Starman, 2021), (Tzima, Styliaras, & Bassounas, 2021)], par exemple pour découvrir les emplacements de CH (Bujari, Ciman, & Gaggi, 2017) et améliorer l'apprentissage expérientiel (Vlizon, Sharamyeva, & Kotsopoulos, 2021). En (Evangelidis, Sylaiou, & Papadopoulos, 2020), une réalité mixte (RM) assistée par les technologies géoinformatiques pour relancer les événements historiques a été présentée.

L'autre forme de gamification avec un grand potentiel pour révolutionner la façon dont les utilisateurs interagissent avec HC est le storytelling. La narration numérique est définie comme l'évolution de la narration traditionnelle qui combine l'art ancien de raconter une histoire et les possibilités technologiques actuelles (Vert, et al., 2021). Les technologies de réalité étendue (XR) ont été utilisées avec succès dans des applications DS interactives pour vivre des événements passés et rencontrer des personnages historiques (Rizvić, et al., 2021) ou pour améliorer l'appréciation de la poésie chinoise classique (Zhao & Ma, 2020).

Bien que la gamification ait été discutée comme un moyen d'encourager la participation, les technologies immersives ont également un potentiel clair pour soutenir le public intéressé par l'HC, non seulement pour les éléments tangibles de l'HC, tels que les artefacts, les bâtiments, des sites historiques, mais aussi des HC intangibles pour mettre en valeur les traditions sociales (Jofresa, Xirau, Ereddam, & Vicente, 2019) comme les sports traditionnels (Tisserand, et al., 2017) ou l'artisanat (Hauser, et al., 2022). La réalité augmentée et la gamification peuvent être

considérées comme des domaines connexes, et leur succès est étroitement lié à l'amélioration des expériences des gens et à la promotion de l'apprentissage et de la collaboration. L'utilisation de la Réalité augmentée en HC sera extrêmement diversifiée en termes de contenu, de types d'interaction, de lieux et de contextes, ce qui permettra des expériences multisensorielles (Marto, Gonçalves, Melo, & Bessa, 2022).

## 7.8 Patrimoine culturel intangible

Le patrimoine culturel immatériel (CIH) joue un rôle important dans la vie de chaque communauté parce qu'il est une expression de l'histoire, des traditions et de la culture des gens qui ont vécu et travaillé dans une certaine région. L'UNESCO inclut dans l'ICH les éléments suivants : « traditions orales, arts de la scène, pratiques sociales, rituels, événements festifs, connaissances et pratiques concernant la nature et l'univers ou les connaissances et compétences nécessaires pour produire des métiers traditionnels » (What is Intangible Cultural Heritage, s.d.).

Les solutions numériques pourraient contribuer à préserver et à protéger l'ICH. La technologie de la réalité augmentée a un grand potentiel pour améliorer la perception du patrimoine culturel immatériel (Lu, Wang, & Chen, 2020). Dans (Zhao Z. , 2017), les auteurs ont proposé un système de Réalité augmentée conçu pour la protection numérique de l'ICH qui utilise un terminal intelligent pour fusionner son contenu et des images vidéo réalistes dans la scène.

Les technologies de Réalité augmentée peuvent être utilisées pour superposer la décoration virtuelle liée à l'ICH dans l'espace de vie, ce qui apporte une nouvelle perspective agréable à l'expérience utilisateur (Yang & Zhao, 2017) De plus, selon les technologies de la réalité augmentée, l'environnement peut devenir semblable à une scène où l'on peut apprendre des danses traditionnelles (par exemple, la danse aborigène australienne Al Ardha (Khan, 2014) ou la danse traditionnelle grecque « Syrtos in Three » (Ziagkas, et al., 2020)).

Dans le domaine de l'affichage numérique pour le contenu lié à l'ICH, l'équipement de Réalité augmentée peut fournir des expériences virtuelles d'essai de costumes et d'accessoires traditionnels (Wen & Chen, 2016). En faisant l'expérience de l'environnement AR, les utilisateurs pouvaient se transposer dans une atmosphère traditionnelle (Xie & Tang, 2018). Les murs Réalité augmentée interactifs qui combinent des marqueurs, une cartographie de projection et des animations sont également adaptés pour présenter le patrimoine folklorique, les traditions nationales et les légendes (Laštovička-Medin, 2019). Les visiteurs peuvent également faire l'expérience d'une impression immergée en utilisant HMD en conjonction avec des matériaux liés à l'ICH (Huang, Xiang, & Li, 2019). Malgré tous les avantages de l'utilisation de la technologie AR, un inconvénient est l'augmentation du temps nécessaire au processus de production des applications Réalité augmentée (Viinikkala, et al., 2016).

## Conclusion

Ce chapitre a présenté un ensemble d'aspects relatifs aux systèmes de réalité augmentée. Nous avons discuté de différentes techniques liées à la réalité augmentée. Un aperçu de chacune d'elles a été présenté, en identifiant et en mettant en évidence les principales caractéristiques de chaque technique. En outre, nous avons expliqué en détail les principaux problèmes liés au patrimoine virtuel en réalité augmentée. Nous avons introduit les principales techniques de reconstruction 3D utilisées dans le patrimoine culturel. Nous nous sommes

concentrés sur les principaux enjeux de la réalité augmentée pour le patrimoine culturel. Nous avons présenté quelques recherches liées à ces domaines et mis en évidence le principal problème de chaque recherche.

L'objectif principal de cette étude était d'identifier et de donner un aperçu des sujets les plus importants de la dernière décennie dans le domaine de la réalité augmentée et du patrimoine culturel. En résumé, les résultats ont montré un intérêt croissant pour les applications de la Réalité augmentée dans de nombreux contextes de HC, des reconstructions 3D au patrimoine culturel immatériel.

La reconstruction 3D d'objets du patrimoine culturel est une tâche complexe. Au cours de la dernière décennie, plusieurs techniques de reconstruction 3D telles que la photogrammétrie, le balayage laser, le Lidar et la lumière structurée ont été utilisées avec succès pour de nombreux artefacts HC.

Le patrimoine numérique bénéficie largement des technologies de la réalité augmentée, dont l'adoption récente et généralisée en fait un défi majeur pour la compréhension et l'étude du passé. Le développement de la Réalité augmentée a donné lieu à d'importantes recherches visant à répondre aux besoins des musées, des archives et des établissements du patrimoine. Les applications développées démontrent que la réalité augmentée peut améliorer plusieurs tâches dans le patrimoine numérique, à savoir la visualisation des artefacts reconstruits, la documentation et la compréhension contextuelle. De plus, grâce à la technologie de la réalité augmentée, le patrimoine numérique combine la production des industries créatives avec la numérisation des artefacts de HC afin d'enrichir l'expérience du visiteur grâce à des archives numériques animées, mêlant les vestiges du monde réel aux médias numériques. Il est vraiment important que les touristes aient la possibilité non seulement de toucher le passé à travers des répliques imprimées en 3D de pièces qui transcendent l'histoire ancienne, mais aussi d'expérimenter et ainsi d'être immergés culturellement dans des installations interactives impliquées dans des activités humaines perdues.

Les musées virtuels de Réalité augmentée permettent un niveau plus profond d'implication et de connexion entre les musées et leurs visiteurs. La réalité augmentée a été utilisée pour rendre le contenu du musée plus intéressant et attrayant pour les utilisateurs. Pour créer des expériences de AR, une variété de techniques sont utilisées afin de rendre une grande quantité d'informations sur les artefacts, principalement à l'aide d'un appareil mobile. De cette façon, les technologies de Réalité augmentée ont le potentiel d'accroître l'accès et l'échange d'information qui décrit les artefacts de HC de façon dynamique. Dans le même temps, la Réalité augmentée commence à être largement utilisée pour offrir de nouveaux publics dans les musées grâce à des environnements multidimensionnels qui transposent les visiteurs des expériences physiques aux expériences multimodales numériques. Cependant, l'utilisation des technologies de Réalité augmentée demeure limitée par quelques problèmes, dont la plupart sont liés aux compétences techniques complexes nécessaires pour développer ces applications.

Le potentiel de la technologie Réalité augmentée disponible a été prouvé en HC, tant du point de vue économique que social. Tandis que les gadgets spécifiques de la Réalité augmentée déclenchent la volonté d'utiliser, ils se sont avérés être plus appropriés dans les installations d'intérieur plutôt que ceux d'extérieur. La technologie mobile reste l'une des solutions les plus accessibles et prometteuses pour la visualisation Réalité augmentée orientée utilisateur de l'HC. Pour obtenir une application de Réalité augmentée efficace du point de vue de

l'expérience utilisateur, il faut tenir compte de certaines considérations de conception. La convivialité, la visualisation intuitive du contenu de la réalité augmentée et la conception des interfaces UX doivent toutes être considérées comme importantes. La jouissance et l'utilité sont des aspects importants, peu importe si nous tenons compte de la liberté de l'utilisateur d'interagir au moyen de la technologie de la réalité augmentée, où les lunettes intelligentes sont préférées en raison des mains libres de l'utilisateur et des métaphores gestuelles mises en œuvre, plutôt que de l'interaction avec le mobile multi-écrans tactiles. De plus, l'élaboration d'une application de Réalité augmentée pour l'HC exige une connaissance et une compréhension des aspects culturels et historiques.

L'éducation a été à l'honneur au cours de la dernière décennie, surtout ces dernières années. Il est particulièrement difficile d'attirer l'attention des élèves et de la maintenir dans le contexte de l'utilisation intensive des médias sociaux et des ordinateurs ou des appareils mobiles. Les progrès technologiques récents doivent être exploités pour ajouter de la valeur au processus éducatif, et la réalité augmentée, la réalité virtuelle et la réalité mixte ne sont pas des exceptions. Dans les outils éducatifs liés au patrimoine culturel, la réalité augmentée joue un rôle important, en permettant une immersion complète dans les contextes historiques reconstruits en 3D et/ou une interaction libre avec les artefacts virtuels reconstruits. La gamification est combinée avec la réalité augmentée/réalité virtuelle, assurant une expérience collaborative et engageante, dans laquelle l'apprentissage se déroule de manière transparente. En même temps, grâce à l'ajout de caractéristiques visuelles et auditives aux collections numériques, la Réalité augmentée accroît l'accès des élèves aux ressources en HC et leur compréhension.

Il est difficile de trouver des expériences HC en dehors d'une expérience touristique. Nous ne pouvons pas imaginer des vacances sans objectifs patrimoniaux intéressants à visiter, localement ou à l'étranger. Les applications de Réalité augmentée peuvent aider le tourisme en améliorant l'expérience des visiteurs et la perception de l'information sur l'HC. En ajoutant des visuels 3D à l'information existante, les applications touristiques basées sur la réalité augmentée peuvent également donner des avantages tangibles en augmentant la valeur des attractions HC existantes.

La gamification est intensément exploitée dans les ressources pédagogiques électroniques, en particulier dans le contexte de l'apprentissage en ligne. Les applications éducatives du patrimoine culturel ne font pas exception, d'autant plus que les outils de Réalité augmentée utilisés pour développer de telles solutions permettent une intégration transparente des comportements ludiques. Qu'il s'agisse d'une approche de jeu sérieux ou de narration, les domaines d'application de l'éducation en HC bénéficient grandement de la gamification. La mobilisation et la collaboration des utilisateurs, qui sont cruciales pour les résultats finaux de tout processus éducatif, sont favorisées par l'intégration d'idées liées aux jeux dans les applications de l'HC. La réalité augmentée et la gamification peuvent constituer un outil viable pour faire participer les participants non seulement aux éléments tangibles de l'habitat essentiel, comme les artefacts, les bâtiments et les lieux historiques, mais aussi à l'héritage culturel immatériel.

L'ICH est en quelque sorte une forme ultime d'expression de l'HC parce qu'elle englobe les artefacts, leurs fonctionnalités, les processus dans lesquels ils sont utilisés et les buts de leur utilisation. Même si nous parlons de traditions (orales et gestuelles), d'arts de la scène, de pratiques sociales, de rituels, de connaissances ou de pratiques axées sur tout type d'activité

humaine (p. ex., les compétences pour produire des métiers traditionnels), aborder l'ICH avec la technologie de la réalité augmentée combinée à la validation de l'action de l'utilisateur (éventuellement basée sur une sorte d'ontologie) pourrait transposer de façon crédible l'expérience d'« immersion culturelle ».

Le chapitre suivant est consacré à l'explication de la méthode que nous proposons pour restreindre la zone de mise en correspondance des algorithmes stéréoscopiques afin de diminuer les temps de calcul nécessaires pour réaliser une reconstruction 3D. La méthode en question a fait l'objet d'une publication dans le journal « Int. J. of Computational Vision and Robotics » en 2022 ; est sera détaillée lors du reste de cette thèse avant de passer à la partie qui concerne nos conclusions et perspectives.



# Chapitre VI

---

**Une combinaison d'approches de  
'mappage de caractéristiques' et de  
'blocs' pour réduire la zone  
d'appariement des algorithmes  
stéréoscopiques.**

---

## **1 Introduction**

Dans ce chapitre, nous proposons une nouvelle approche pour restreindre le champ d'appariement des algorithmes stéréoscopiques (Djaber & Nour Eddine, 2022). Il a été constaté que le calcul de la carte de disparité implique l'utilisation de l'image entière pour un large éventail de méthodes stéréoscopiques, ce qui conduit à un calcul de temps supplémentaire et à des artefacts visuels qui apparaissent dans les résultats. Sur la base de cette observation, nous avons dérivé une approche qui réduit considérablement le temps de calcul des algorithmes stéréoscopiques et évite les bruits pouvant apparaître dans le résultat.

L'approche proposée introduit une forte association entre les bords de silhouettes et les algorithmes stéréoscopiques en utilisant uniquement les informations géométriques présentes dans les images pour restreindre la zone de mise en correspondance. Elle vise essentiellement à limiter la zone d'appariement à la géométrie exacte de l'objet analysé, évitant ainsi des temps de calcul supplémentaires et des bruits indésirables. Nous n'avons pas utilisé d'algorithmes de codage complexe ni d'équipement coûteux, et nous avons obtenu des résultats acceptables en termes de temps et de précision.

La reconstruction 3D à partir d'images (Aharchi et Ait Kbir, 2020 ; Han et al., 2021) est un sujet de recherche actif. Ces dernières années, les algorithmes et les méthodes dans ce domaine progressent encore. Plusieurs méthodes et techniques utilisant uniquement les informations géométriques et photométriques disponibles sur les photographies sont proposées (Hou et al., 2022 ; Durou et al., 2020 ; Kordelas et al., 2010 ; Zollhöfer et al., 2018). Elles visent toutes à récupérer la géométrie 3D pour permettre l'observation de la scène à partir de points de vue arbitraires et à manipuler l'objet avec des outils 3D spéciaux.

Inférer des modèles 3D précis en peu de temps est un sujet de recherche imminent en vision par ordinateur en raison de l'évolution technologique des nouvelles caméras numériques de haut niveau. Un facteur important est la demande croissante de méthodes de numérisation de hautes performances pour prendre en charge les applications multimédias. Cependant, les méthodes stéréo-multivues souffrent du temps supplémentaire nécessaire pour réaliser une reconstruction partielle.

La stéréovision (Perkins, 1970) est une technique qui permet de reconstituer partiellement la structure 3D d'un objet ou d'une scène, à partir de deux images décalées -prises de points de vue différents-. Le calcul des couples correspondants est très compliqué et nécessite un temps de traitement très important. De plus, ce processus devient assez délicat, surtout dans le cas où des images de haute résolution sont utilisées. L'idée centrale des méthodes de stéréoscopie multivues<sup>1</sup> (Lane et Thacker, 1996) est de comparer de petites régions d'image et de rechercher des correspondances avec des apparences similaires ou identiques sur plusieurs images. Cependant, cela conduit à un temps de calcul supplémentaire (puisqu'ils utilisent l'intégralité des images pour les calculs) et provoque une apparition d'artefact en dehors de l'objet reconstruit.

Cependant, les méthodes stéréo multivues souffrent du temps supplémentaire nécessaire pour réaliser une reconstruction partielle. Il a été constaté que le calcul de la carte de disparité implique l'utilisation de l'image entière pour une large gamme d'approches stéréoscopiques, ce qui conduit à un temps de calcul élevé ou implique l'utilisation de matériel de haute performance.

Nous proposons comme contribution de limiter la zone d'appariement des algorithmes stéréoscopiques afin de ne pas faire de calculs en dehors des bords des silhouettes de l'objet à reconstruire, ce qui est possible en raison de la forte connexion qui relie la géométrie et le modèle

3D. Partant de cette idée, les temps de calcul atteints par l'algorithme proposé ont en fait été réduits par rapport aux algorithmes stéréoscopiques classiques.

Dans ce qui suit, nous traiterons notre approche proposée sous le titre : "Une combinaison d'approches de 'mappage de caractéristiques' et de 'blocs' pour réduire la zone d'appariement des algorithmes stéréoscopiques". Une revue de travaux connexes est présentée dans la section 2. La section 3 explique notre méthode proposée. Ensuite, une discussion des résultats obtenus est établie dans la section 4. Enfin, la conclusion est présentée dans la section 5.

## **2 Travaux connexes**

La reconstruction 3D existe depuis des décennies (Ham et al., 2019), et plusieurs algorithmes de haute qualité ont été développés dans la littérature. Ce qui a conduit à une croissance exponentielle de ces techniques. Nous avons essayé lors du chapitre 1 de cette thèse de dresser une classification multicritère des méthodes de reconstruction 3D. Mais en raison de la grande diversité de propriétés et d'hypothèses qui différencient chaque algorithme, il est - malheureusement - assez difficile de placer une méthode dans une classe spécifique ou même de proposer une taxonomie générale.

Plusieurs tentatives de classification des différentes techniques ou d'établissement d'un état de l'art global ont été proposées. Cependant, les techniques de reconstruction 3D se divisent en deux grandes classes principales: les méthodes actives (Buelthoff et Yuille, 1991) qui reconstruisent le profil 3D par approximation numérique et interférence active avec un objet (télémètres mécaniques ou radiométriques, par exemple lumière structurée, télémètre laser, etc.) les méthodes passives (Ham et al., 2019 ; Siudak et Rokita, 2014) infèrent la forme en analysant uniquement les signaux d'image et ne nécessitent donc aucun autre équipement que les caméras monoculaires ou stéréo conventionnelles.

Nous n'avons pas utilisé d'équipements de radiométrie active coûteux pour appliquer notre approche, et nous nous bornons à citer quelques références liées aux méthodes passives. Plusieurs méthodes sont proposées, et les plus populaires d'entre elles sont les formes à partir de X (Buelthoff et Yuille, 1991) et la stéréoscopie. L'objectif principal des méthodes stéréoscopiques est de générer une carte de disparités (Hamzah et Ibrahim, 2016) qui permet d'obtenir des nuages de points par triangulation (Hartley et Zisserman, 2004). Cependant, leurs principaux inconvénients sont : le temps de calcul relativement élevé, les résultats partiels et l'apparition de bruit dans les résultats finaux. En outre, ces méthodes souffrent d'une perte d'informations importante, d'une complexité de calcul et de l'incapacité à détecter les régions concaves ou occlusives. Les deux approches ci-dessus sont complémentaires, et plusieurs travaux (Djaber, 2012 ; Lin et al., 2014 ; Lin et Wu, 2008) proposent leur combinaison pour pallier les inconvénients de chacune d'entre elles. L'article de Lane et Thacker (1996) sur l'appariement stéréo montre une douzaine de méthodes de 1973 à 1992, sans élaborer de comparaison. Des recherches importantes ont été menées par Scharstein et Szeliski (2002), où ils ont comparé plusieurs algorithmes et leurs performances sur la base de plusieurs mesures. Dans une autre étude, Seitz et al. (2006) ont présenté une classification de ces approches de reconstruction basée sur un ensemble de caractéristiques. Selon les auteurs, les techniques stéréoscopiques multi-vues peuvent être classées en fonction de la manière dont le résultat final est représenté, par exemple à travers des nuages de points, des voxels, des maillages ou des cartes de profondeur. Selon Bebeslea-Sterp et al. (2017), la littérature contient probablement une grande variation d'algorithmes et il est presque difficile de faire une comparaison complète des méthodes d'appariement stéréoscopique. Ils ont présenté dans leur travail plusieurs algorithmes, en se concentrant sur leurs idées fondamentales. Aussi, ils ont proposé une classification récente, où ils ont parlé de méthodes locales, de méthodes globales et de méthodes semi-globales.

Les méthodes globales sont comparativement plus lentes mais donnent de meilleurs résultats de correspondance stéréo. En fait, ils minimisent une fonction énergétique globale, qui combine des données et des termes de lissage, pour toutes les valeurs de disparité (Hamzah et Ibrahim, 2016 ; Scharstein et Szeliski, 2002 ; Brown et al., 2003).

Le principal avantage des méthodes locales est leur rapidité de calcul au détriment de la qualité et de la stabilité de l'appariement stéréo. Cependant, ils sont sensibles aux zones ambiguës dans les images, par exemple, les régions d'occlusion et les parties faiblement texturées (Hamzah et Ibrahim, 2016). Les méthodes locales sont basées sur trois approches différentes : bloc, gradient et feature matching (Hamzah et Ibrahim, 2016).

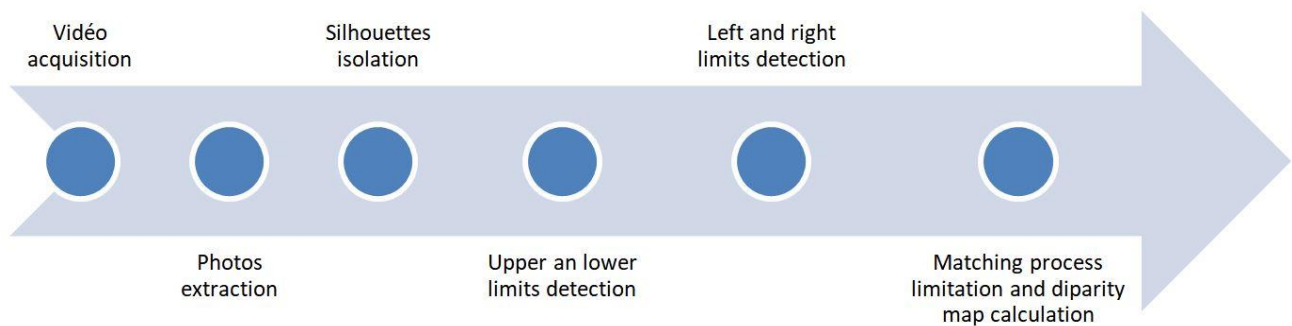
Les méthodes d'appariement de blocs comparent de petites parties - généralement carrées - d'une image avec une série de blocs d'une autre image. Les méthodes d'appariement des caractéristiques limitent les régions pour le calcul de la disparité aux caractéristiques globale de l'image sélectionnée telles que les bords, les courbes ou les points et aboutissent généralement à des cartes de disparité dispersées (Hamzah et Ibrahim, 2016 ; Scharstein et Szeliski, 2002 ; Brown et al., 2003).

Les méthodes de gradient supposent que la luminosité d'un même point 3D visible dans deux images stéréoscopiques est constante et permettent de petites disparités locales entre les images stéréo (Hamzah et Ibrahim, 2016).

Néanmoins, il existe des méthodes qui ne peuvent être strictement rattachées à l'une des principales catégories. Un groupe intermédiaire de méthodes utilise des algorithmes semi-globaux (Sun et al., 2011 ; Hirschmuller, 2005 ; Morales et Lopez, 2016). Ils combinent l'approche de mappage de blocs avec une contrainte d'approximation de lissage global. Il a été montré que les méthodes d'appariement semi-globales produisent des cartes de disparités de meilleure qualité que celles obtenues par les méthodes locales et ont une complexité de calcul nettement inférieure à celle des méthodes globales (Morales et Lopez, 2016).

### **3 Méthode proposée**

Dans cette section, la méthode proposée illustrée à la figure 72 est décrite en détail. Son entrée est un ensemble d'images, et sa sortie est une carte de disparité calculée après la restriction de la zone de mise en correspondance. Les contours des objets sont estimés et utilisés pour restreindre le champ correspondant afin qu'il s'y adapte exactement.



*Figure 72: Pipeline de la méthode proposée.*

Notre objectif est d'isoler un objet de son arrière-plan et de réduire la zone de mise en correspondance pour s'adapter exactement aux silhouettes extraites. L'ensemble du processus comporte quatre phases essentielles :

1. Extraction des images de la vidéo ;
2. Isolation de l'objet et extraction de ses silhouettes ;
3. Restriction de la zone d'appariement ;
4. Application de la stéréoscopie à la zone restreinte.

La méthode proposée utilise les silhouettes d'objets pour définir un champ de recherche restreint. Par conséquent, notre optimisation consiste à limiter les calculs de correspondants aux points situés à l'intérieur des silhouettes pour éviter les extra-calculs et éliminer les éventuelles anomalies pouvant apparaître en dehors de la frontière externe du résultat.

Les résultats obtenus seront analysés dans la section 4. Dans les sous-sections suivantes, nous détaillons chaque étape de la méthode proposée.

### 3.1 Extraction d'images et de silhouettes

Le suivi d'un objet dans une séquence d'images et son isolation de son arrière-plan est une opération assez complexe et nécessite de spécifier un ensemble de caractéristiques permettant de segmenter chaque image (Gwet et al., 2018). Nous avons utilisé un mini studio comme système de prise de vue (inspiré de celui de la figure 73). L'idée de base est de placer un objet dans une boîte carrée avec une seule couleur de fond (les couleurs habituellement utilisées sont le blanc, le bleu ou le vert). Des sources lumineuses sont placées de part et d'autre du mini-studio pour éliminer l'ombre de l'objet.



*Figure 73 : Exemple de mini-studio.*

Le but du mini-studio est de faciliter l'isolation de l'objet de son arrière-plan et d'obtenir des images sans effets d'ombre. Un fond blanc a été utilisé pour réduire la complexité de cette opération et assurer un bon éclairage de l'objet.

On sait que chaque vidéo est composée d'un ensemble de trames ; leur nombre est lié au type de caméra (7 FPS dans notre cas). Deux objets réels ont été utilisés : Minion et Schtroumpf (Figure 74).



Figure 74 : Objets réels, (a) Minion (b) Schtroumpf

Chaque vidéo filmée a une durée de 17 secondes avec une taille de 640\*640 pixels pour chaque image. La méthode proposée est appliquée à une séquence d'images extraites d'une vidéo. L'objet à reconstruire est placé sur une table tournante à vitesse fixe. Les prises de vue sont captées tous les 3 degrés par une caméra fixe positionnée devant l'objet. La rotation complète de l'objet donne un ensemble de 121 images (Figure 75).



Figure 75 : Un sous-ensemble d'images originales

L'objectif principal de l'approche proposée est de réduire la zone de recherche pour diminuer le temps de calcul des algorithmes stéréoscopiques. L'algorithme proposé fonctionne également sur des angles plus grands, mais l'augmentation de l'angle de rotation réduira le nombre d'images à analyser au prix d'une perte considérable d'informations. Par exemple, utiliser un angle de  $9^\circ$  ne donne que 41 photos pour chaque rotation et implique une importante perte d'informations.

La minimisation de l'angle de rotation augmente le nombre d'images ainsi que le temps de calcul nécessaire pour effectuer une opération nécessitant moins d'images. Par exemple, utiliser un angle de rotation de  $1^\circ$  implique de traiter 361 images. Bien que le degré de détail soit meilleur dans ce cas, ce nombre reste important dans un contexte de reconstruction 3D et nécessite des calculs supplémentaires qui peuvent être évités.

Comme compromis entre la diminution de l'angle de rotation impliquant une perte importante d'information, et l'augmentation de l'angle de rotation impliquant des calculs supplémentaires. L'angle de  $3^\circ$  a été choisi car il donne 121 images, ce qui est considéré comme suffisant pour assurer la conservation des informations nécessaires au processus d'appariement sans qu'il soit nécessaire d'effectuer des calculs supplémentaires conduisant aux mêmes résultats.

### 3.2 Isolation d'objet et extraction de ses silhouettes

Une image colorée constitue une conjonction des trois matrices essentielles : rouge, vert et bleu. La séparation de ces matrices permet d'obtenir une silhouette appropriée en utilisant un seuil défini en calculant des valeurs minimales et maximales pour chaque matrice. La silhouette peut être obtenue par comparaison avec des valeurs minimales et maximales ; si un pixel appartient à cet intervalle, il est transformé en blanc car il appartient au fond. Sinon, le pixel deviendra noir. Les images monochromes prétraitées obtenues (dont la matrice représentative ne contient que les valeurs 0 et 1) sont enregistrées pour être utilisées comme entrée de l'algorithme de restriction. La figure 76 illustre quelques silhouettes obtenues.



Figure 76 : Exemple de silhouettes obtenues

### 3.3 Restriction de la zone de correspondance

La méthode proposée limite les calculs pour les points situés exactement à l'intérieur de l'objet, évitant ainsi des calculs supplémentaires et l'apparition de bruit dans les résultats. Dans cette étape, la zone de calcul est limitée verticalement, en recherchant la limite supérieure  $L_s$  et la limite inférieure  $L_i$  représentant les bords où l'objet commence par le haut et se termine par le bas (Figure 77).

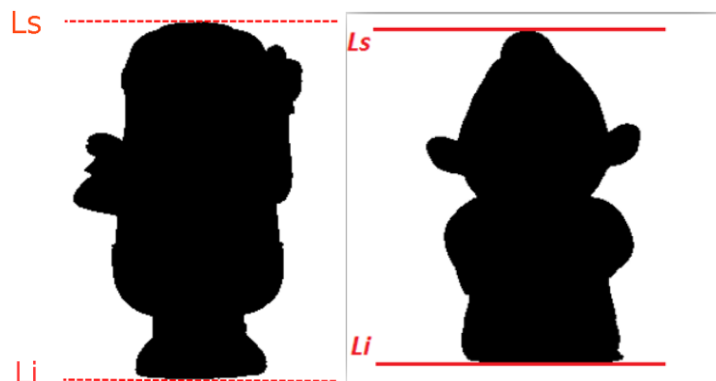


Figure 77 : Délimitation verticale

Un balayage ligne par ligne de la première silhouette est effectué pour calculer les limites supérieure et inférieure. Le premier pixel noir détecté en descente représente le début de l'objet à partir du haut ; son numéro de ligne représente la limite supérieure. Le balayage commence à partir du bas de l'image

pour calculer la limite inférieure. Ces deux valeurs sont identiques pour toutes les images et sont calculées une seule fois (pour un même objet).

Tous les futurs calculs seront effectués dans la zone délimitée par les deux lignes  $L_s$  et  $L_i$ .

La restriction effective de la zone de mise en correspondance commence. Un ensemble de 120 silhouettes est disponible (les images 120 et 121 sont identiques), mais seules 60 d'entre elles sont utilisées. Une silhouette de nombre  $k$  est utilisée simultanément pour calculer les bords de l'objet dans l'image  $k$  ainsi que ceux dans l'image  $k + 60$  ; en effet, le côté gauche de l'image numéro  $k + 60$  représente exactement le côté droit de l'image  $k$  puisqu'il s'agit d'une rotation de  $180^\circ$  ( $3^\circ * 60$  images). Il est donc inutile de refaire les calculs pour deux images symétriques inversées (Figure 78).

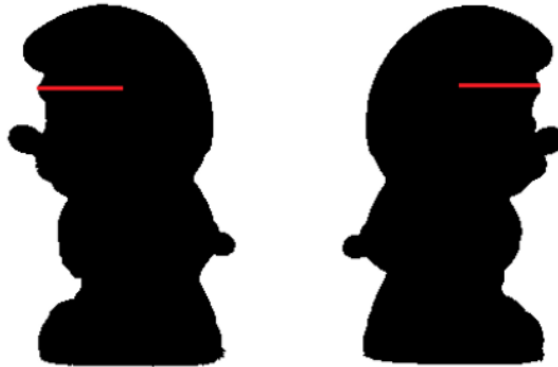


Figure 78 : Détection des bords

Pour déduire les limites gauches de l'objet ; une recherche horizontale est effectuée à partir du début de l'image dans la zone délimitée par  $L_s$  et  $L_i$  (bornes supérieure et inférieure).

Pour chaque ligne, le premier pixel noir détecté représente la limite gauche de l'objet pour cette ligne. Le même principe est réalisé pour les limites droites, mais dans ce cas, le balayage commence de l'autre côté de l'image. L'algorithme relatif est présenté ci-dessous.

---

**Algorithm 1** LeftRightLimits (image)

---

```

1: for line ←  $L_i$  to  $L_s$  do
2:   while (column ≤ imageSize and pixel ≠ black ) do
3:     Column ← column + 1
4:     if pixel is black then
5:       Store the pixel's position in the leftLimit array
6:     end if
7:   end while
8:   while (column ≥ 1 and pixel ≠ black ) do
9:     Column ← column - 1
10:    if pixel is black then
11:      Store the pixel's position in the rightLimit array
12:    end if
13:  end while
14: end for

```

---

Ce processus est effectué pour chaque ligne appartenant à la zone comprise entre  $L_s$  et  $L_i$  pour l'ensemble des 60 images. Les résultats sont ensuite stockés dans une matrice tridimensionnelle nommée arêtes. L'intégration du parallélisme constitue l'une de nos perspectives d'avenir. Par exemple, la fonction LeftRightLimits peut être exécutée en parallèle pour calculer simultanément plusieurs bords d'image, de même que les deux boucles for.



### 3.4 Étape finale

L'appariement stéréoscopique est gourmand en termes de temps de calcul, donc une bonne disparité nécessite un temps de calcul considérable ou un système puissant de calcul. Le système visuel doit déterminer quelles parties des deux images rétiniennes correspondent à la disparité d'estimation (Banks et al., 2004). A cet effet, une fenêtre rectangulaire appelée fenêtre de corrélation est utilisée pour effectuer la mise en correspondance. Pour les objets mal texturés ou contenant des zones homogènes, la fenêtre de corrélation doit être aussi petite que possible pour éviter toute confusion possible et augmenter la probabilité de correspondances de petites régions pour des zones similaires.

Un exemple de fenêtre de corrélation est représenté par le petit carré rouge sur la figure 79 (b).

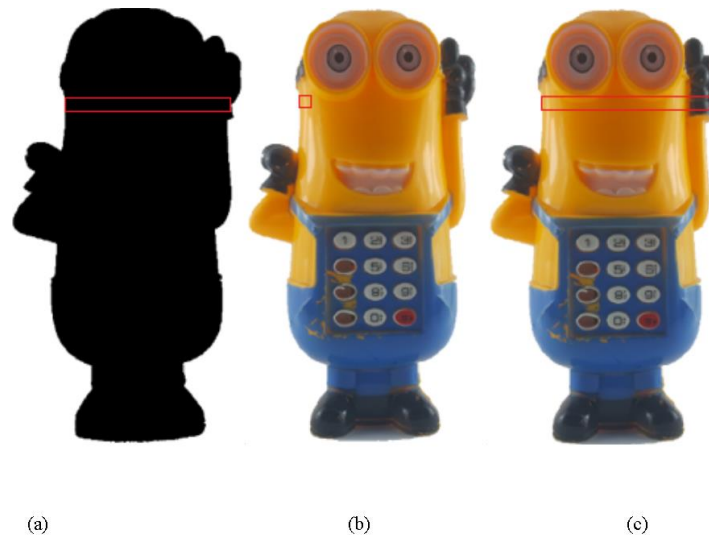


Figure 79 : La zone de recherche effective

(a) Utilisation de la silhouette pour définir la zone de recherche (dans l'espace délimité par des frontières) ;

(b) Image de droite dans laquelle la fenêtre de corrélation de 3\*3 pixels (carré rouge) sera positionnée au début de la boîte de recherche ;

(c) La zone de recherche effective des correspondants dans l'image de gauche.

À ce stade, les valeurs stockées dans la matrice des bords peuvent être utilisées pour affiner la zone de recherche. En effet, cette matrice contient tous les pixels appartenant à la bordure de l'objet, et les calculs ne doivent pas dépasser ces limites. L'intérêt principal est donc de passer directement à la phase d'appariement. Pour le calcul des paires correspondantes, une approche locale (qui ne prend en compte qu'une zone entourant chaque point) qui combine les approches « feature matching » et « blocs » a été utilisée pour les calculs des paires correspondantes.

L'appariement de caractéristiques permet de limiter les régions de calcul de la disparité aux caractéristiques globales de l'image sélectionnées ; cela a été fait par l'étape décrite ci-dessus, où nous avons détecté les bords de l'objet. Le principe de base de l'approche par blocs est de choisir une fenêtre bien définie (fenêtre de corrélation) dans la première image et de rechercher sa correspondance dans la deuxième image. Dans ce cas d'étude, nous avons recherché des correspondants uniquement à l'intérieur des bords de l'objet (Figure 79), et le scan a été réalisé dans deux directions (de gauche à droite - et vice versa).

Les objets utilisés contiennent de nombreuses zones homogènes et des parties mal texturées, et donc la recherche de correspondants est assez délicate. Une petite fenêtre de corrélation (3\*3 pixels) a été choisie pour surmonter ce problème.

La méthode utilisée pour réaliser la phase d'appariement est basée sur le SAD<sup>1</sup>. La similarité entre deux blocs distincts est obtenue en calculant la différence absolue entre chaque bloc dans la première image et le bloc correspondant dans la deuxième image. Ce processus est répété pour calculer la différence absolue entre un bloc de la première image et tous les blocs de la même ligne épipolaire dans la deuxième image ; la valeur de différence minimale obtenue signifie que ces deux blocs sont correspondants.

Tous les calculs de mise en correspondance sont effectués uniquement dans les bords précalculés (le résultat Bords de la troisième étape). Au lieu de parcourir toute la largeur et la hauteur des images, on utilise simplement les parties délimitées par les silhouettes des objets (valeurs déjà calculées des limites supérieure, inférieure, gauche et droite). La figure 79 montre comment utiliser la silhouette d'objet pour restreindre la zone de correspondance de l'algorithme stéréoscopique.

Cet algorithme proposé ne recherche pas les correspondances dans l'image entière ; cela réduit le coût de calcul et donne des résultats visuels similaires -dans la plupart des cas- aux résultats des algorithmes stéréoscopiques originaux. Par conséquent, ces étapes éliminent les éventuels artefacts et évitent les éventuelles opérations de post-traitement.

---

#### Algorithm 2 restrictedStereoision

---

```

1: Load image directory
2: Specify the size of the correlation window
3: for  $i \leftarrow 1$  to 120 do
4:   Read(img[i],img[i+1])
5:   Assign the disparity value
6:   for each 3 line do
7:     Calculate the min and max values from the left image
8:     for  $line \leftarrow L_s$  to  $L_i$  do
9:       for  $column \leftarrow min$  to  $max$  do
10:        Place the correlation window in the first image
11:        Use SAD to search the corresponding windows within the
object edges in the other image
12:        Choose the window with the minimum value as the corre-
sponding window
13:        Save result
14:      end for
15:    end for
16:  end for
17: end for

```

---

## 4 Résultats expérimentaux

L'objet à reconstituer était posé sur un plateau tournant. Une caméra fixe a été utilisée pour capturer une vidéo de la scène. La rotation complète de l'objet a fourni un ensemble de 121 images.

### 4.1 Environnement matériel et de programmation

Les vidéos capturées résultent de l'utilisation d'un smartphone SAMSUNG Galaxy Note 5 avec un appareil photo de 16 mégapixels et une distance focale de 28 mm, une ouverture de 1,9 et un objectif de 1 /2.6 capteur. L'ordinateur utilisé est un LENOVO W540 avec un processeur Intel Core i7 vPro 4800MQ 2,70 GHz, 16 Go de Réalité augmentée M et une carte graphique Nvidia Quadro

K2100M. Le système d'exploitation installé sur cette machine est WINDOWS 10 Pro 64 bits. La version 2016 de l'environnement de programmation MATLAB a été choisie pour implémenter les différents algorithmes.

## 4.2 Extraction d'images de la vidéo

Le but de la vidéo est d'extraire un nombre suffisant d'images de haute résolution en quelques secondes. La vidéo filmée avait une durée de 17 secondes et le degré de rotation était de  $3^\circ$ , ce qui nous a permis d'obtenir un ensemble de 121 images pour une rotation complète de chaque objet.

On sait que chaque vidéo est composée d'un jeu de trames ; leur nombre est lié au type de caméra utilisée (7 FPS dans notre cas). Les images extraites sont de la taille de  $640 \times 640$  pixels. La technique utilisée pour extraire les images de la vidéo est assez simple ; chaque cadre est transformé en image '.png' puis enregistré dans un répertoire spécifique.

## 4.3 Extraction de silhouettes

L'objectif principal de cette recherche n'est pas la segmentation d'images, car nous n'avons pas utilisé d'algorithmes de segmentation de haut niveau pour extraire les silhouettes d'un objet. Pour diminuer la complexité de cette tâche, un mini-studio à fond blanc a été utilisé pour faciliter la séparation de l'objet en détectant simplement les différences avec le fond. Malheureusement, nous avons détecté, lors de l'acquisition vidéo, que les effets lumineux et l'environnement extérieur provoquent des bruits et que la couleur de fond n'est pas purement blanche comme souhaité. Cette phase ne s'est pas faite en simplicité car nous avons détecté lors de l'acquisition vidéo que les effets lumineux et l'environnement extérieur provoquent du bruit ; la couleur de fond n'est pas purement blanche comme souhaité.

Ce processus peut être effectué en parallèle ; ainsi, les images sont traitées simultanément, et le temps de calcul de traitement séquentiel de toutes les 121 images colorées a été considérablement réduit.

- Pour le premier objet 'Minion', le temps de traitement séquentiel est d'environ 44 secondes (avec un temps moyen de 0,73 seconde pour chaque photo). Cependant, lors de l'utilisation de l'algorithme parallèle, le temps global devient seulement 17 secondes (avec un temps moyen de 0,28 seconde pour chaque photo). La réduction est alors d'environ 61,36% du temps de traitement séquentiel.
- L'objet 'Schtroumpf' a nécessité un temps global de 65 secondes pour appliquer l'algorithme séquentiel aux 60 photos ; le temps moyen ici est d'environ 1,08 seconde pour chaque image. L'utilisation de l'algorithme parallèle a réduit ce temps à 45 secondes (une moyenne de 0,75 seconde pour chaque photo). Et 31% du temps de traitement a été économisé.

## 4.4 Stéréoscopie sans notre méthode de restriction

Dans un premier temps et pour faire des comparaisons, l'algorithme stéréoscopique a été appliqué à l'ensemble de l'image (comme dans le cas de la plupart des méthodes stéréoscopiques classiques) ; ainsi, la recherche des couples correspondants s'effectue par un balayage horizontal et vertical de tous les pixels de l'image. De ce fait, la zone de recherche est plus grande que l'objet à reconstruire, ce qui implique une augmentation du temps de calcul nécessaire.

- Pour l'objet Minion, le temps global nécessaire (sans utiliser notre méthode de restriction) pour l'application de l'algorithme stéréoscopique sur les 120 images traitées est de 1168,49 secondes (un temps moyen de 9,73 secondes pour chaque image).

- L'objet Schtroumpf a nécessité un temps total de 1152,686 secondes pour appliquer l'algorithme stéréoscopique aux 120 images (lors de l'utilisation de la méthode sans restriction). Le temps moyen ici est d'environ 9,61 secondes pour chaque image.

### 4.5 Application de notre approche restrictive

L'efficacité de l'algorithme proposé a été vérifiée en l'appliquant au calcul de correspondants utilisant les 120 images (dimensionnées par 640\*640 pixels) des mêmes objets mais avec la limitation à leurs silhouettes.

Les résultats suivants ont été obtenus : seulement

- 351,6 secondes (y compris le temps de pré-traitement) ont été nécessaires pour appliquer la stéréoscopie à toutes les 120 images de l'objet minion. Le temps de traitement moyen pour chaque image était de 2,93 secondes.
- 354 secondes (y compris le temps de pré-traitement) est le temps de traitement stéréoscopique des 120 photos de l'objet Schtroumpf. Le temps moyen ici est d'environ 2,95 secondes pour chacun d'eux.

### 4.6 Comparaison

En analysant les résultats obtenus, il est clair qu'un temps considérable est gagné en appliquant notre méthode de restriction au processus de recherche de correspondants. Les figures 80 à 81 montrent graphiquement les différences de temps lors de l'application de l'algorithme stéréoscopique à l'image entière et lors de son application en conjonction avec la méthode proposée, où les figures 9 (a) et 10 (a) représentent le temps total pour traiter les 120 images de chaque objet, et les figures 9 (b) et 10 (b) représentent le temps moyen de traitement de chaque image.

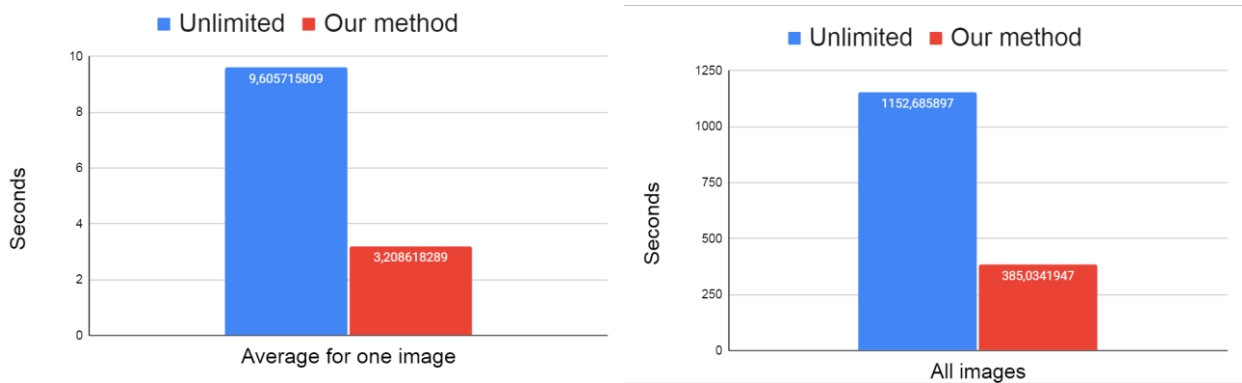


Figure 80 : Résultats obtenus pour l'objet 'Minion'

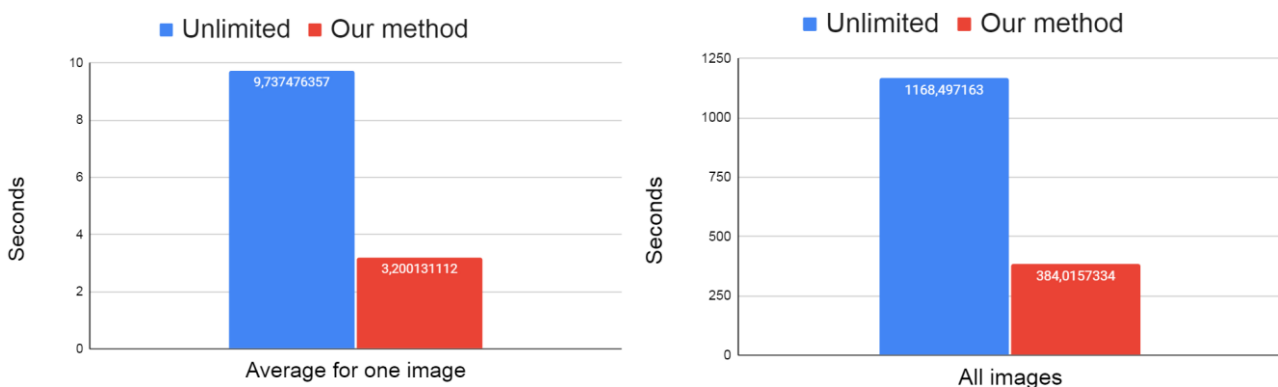
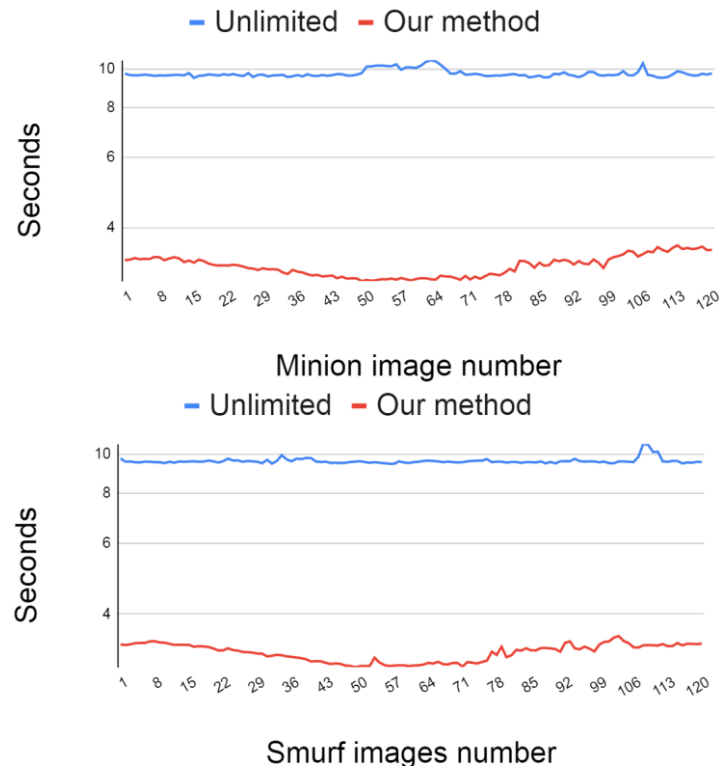


Figure 81 : Résultats obtenus pour l'objet 'Schtroumpf'

- Les figures 80 (b) et 81 (b) montrent que le taux de réduction du temps atteint 67,14% pour l'objet 'Minion' (et 66,59% pour l'objet 'Schtroumpf') lorsque nous utilisons notre proposition pour restreindre la zone de recherche. La couleur bleue indique le temps global nécessaire pour appliquer l'algorithme stéréoscopique d'origine aux 120 images. La couleur rouge représente le temps global d'utilisation de notre méthode.
- Les figures 80(a) et 81 (a) montrent le temps de calcul moyen pour chaque image. La couleur bleue correspond au temps de l'algorithme stéréoscopique d'origine, et le rouge correspond au temps calculé lors de l'application de notre approche.



**Figure 82 : Détails du temps obtenu lors de l'application de la stéréoscopie à chaque image de les 120 images, résultats (a) pour l'objet Minion et (b) pour l'objet Schtroumpf**

- La figure 82 montre bien la différence des temps de calcul de correspondants (pour chaque image) entre l'algorithme stéréoscopique libre et celui limité par la méthode que nous proposons.
- Les lignes rouges (application de notre méthode) n'ont pas dépassé le seuil de 3,61 secondes pour l'objet 'Minion' et 3,53 secondes pour l'objet 'Schtroumpf'.
- Les lignes bleues (l'algorithme original) ne descendent pas en dessous de 9,49 secondes pour 'Minion' (et 9,46 secondes pour 'Smurf') et atteignent jusqu'à 10,4 secondes.

La comparaison confirme notre hypothèse et prouve que la méthode proposée réduit considérablement le temps des calculs de correspondants lors de la reconstruction d'un seul objet à l'intérieur d'une seine.

## 4.7 Carte des disparités

Les cartes des disparités obtenues en appliquant l'algorithme original et celle obtenue après l'application de notre proposition sont presque les mêmes, et aucune différence visuelle n'est détectée ; cela est principalement dû au fond blanc des images acquises.

L'algorithme proposé élimine les éventuels artefacts pouvant apparaître lors de l'utilisation d'un arrière-plan non uniforme puisque les calculs étaient limités à la zone exacte à l'intérieur des silhouettes. La figure 83 présente deux exemples de cartes des disparités obtenues en appliquant notre méthode.

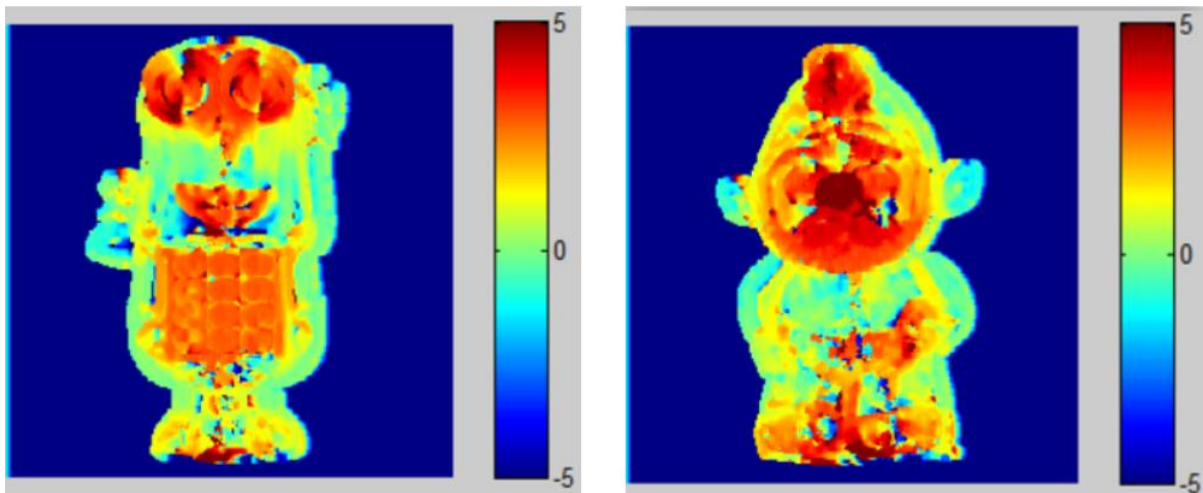


Figure 83 : Cartes de disparités obtenues

## 5 Discussion

L'augmentation des algorithmes de stéréovision est notre objectif de travail ; nous visons pour notre proposition à minimiser le temps de reconstruction de l'objet. Le principe est de restreindre l'espace de recherche à une petite partie des images pour optimiser l'algorithme d'appariement.

La méthode proposée repose sur deux idées principales pour minimiser les temps de calcul : La première est de limiter le calcul de la carte de disparités ; en utilisant les informations géométriques présentes dans les images. La seconde consiste à réduire de moitié le nombre de silhouettes analysées.

Nous pouvons résumer les avantages suivants :

- Utilisation d'objets réels ;
- Simplicité de mise œuvre de l'approche proposée ;
- Accélération du processus de mise en correspondance ;
- Réduction des éventuelles correspondances erronées ;
- Réduction des artefacts pouvant apparaître dans le résultat.

La principale préoccupation de notre approche proposée est la récupération de la forme 3D complète d'un objet. Pour cela, l'objet doit effectuer une rotation complète pour obtenir un maximum de détails, et c'est pourquoi nous avons utilisé le mini-studio, le plateau tournant et la caméra d'enregistrement vidéo.

Tout d'abord, toutes les photos d'objets ont été utilisées, et les temps de calcul étaient importants. Ainsi, une redondance de calculs a été détectée. Pour éliminer les calculs répétés, le nombre de photos a été réduit de moitié en prenant en compte la symétrie inversée mentionnée. Cette réduction a donné 60 silhouettes en 44 secondes pour l'objet 'Minion' et 65 secondes pour l'objet 'Schtroumpf'. Mais les temps de calcul restent inacceptables. Par conséquent, un parallélisme partiel a été introduit pour traiter les 3 matrices de couleurs rouge, vert et bleu tandis que tous les autres calculs ont été effectués séquentiellement. Ceci nous a permis de réduire les temps de calcul pour obtenir des résultats en seulement 17 secondes pour l'objet 'Minion' (une réduction de 61,36%) et en seulement 45 secondes (un taux de réduction de 31%) pour l'objet 'Schtroumpf'.

Malgré cette amélioration, nous sommes convaincus que de meilleurs temps peuvent être obtenus en introduisant un parallélisme intégral à différents niveaux de calcul (images, lignes ou pixels, lors de l'extraction de silhouettes, lors du calcul des bordures d'objets et lors de la mise en correspondance).

L'utilisation d'algorithmes spécialisés et de techniques de segmentation conduira probablement à des résultats en temps réel. Par ailleurs, l'introduction d'équipements dédiés tels que les GPU ou les processeurs parallèles permettrait d'améliorer d'une manière encore plus avantageuse le gain en temps de calcul.

Les résultats obtenus et quelques statistiques pour appuyer notre proposition ont été présentés. Après avoir implémenté les deux algorithmes de stéréovision, celui d'origine et celui limité par la méthode proposée, nous avons constaté que notre proposition présente un avantage significatif en termes de réduction du temps nécessaire au processus d'appariement.

## **6 Conclusions**

Dans ce chapitre, nous avons proposé une méthode qui combine deux approches stéréoscopiques locales pour restreindre la zone de recherche des correspondants. La première est la technique de « correspondance des caractéristiques » et la seconde est la technique de « blocs ». L'idée principale est de faire des calculs uniquement dans la zone délimitée par les arêtes des silhouettes.

Dans un premier temps, nous avons détaillé notre démarche de travail, en commençant par l'extraction d'images d'une vidéo d'objets réels à l'aide d'un mini studio. Nous avons également présenté la technique choisie pour isoler l'objet de son arrière-plan et extraire ces silhouettes.

Ensuite, nous avons expliqué l'algorithme restrictif proposé où nous avons utilisé 60 silhouettes et réduit la zone de correspondance des algorithmes stéréoscopiques. Après cela, nous avons présenté les résultats obtenus et quelques statistiques relatives. Par ailleurs, nous avons achevé notre manuscrit par une discussion où nous avons relaté l'ensemble de notre démarche.

En perspective de notre travail, nous prévoyons d'étendre et de généraliser le parallélisme dans le but de le rendre total (à différents niveaux) pour améliorer les temps de calcul. Nous essaierons également d'utiliser des algorithmes avancés dédiés à la segmentation d'images, et nous essayerons d'introduire les capacités de calcul des GPU.

---

# *Conclusion Générale et Perspectives*

---



Nous nous sommes intéressés à la problématique de la reconstruction 3D à partir d'images qui appartiennent au domaine de l'informatique graphique et de la vision par ordinateur, laquelle 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 cette thèse, nous avons présenté quelques notions de base liées au domaine telles que : le système de coordonnées, le processus de 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 coût et avec une meilleure qualité.

La réalité augmentée au profil de l'héritage culturel a fait une partie intégrante de notre thèse. A cet effet, nous avons détaillé les aspects y afférant. Nous avons aussi présenté un état de l'art des recherches dans ce contexte. En effet, cette étude couvre un large éventail d'applications de réalité augmentée dans le contexte de l'héritage culturel étalées sur les 10 dernières années, cette association ayant fait plusieurs articles universitaires et qui utilise comme piliers principaux huit des mots-clés / thèmes de recherche les plus utilisés dans la production scientifique relatives au domaine, à savoir « la reconstruction 3D », « le patrimoine numérique », « le musée virtuel », « l'expérience utilisateur », « l'éducation », « le tourisme », « les jeux » et « le patrimoine culturel immatériel ».

En abordant et investissant ce domaine, nous avons pu élaborer une classification multicritère des différentes méthodes de reconstruction 3D existantes, cette classification ayant pour finalité de bien choisir la méthode de reconstruction 3D qui convient au cas étudié.

Parmi les méthodes *IBMR* existantes, on trouve la méthode de l'enveloppe visuelle et celle de la stéréovision. Après l'étude que nous avons réalisée sur ces deux méthodes, nous avons constaté que ces deux approches ont quelques limitations qui peuvent être traitées si on les combine ensemble pour obtenir des objets 3D plus réalistes et dans des temps minimaux.

Avant de passer à la méthode proposée, nous avons présenté quelques travaux qui utilisent l'idée de la combinaison de l'enveloppe visuelle avec la stéréovision, en détaillant le déroulement de chaque travail et les résultats obtenus.

Dans notre cas, nous avons proposé une méthode hybride de reconstruction 3D hybride qui consiste aussi à combiner l'enveloppe visuelle avec la stéréovision. L'enveloppe visuelle est une première forme de l'objet qui limite au maximum l'espace englobant cet objet. Cette forme est utilisée par la suite dans la phase de la stéréovision pour réduire les calculs inutiles en dehors de la zone de l'objet.

L'approche proposée introduit une forte association entre les bords de silhouettes et les algorithmes stéréoscopiques en utilisant uniquement les informations géométriques présentes dans les images pour restreindre la zone de mise en correspondance. Elle vise essentiellement à limiter la zone d'appariement à la géométrie exacte de l'objet analysé, évitant ainsi des temps de calcul supplémentaires et des bruits indésirables. Nous n'avons pas utilisé d'algorithmes de codage complexe ni d'équipement coûteux, et nous avons obtenu des résultats acceptables en termes de temps et de précision.

Nous proposons ainsi comme contribution de limiter la zone d'appariement des algorithmes stéréoscopiques afin de ne pas faire de calculs en dehors des bords des silhouettes de l'objet à

reconstruire, ce qui est possible en raison de la forte connexion qui relie la géométrie et le modèle 3D. Partant de cette idée, les temps de calcul atteints par l'algorithme proposé ont en fait été réduits par rapport aux algorithmes stéréoscopiques classiques.

Pour la validation de notre approche, nous avons capturé deux vidéos de deux objets réels à l'aide d'un mini studio. L'utilisation de vidéo facilite la récupération instantanée d'un grand nombre d'images, et nous avons extrait un ensemble de 121 photographies pour chaque objet.

Pour l'extraction des silhouettes, nous avons adopté un algorithme parallèle (c'est-à-dire traitant les images d'une façon parallèle) qui permet de séparer l'objet de son arrière-plan, ces silhouettes sont utilisées par la suite pour l'estimation de l'enveloppe visuelle.

L'algorithme de l'enveloppe visuelle implémenté, donne comme résultats des formes très proches de celles des objets à reconstruire et l'utilisateur ne détecte aucune lenteur. Le problème rencontré dans cette approche est qu'elle ne peut pas traiter les zones concaves sur sa surface. Cette anomalie est raffinée avec la stéréovision qui est, pour sa part, optimisée par l'enveloppe visuelle qui limite les correspondances incorrectes en appliquant la recherche seulement aux pixels qui se trouvent dans la bordure de l'enveloppe visuelle. Nous avons obtenu un excellent taux de réduction du temps de calcul qui a atteint 69 % du temps global requis sans limitation de la zone de recherche.

Notre but principal était de réduire de façon significative le coût du traitement. Après l'implémentation d'un algorithme de stéréovision qui utilise la globalité de l'image et notre algorithme qui combine l'enveloppe visuelle avec la stéréovision, nous avons montré que l'hybridation de ces deux méthodes permet de minimiser le temps de calcul avec un taux de 69%.

Ceci n'a pas eu pour conséquences d'uniquement réduire le temps de calcul. En fait, la carte de disparité que nous avons obtenue ne contient aucun bruit ou fausses correspondances, malgré l'utilisation d'objets avec beaucoup de zones homogènes qui sont très difficiles à différencier par le processus de mise en correspondance qui est assez délicat dans le cas de zones homogènes et d'objets mal texturés. D'autre part, l'introduction de la notion de parallélisme a permis d'accélérer quelques tâches comme par exemple l'extraction des silhouettes. De plus, l'utilisation d'un grand nombre de silhouettes nous a permis d'aboutir à des résultats d'enveloppe visuelle satisfaisants.

Parmi les perspectives immédiates de notre travail, nous pouvons citer les suivantes :

- Nous envisageons aussi l'implémentation d'un algorithme d'isolation des objets de leurs arrière-plans sans avoir recours au mini studio.
- Nous espérons compléter l'algorithme de sauvegarde au format obj, en ajoutant des facettes entre les différents sommets. De cette façon, l'objet 3D résultant peut même être imprimé facilement avec des imprimantes 3D.
- Nous comptons aller vers l'implémentation d'une application Android de reconstruction 3D en temps réel.
- La publication d'articles dans le contexte de « la reconstruction 3d du patrimoine culturel et l'utilisation de ces modèles 3D dans le cadre de la réalité augmentée ; fait l'objet de notre plus grande priorité. En effet, nous avons eu l'opportunité de reconstruire plusieurs objets culturels en utilisant des techniques différentes (scanner 3D, photogrammétrie gratuite/payante...), et nous avons même utilisé les modèles 3D obtenus dans des contextes de réalité virtuelle (casque Oculus go) et de réalité augmentée (application Android, casque Oculus Quest 2).
- En fin, nous envisageons de réaliser un musée virtuel en mettant en œuvre nos connaissances.

---

# *Références*

---

## Références

- A.Rauschnabel, P., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior*, 133. ELSEVIER. doi:<https://doi.org/10.1016/j.chb.2022.107289>
- Abate, A., Barra, S., Galeotafiore, G., Díaz, C., Aura, E., Sánchez, M., & Vendrell, E. (2018). An Augmented Reality Mobile App for Museums: Virtual Restoration of a Plate of Glass. *7th International Conference on Digital Heritage, EuroMed 2018*, 11196, pp. 539–547. Springer. doi:[https://doi.org/10.1007/978-3-030-01762-0\\_47](https://doi.org/10.1007/978-3-030-01762-0_47)
- Aggarwal, R. .. (2019). augmented Reality and its effect on our life. *9th International Conference on Cloud Computing, Data Science & Engineering* (pp. 510-515). Computer Science. doi:10.1109/CONFLUENCE.2019.8776989
- Alexandre, B. (2014). Reconstruction automatique de maquettes numériques 3D. *Traitement du signal et de l'image*. Université Paris-Est.
- Amparo, N. A., Felipe Buill, P., Joaquín Regot, M., & Andrés de Mesa, G. (2012). Generation of virtual models of cultural heritage. *Journal of Cultural Heritage*, 13(1), 103-106. doi:10.1016/j.culher.2011.06.004
- Angeline, M., & Richard, H. (2005). Shape from non-homogeneous, non-stationary, anisotropic, perspective texture. *The British Machine Vision Conference* (pp. 69-78). Oxford: Oxford Brookes University.
- Angelopoulou, A., Economou, D., Bouki, V., Psarrou, A., Jin, L., Pritchard, C., & Kolyda, F. (2011). Mobile augmented reality for cultural heritage. *Proceedings of the International Conference on Mobile Wireless Middleware, Operating Systems, and Applications*, 93, pp. 15–22. Springer. doi:[https://doi.org/10.1007/978-3-642-30607-5\\_2](https://doi.org/10.1007/978-3-642-30607-5_2)
- Arena, F., Collotta, M., Pau, G., & Termine, F. (2022). An Overview of Augmented Reality. *Computers*. doi:10.3390/computers11020028
- Asada, N., Fujiwara, H., & Matsuyama, T. (1998). Edge and Depth from Focus. *International Journal of Computer Vision*, 26(2), 153–163. doi:10.1023/A:1007996810301
- Augello, A., Infantino, I., Pilato, G., & Vitale, G. (2021). Site experience enhancement and perspective in cultural heritage fruition—A survey on new technologies and methodologies based on a “four-pillars” approach. *Future Internet*, 13(4). doi:<https://doi.org/10.3390/fi13040092>
- Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & Macintyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications*, 21, pp. 34-47. doi:10.1109/38.963459
- Balduini, M., Celino, I., Dell’Aglio, D., Della Valle, E., Huang, Y., Lee, T., . . . Tresp, V. (2012). BOTTARI: An augmented reality mobile application to deliver personalized and location-based recommendations by continuous analysis of social media streams. *Journal of Web Semantics*, 16, 33–41. doi:<https://doi.org/10.1016/j.websem.2012.06.004>
- Barrile, V., Bilotta, G., Meduri, G., De Carlo, D., & Nunnari, A. (2017). Laser scanner technology, ground-penetrating radar and augmented reality for the survey and recovery of artistic; archaeological and cultural heritage. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 123–127. doi:<https://doi.org/10.5194/isprs-annals-IV-4-W4-123-2017>
- Barrile, V., Nunnari, A., & Ponterio, R. (2016). Laser scanner for the Architectural and Cultural Heritage and Applications for the Dissemination of the 3D Model. *Social and Behavioral Sciences*, 223, pp. 555–560. doi:<https://doi.org/10.1016/j.sbspro.2016.05.342>

- Basaraba, N., Conlan, O., Edmond, J., & Arnds, P. (2019). Digital narrative conventions in heritage trail mobile apps. *New Review of Hypermedia and Multimedia*, 25(1-2), 1-30. doi:<https://doi.org/10.1080/13614568.2019.1642963>
- Baumgart, A., & Bruce G. (1974). Geometric Modeling for Computer Vision. *PhD thesis*. University of Stanford.
- Bec, A., Moyle, B., Timms, K., Schaffer, V., Skavronskaya, L., & Little, C. (2019). Management of immersive heritage tourism experiences: A conceptual model. *Tourism Management*, 72, 117–120. doi:<https://doi.org/10.1016/j.tourman.2018.10.033>
- Bekele, M., & Champion, E. (2019). A Comparison of Immersive Realities and Interaction Methods: Cultural Learning in Virtual Heritage. *Frontiers in Robotics and AI*, 6. doi:<https://doi.org/10.3389/frobt.2019.00091>
- Benoît, L. (2009). Enveloppes Visuelles et Reconstruction 3D.
- Berthold, H. (1989). Obtaining shape from shading information. *Shape from Shading*, 123-171.
- Besbes, B., Collette, S., Tamaazousti, M., Bourgeois, S., & Gay-Bellile, V. (2012). An interactive augmented reality system: A prototype for industrial maintenance training applications. *2012 IEEE International Symposium on Mixed and Augmented Reality* (pp. 269–270). IEEE. doi:10.1109/ISMAR.2012.6402568
- Beury, J.-N. (2014). *Système de coordonnées cartésiennes, cylindriques, sphériques*.
- Blanc, J. (1998). *Synthèse de nouvelles vues d'une scène 3D à partir d'images existantes*. Grenoble: Institut National Polytechnique de Grenoble.
- Blanco-Pons, S., Carrión-Ruiz, B., Lerma, J., & Villaverde, V. (2019). Design and implementation of an augmented reality application for rock art visualization in Cova dels Cavalls (Spain). *Journal of Cultural Heritage*, 177–185. doi:<https://doi.org/10.1016/j.culher.2019.03.014>
- Boboc, R. G., Elena, B., Florin, G., Norina, P., & Dorin-Mircea, P. (2022). Augmented Reality in Cultural Heritage: An Overview of the Last Decade of Applications. *Applied Sciences*, 12(9). doi:<https://doi.org/10.3390/app12199859>
- Boboc, R., Duguleană, M., Voinea, G.-D., Postelnicu, C.-C., Popovici, D.-M., & Carrozzino, M. (2019). Mobile Augmented Reality for Cultural Heritage: Following the Footsteps of Ovid among Different Locations in Europe. *Sustainability*, 11(4). doi:<https://doi.org/10.3390/su11041167>
- Bottino, A., & Laurentini, A. (2001). A Silhouette Based Technique for the Reconstruction of Human Movement. *Computer Vision and Image Understanding*, 83(1), 79-95. doi:<https://doi.org/10.1006/cviu.2001.0918>
- Boufama, B. (1994). *Reconstruction tridimensionnelle en vision par ordinateur : cas des cameras non etalonnees*. Grenoble: Institut National Polytechnique de Grenoble.
- Brandou, V. (2008). *Stereo Vision and 3D/4D Reconstruction*. Nice: Nice Sophia-Antipolis.
- Bujari, A., Ciman, M., & Gaggi, O. (2017). Using gamification to discover cultural heritage locations from geo-tagged photos. *Personal and Ubiquitous Computing*, 21, 235–252. doi:<https://doi.org/10.1007/s00779-016-0989-6>
- Butnariu, S., Duguleana, M., Brondi, R., Florin, G., Postelnicu, C., & Carrozzino, M. (2018). An interactive haptic system for experiencing traditional archery. *Acta Polytechnica Hungarica*, 15(5).

- Campi, M., di Luggo, A., Palomba, D., & Palomba, R. (2019). DIGITAL SURVEYS AND 3D RECONSTRUCTIONS FOR AUGMENTED ACCESSIBILITY OF ARCHAEOLOGICAL HERITAGE. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 205–212. doi: <https://doi.org/10.5194/isprs-archives-XLII-2-W9-205-2019>
- Camurri, A., & Volpe, G. (2016). The intersection of art and technology. *IEEE MultiMedia*, 23(1), 10–17.
- Ceccacci, S., Generosi, A., Leopardi, A., Mengoni, M., & Mandorli, A. (2021). The role of haptic feedback and gamification in virtual museum systems. *Journal on Computing and Cultural Heritage*, 14(3), 1–14. doi:<https://doi.org/10.1145/3453074>
- Ch'ng, E., Cai, S., Leow, F.-T., & Zhang, T. (2019). Adoption and use of emerging cultural technologies in China's museums. *Journal of Cultural Heritage*, 37, 170–180. doi:<https://doi.org/10.1016/j.culher.2018.11.016>
- Chatzigrigoriou, P., Nikolakopoulou, V., Vakkas, T., Vosinakis, S., & Koutsabasis, P. (2021). Is Architecture Connected with Intangible Cultural Heritage? Reflections from Architectural Digital Documentation and Interactive Application Design in Three Aegean Islands. (MDPI, Éd.) *Heritage*, 4(2), 664-689. doi: <https://doi.org/10.3390/heritage4020038>
- Chehata, N., Jung, F., Pierrot-Deseilligny, M., & Stamon, G. (2004). Détection et reconstruction de facettes 3D par approche hiérarchique par régions à partir de couples d'images satellite en milieu urbain.
- Comunità, M., Gerino, A., Lim, V., & Picinali, L. (2021). Design and Evaluation of a Web- and Mobile-Based Binaural Audio Platform for Cultural Heritage. *Applied Sciences*, 11(4), 15-40. doi: <https://doi.org/10.3390/app11041540>
- Crowley, J. (2010). *Formation et Analyse d'Images*.
- Cruz, D., Sevilla, J., San Gabriel, J., Cruz, A., & Caselis, E. (2018). Design and Development of Augmented Reality (AR) Mobile Application for Malolos' Kameztizuhan (Malolos Heritage Town, Philippines). *IEEE Games, Entertainment, Media Conference* (pp. 1-9). IEEE. doi:10.1109/GEM.2018.8516272
- Cruz-Neira, C., Sandin, D., DeFanti, T., Kenyon, R., & Hart, J. (1992). The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6), 64–72. doi:<https://doi.org/10.1145/129888.129892>
- Cultural heritage at risk: United States*. (2022, 12 10). Récupéré sur [www.khanacademy.org](https://www.khanacademy.org/humanities/special-topics-art-history/arches-at-risk-cultural-heritage-education-series/xa0148fd6a60f2ff6:cultural-heritage-endangered-round-the-world/a/cultural-heritage-at-risk-united-states): <https://www.khanacademy.org/humanities/special-topics-art-history/arches-at-risk-cultural-heritage-education-series/xa0148fd6a60f2ff6:cultural-heritage-endangered-round-the-world/a/cultural-heritage-at-risk-united-states>
- Czernuszenko, M., Pape, D., Sandin, D., DeFanti, T., Dawe, G., & Brown, M. (1997). The ImmersaDesk and Infinity Wall projection-based virtual reality displays. *ACM SIGGRAPH Computer Graphics*, 31(2), 46–49. doi:<https://doi.org/10.1145/271283.271303>
- Damala, A., Hornecker, E., Van der Vaart, M., van Dijk, D., & Ruthven, I. (2016). The Loupe: Tangible augmented reality for learning to look at Ancient Greek Art. *editerranean Archaeology and Archaeometry*, 16(5), 73–85. doi:10.5281/zenodo.204970
- Damien, P. (2004). *Renderu en temps réel de scène complexe*.
- Debevec, P., Taylor, C., & Malik, J. (1996). Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach. *SIGGRAPH '96, proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, 11-20. doi:<https://doi.org/10.1145/237170.237191>

- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, (pp. 9–15). doi:<https://doi.org/10.1145/2181037.2181040>
- Djaber, R., & Chaouki, B. M. (2011). Multi-criteria classification of 3D reconstruction approaches. *ICIST 2011, 1st International Conference on Information Systems and Technologies*. Tebessa: IEEE.
- Djaber, R., & Nour Eddine, D. (2022). A combination of 'feature mapping' and 'block' approaches to reduce the matching area of stereoscopic algorithms. *International Journal of Computational Vision and Robotics*. doi:10.1504/IJCVR.2022.10048705
- Dyer, C. (2001). Volumetric Scene Reconstruction From Multiple Views. doi:10.1007/978-1-4615-1529-6\_16
- El-Hakim, S., Beraldin, J.-A., Picard, M., & Godin, G. (2004). Detailed 3D reconstruction of large-scale heritage sites with integrated techniques. *IEEE Computer Graphics and Applications*, 24(3), 21-29. doi:10.1109/MCG.2004.1318815
- Etzeberria, A., Asensio, M., Vicent, N., & Cuenca, J. (2012). Mobile devices: A tool for tourism and learning at archaeological sites. *International Journal of Web Based Communities*, 8(1), 57–72. doi:<https://doi.org/10.1504/IJWBC.2012.044682>
- Evangelidis, K., Sylaiou, S., & Papadopoulos, T. (2020). Mergin' Mode: Mixed Reality and Geoinformatics for Monument Demonstration. *Applied Sciences*, 10(11). doi:<https://doi.org/10.3390/app10113826>
- Fakhfakh, N. (2011). *Détection et localisation tridimensionnelle par stéréovision d'objets en mouvement dans des environnements complexes : application aux passages à niveau*.
- Fakhfakh, N., Debled-Rennesson, I., Domenjoud, E., Kerautret, B., & Even, P. (1999). Discrete Geometry for Computer Imagery. *8th International Conference, DGCI'99*. Marne-la-Vallee. doi:10.1007/978-3-642-19867-0
- Fenu, C., & Pittarello, F. (2018). Svevo tour: The design and the experimentation of an augmented reality application for engaging visitors of a literary museum. *International Journal of Human-Computer Studies*, 114, 20–35. doi:<https://doi.org/10.1016/j.ijhcs.2018.01.009>
- Ferdani, D., Fanini, B., Piccioli, M., Carboni, F., & Vigliarolo, P. (2020). 3D reconstruction and validation of historical background for immersive VR applications and games: The case study of the Forum of Augustus in Rome. *Journal of Cultural Heritage*, 43, 129–143. doi:<https://doi.org/10.1016/j.culher.2019.12.004>
- Frank, H., Bernhard, E., & Benjamin, N. (2000). Reconstruction and virtual model of the Schickard calculator. *Journal of Cultural Heritage*, 1(4), 335-340. doi:[https://doi.org/10.1016/S1296-2074\(00\)01090-6](https://doi.org/10.1016/S1296-2074(00)01090-6)
- Frédéric, J., Noirfalise, E., & Sturm, P. (2011). Calibrage itératif de caméras à partir de scènes planes. *14ème Congrès AFCET de Reconnaissance des Formes et Intelligence Artificielle*.
- Frischer, B., Abernathy, D., Giuliani, F., Scott, R., & Ziemssen, H. (2006). A New Digital Model of the Roman Forum. *Journal of Roman Archaeology*, 163–182. Consulté le 12 2022, 2022, sur [https://www.academia.edu/36574837/Frischer\\_et\\_al\\_Roman\\_Forum\\_2006\\_pdf](https://www.academia.edu/36574837/Frischer_et_al_Roman_Forum_2006_pdf)
- Fritsch, D., & Klein, M. (2018). 3D preservation of buildings – Reconstructing the past. *Multimedia Tools and Applications*, 9153–9170. doi:<https://doi.org/10.1007/s11042-017-4654-5>
- Gaitatzes, A., Christopoulos, D., Voulgari, A., & Roussou, M. (2000). Hellenic Cultural Heritage through Immersive Virtual Archaeology. *6th International Conference on Virtual Systems and Multimedia*, (pp. 57–64).

- Gales, G. (2011). *Mise en correspondance de pixels pour la stéréovision binoculaire par propagation d'appariements de points d'intérêt et sondage de régions*.
- Galmangoda, G., Gajanayake, P., Indika, K., Jayaweera, Y., & Raj, N. (2016). Augmented Reality to Reconstruct Sri Lankan Cultural Heritage in Prime State: HeladivaAR. *Symposium on Statistical & Computational Modelling with Applications*, (pp. 40-44). doi:10.13140/RG.2.2.11936.81926
- Gao, Q., Wan, T., Tang, W., & Chen, L. (2017). A Stable and Accurate Marker-Less Augmented Reality Registration Method. *2017 International Conference on Cyberworlds*, (pp. 41-47). doi:10.1109/CW.2017.44
- Garau, C. (2014). From Territory to Smartphone: Smart Fruition of Cultural Heritage for Dynamic Tourism Development. *Planning Practice & Research*, 29(3), 238–255. doi:https://doi.org/10.1080/02697459.2014.929837
- Garau, C., & Ilardi, E. (2014). The “Non-Places” Meet the “Places”: Virtual Tours on Smartphones for the Enhancement of Cultural Heritage. *Journal of Urban Technology*, 21(1), 79–91. doi:https://doi.org/10.1080/10630732.2014.884384
- Gherghina, A., Olteanu, A., & Tapus, N. (2013). A marker-based augmented reality system for mobile devices. *11th RoEduNet International Conference*, (pp. 1-6). doi:10.1109/RoEduNet.2013.6511731
- Girbacia, F., Butnariu, S., Orman, A., & Postelnicu, C. (2013). Virtual restoration of deteriorated religious heritage objects using augmented reality technologies. *European Journal of Science and Theology*, 9(2), 223–231.
- Gloria, H. (2012). Shape from Silhouette Consensus. *Pattern Recognition*, 3231-3244.
- Gomes., L., Bellon., O. R., & Silva, L. (2014). 3D reconstruction methods for digital preservation of cultural heritage: A survey. *Pattern Recognition Letters*, 50, 3-14. doi:https://doi.org/10.1016/j.patrec.2014.03.023
- González Vargas, J., Fabregat, R., Carrillo-Ramos, A., & Jové, T. (2020). Survey: Using Augmented Reality to Improve Learning Motivation in Cultural Heritage Studies. *Applied Sciences*, 10(3). doi:https://doi.org/10.3390/app10030897
- González-Rodríguez, M., Díaz-Fernández, M., & Pino-Mejías, M. (2020). The impact of virtual reality technology on tourists' experience: A textual data analysis. *Soft Computing*, 24, 13879–13892.
- Hammady, R., Ma, M., & Temple, N. (2016). Augmented Reality and Gamification in Heritage Museums. *Lecture Notes in Computer Science()*, 9894, 181–187. doi:https://doi.org/10.1007/978-3-319-45841-0\_17
- Han, D.-I., tom Dieck, M., & Jung, T. (2018). User experience model for augmented reality applications in urban heritage tourism. *Journal of Heritage Tourism*, 13(1), 46–61. doi:https://doi.org/10.1080/1743873X.2016.1251931
- Han, D.-I., Tom Dieck, M., & Jung, T. (2019). Augmented Reality Smart Glasses (ARSG) visitor adoption in cultural tourism. *Leisure Studies*, 38(5), 618–633. doi:https://doi.org/10.1080/02614367.2019.1604790
- Hartley, R. I., & Sturm, P. (1997). Triangulation. *Computer Vision and Image Understanding*, (pp. 146-157). doi:https://doi.org/10.1006/cviu.1997.0547
- Hauser, H., Beisswenger, C., Partarakis, N., Zabulis, X., Adami, I., Zidianakis, E., . . . al., e. (2022). Multimodal narratives for the presentation of silk heritage in the museum. *Heritage*, 5(1), 461–487. doi:https://doi.org/10.3390/heritage5010027



- Hoang, T., & Cox, T. (2018). Alternating reality: An interweaving narrative of physical and virtual cultural exhibitions. *Presence: Teleoperators and Virtual Environments*, 26(4), 402–419. doi:[https://doi.org/10.1162/PRES\\_a\\_00307](https://doi.org/10.1162/PRES_a_00307)
- Hoshang, K., & Mohd, S. S. (2013). Survey of shadow volume algorithms in computer graphics. *IETE Technical Review*, 30(1), 38-46. doi:10.4103/0256-4602.107338
- Hoshang, K., Abdennour, E. R., Mohd, S. S., & Tanzila, S. (2018). ReVitAge: Realistic virtual heritage taking shadows and sky illumination into account. *Journal of Cultural Heritage*, 32, 166-175. doi:<https://doi.org/10.1016/j.culher.2018.01.020>
- Huang, W., Xiang, H., & Li, S. (2019). The application of augmented reality and unity 3D in interaction with intangible cultural heritage. *Evolutionary Intelligence*, 12, 1-9. doi:<https://doi.org/10.1007/s12065-019-00314-6>
- Ibañez-Etxeberria, A., Gómez-Carrasco, C., Fontal, O., & García-Ceballos, S. (2020). Virtual environments and augmented reality applied to heritage education. *Applied Sciences*, 10(7). doi:<https://doi.org/10.3390/app10072352>
- Indrawan, I., a Bayupati, I. P., & Putri, D. (2018). Markerless Augmented Reality Utilizing Gyroscope to Demonstrate the Position of Dewata Nawa Sanga. *International Journal of Interactive Mobile Technologies*, 12(1), 19-35. doi:<https://doi.org/10.3991/ijim.v12i1.7527>
- Jacquemin, C., Caye, V., Luca, L., & Favre-Brun, A. (2014). Genius Loci: Digital heritage augmentation for immersive performance. *International Journal of Arts and Technology*, 7(2-3), 223–246. doi:<https://doi.org/10.1504/IJART.2014.060942>
- Jaffar, E., Abubakar, J., & Zulkifli, A. N. (2018). Mobile augmented reality for hearing impaired museum engagement (MARHIME): A conceptual model. *AIP Conference Proceedings*, (pp. 79–96). doi:10.1063/1.5055433
- Jofresa, R., Xirau, M., Eredam, H., & Vicente, O. (2019). Gamification and Cultural Heritage. *Mixed reality and gamification for cultural heritage*, 2.
- Joo-Nagata, J., Abad, F., Giner, J., & García-Peñalvo, F. (2017). Augmented reality and pedestrian navigation through its implementation in m-learning and e-learning: Evaluation of an educational program in Chile. *Computers & Education*, 111. doi:<https://doi.org/10.1016/j.compedu.2017.04.003>
- Jung, K., Nguyen, V., Piscarac, D., & Yoo, S. (2020). Meet the virtual jeju dol harubang—The mixed VR/AR application for cultural immersion in Korea’s main heritage. *ISPRS International Journal of Geo-Information*, 9(6). doi:<https://doi.org/10.3390/ijgi9060367>
- Jung, T., Lee, H., Chung, N., & Tom Dieck, M. (2018). Cross-cultural differences in adopting mobile augmented reality at cultural heritage tourism sites. *International Journal of Contemporary Hospitality Management*, 30(3), 1621–1645. doi:<https://doi.org/10.1108/IJCHM-02-2017-0084>
- Kallergis, G., Christoulakis, M., Diakakis, A., Ioannidis, M., Paterakis, I., Manoudaki, N., . . . Oungrinis, K. (2020). Open City Museum: Unveiling the Cultural Heritage of Athens Through an-Augmented Reality Based-Time Leap. *Culture and Computing*, 156–171. doi:[https://doi.org/10.1007/978-3-030-50267-6\\_13](https://doi.org/10.1007/978-3-030-50267-6_13)
- Khan, M. (2014). MUSE: Understanding traditional dances. In *Proceedings of the 2014 IEEE Virtual Reality (VR)*. *Proceedings of the 2014 IEEE Virtual Reality (VR)*, (pp. 173–174). doi:10.1109/VR.2014.6802107
- Koleva, B., Schnädelbach, H., Benford, S., & Greenhalgh, C. (2000). Traversable interfaces between real and virtual worlds. *the CHI 2000 Conference on Human factors in computing systems*, (pp. 233-240). doi:10.1145/332040.332437

- Kyriakou, P., & Hermon, S. (2019). Can I touch this? Using natural interaction in a museum augmented reality system. *Digital Applications in Archaeology and Cultural Heritage*, 12. doi:https://doi.org/10.1016/j.daach.2018.e00088
- Landers, R., Auer, E., Collmus, A., & Armstrong, M. (2018). Gamification science, its history and future: Definitions and a research agenda. *Simulation & Gaming*, 49(3), 315–337. doi:https://doi.org/10.1177/1046878118774385
- Laštovička-Medin, G. (2019). The Materiality of Interaction & Intangible Heritage: Interaction Design. *Proceedings of the 2019 8th Mediterranean Conference on Embedded Computing*, 10, pp. 1–6. doi:10.1109/MECO.2019.8760092
- Laudazi, A., & Boccaccini, R. (2014). Augmented museums through mobile apps. *Proceedings of the Workshop on Horizon2020 and Creative Europe vs Digital Heritage: A European Projects Crossover. Flash News Co-Located with the International Conference Museums and the We*, 1336, pp. 12–17.
- Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16, pp. 150-162. doi:10.1109/34.273735
- Lee, J., Lee, H., Jeong, D., Lee, J., Kim, T., & Lee, J. (2021). Developing Museum Education Content: AR Blended Learning. *International Journal of Art & Design Education*, 40(3), 473–491. doi:https://doi.org/10.1111/jade.12352
- Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., . . . Inc, C. (2001). The Digital Michelangelo Project: 3D Scanning of Large Statues. *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*.
- Li, M., Schirmacher, H., H., M. M., & Siedel, -P. (2002). Combining stereo and visual hull information for on-line reconstruction and rendering of dynamic scenes. *EEE Workshop on Multimedia Signal Processing* (pp. 9-12). IEEE. doi:10.1109/MMSP.2002.1203235
- Liarokapis, F., Petridis, P., Andrews, D., & De Freitas, S. (2017). Multimodal serious games technologies for cultural heritage. *Mixed Reality and Gamification for Cultural Heritage*, 371–392. doi:https://doi.org/10.1007/978-3-319-49607-8\_15
- Litvak, E., & Kuflik, T. (2020). Enhancing cultural heritage outdoor experience with augmented-reality smart glasses. *Personal and Ubiquitous Computing*, 24, 873–886. doi:https://doi.org/10.1007/s00779-020-01366-7
- Liu, E., Liu, C., Yang, Y., Guo, S., & Cai, S. (2018). Design and Implementation of an Augmented Reality Application with an English Learning Lesson. *Proceedings of the 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering*, (pp. 494–499). doi:10.1109/TALE.2018.8615384
- Loh, A., & Hartley, R. (2005). Shape from Non-homogeneous, Non-stationary, Anisotropic, Perspective Texture. *Proceedings of the British Machine Vision Conference*.
- Lu, W., Wang, M., & Chen, H. (2020). Research on Intangible Cultural Heritage Protection Based on Augmented Reality Technology. *First International Conference on Computer Applied Science and Information Technology*, 1574. doi:10.1088/1742-6596/1574/1/012026
- Machidon, O., Postelnicu, C., & Girbacia, F. (2016). 3D Reconstruction as a Service – Applications in Virtual Cultural Heritage. *Augmented Reality, Virtual Reality, and Computer Graphics*, 9769. doi:https://doi.org/10.1007/978-3-319-40651-0\_21

- Malik, U., Tissen, L., & Vermeeren, A. (2021). 3D Reproductions of Cultural Heritage Artefacts: Evaluation of significance and experience. *Studies in Digital Heritage*, 5(1), 1–29. doi:<https://doi.org/10.14434/sdh.v5i1.32323>
- Marasco, A., Buonincontri, P., van Niekerk, M., Orlowski, M., & Okumus, F. (2018). Exploring the role of next-generation virtual technologies in destination marketing. *Journal of Destination Marketing & Management*, 9, 138–148. doi:<https://doi.org/10.1016/j.jdmm.2017.12.002>
- Marr, D. (2010). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. doi:<https://doi.org/10.7551/mitpress/9780262514620.001.0001>
- Marto, A., Gonçalves, A., Melo, M., & Bessa, M. (2022). A survey of multisensory vr and ar applications for cultural heritage. *102*, 426–440. doi:<https://doi.org/10.1016/j.cag.2021.10.001>
- Mason, M. (2016). The MIT museum glassware prototype: Visitor experience exploration for designing smart glasses. *Journal on Computing and Cultural Heritage*, 9(3), 1–28. doi:<https://doi.org/10.1145/2872278>
- Matsuyama, T., Nobuhara, S., Takai, T., & Tung, T. (2012). *3D Video and Its Applications*. doi:10.1007/978-1-4471-4120-4
- May, P., Schmidt, W., & Vlachopoulos, D. (2020). The Use of Augmented Reality (AR) In Museum Education: A Systematic Literature Review. *Proceedings of the 14th International Technology, Education and Development Conference*, (pp. 3179-3186).
- Merchán, M., Merchán, P., & Pérez, E. (2021). Good Practices in the Use of Augmented Reality for the Dissemination of Architectural Heritage of Rural Areas. *Applied Sciences*, 11(5). doi:<https://doi.org/10.3390/app11052055>
- Meyer, A., Briceño, H., & Bouakaz, S. (2007). User-Guided Shape from Shading to Reconstruct Fine Details from a Single Photograph. *Computer Vision*.
- Monga, O., & Horaud, R. (1993). Vision par ordinateur, outils fondamentaux.
- Mortara, M., Catalano, C., Bellotti, F., Fiucci, G., Houry-Panchetti, M., & Petridis, P. (2014). Learning cultural heritage by serious games. *Journal of Cultural Heritage*, 15(3), 318–325. doi:<https://doi.org/10.1016/j.culher.2013.04.004>
- Münzer, M. (2020). How can augmented reality improve the user experience of digital products and engagement with cultural heritage outside the museum space? *International Conference Florence Heri-tech: the Future of Heritage Science and Technologies*, 949. doi:10.1088/1757-899X/949/1/012040
- Neto, M. (2002). Image-Based Modeling and Rendering Techniques: A Survey. *RITA*, 37-66.
- Nisi, V., Cesario, V., & Nunes, N. (2019). Augmented reality museum's gaming for digital natives: Haunted encounters in the Carvalhal's palace. *Joint International Conference on Entertainment Computing and Serious Games*, 11863, pp. 28-41. doi:[https://doi.org/10.1007/978-3-030-34644-7\\_3](https://doi.org/10.1007/978-3-030-34644-7_3)
- Nister, D. (2004). Automatic passive recovery of 3D from images and video. *Proceedings. 2nd International Symposium on 3D Data Processing, Visualization and Transmission*, (pp. 438-445). doi:10.1109/TDPVT.2004.1335271
- Noh, Z., Sunar, M. S., & Pan, Z. (2009). A review on augmented reality for virtual heritage system. *International conference on technologies for E-learning and digital entertainment* (pp. 50-61). Springer.

- Novotný, M., Lacko, J., & Samuelčík, M. (2013). Applications of multi-touch augmented reality system in education and presentation of virtual heritage. *25*, 231–235. doi:<https://doi.org/10.1016/j.procs.2013.11.028>
- Nowak, E. (2003). *Amélioration d'une reconstruction 3D par voxels à l'aide de lastéréo-vision*. INREA.
- Okanovic, V., Ivkovic-Kihic, I., Boskovic, D., Mijatovic, B., Prazina, I., Skaljo, E., & Rizvic, S. (2022). Interaction in eXtended Reality Applications for Cultural Heritage. *Applied Sciences*, *12*. doi:[doi:10.3390/app12031241](https://doi.org/10.3390/app12031241)
- Paliokas, I., Patenidis, A., Mitsopoulou, E., Tsita, C., Pehlivanides, G., Karyati, E., . . . al., e. (2020). A Gamified Augmented Reality Application for Digital Heritage and Tourism. *Applied Sciences*, *10*. doi:<https://doi.org/10.3390/app10217868>
- Pang, Y., Yuan, M., Nee, A., Ong, S., & Youcef-Toumi, K. (2006). A markerless registration method for augmented reality based on affine properties. *Conference on Research and Practice in Information Technology*, *50*, pp. 15-22.
- Panou, C., Ragia, L., & Dimelli, D. a. (2018). Outdoors Mobile Augmented Reality Application Visualizing 3D Reconstructed Historical Monuments. *4th International Conference on Geographical Information Systems Theory, Applications and Managemen* (pp. 59-67). SciTePress. doi:[10.5220/0006701800590067](https://doi.org/10.5220/0006701800590067)
- Parfenov, V., Igoshin, S., Masaylo, D., Orlov, A., & Kuliashou, D. (2022). Use of 3D Laser Scanning and Additive Technologies for Reconstruction of Damaged and Destroyed Cultural Heritage Objects. *Quantum Beam Science*, *6*(1), 11. doi:<https://doi.org/10.3390/qubs6010011>
- Partarakis, N., Antona, M., Zidianakis, E., & Stephanidis, C. (2016). Adaptation and Content Personalization in the Context of Multi User Museum Exhibits. *1st Workshop on Advanced Visual Interfaces for Cultural Heritage*, (pp. 5-10).
- Partarakis, N., Zabulis, X., Foukarakis, M., Moutsaki, M., Zidianakis, E., Patakos, A., & Tasiopoulou, E. (2022). Supporting sign language narrations in the museum. *Heritage*, *5*(1), 1–20. doi:<https://doi.org/10.3390/heritage5010001>
- Partarakis, N., Zidianakis, E., Antona, M., & Stephanidis, C. (2015). Art and Coffee in the Museum. Dans *Distributed, Ambient, and Pervasive Interactions* (pp. 370–381). Springer. doi:[https://doi.org/10.1007/978-3-319-20804-6\\_34](https://doi.org/10.1007/978-3-319-20804-6_34)
- Patirupanusara, P. (2012). Marker-Based Augmented Reality Magic Book for Anatomical Education. *International Conference on Computer and Communication Technologies*, (pp. 136-138).
- Patti, I. (2020). Standard Cataloguing of Augmented Objects for a Design Museum. *International Conference Florence Heri-tech: the Future of Heritage Science and Technologies*, 949. doi:[10.1088/1757-899X/949/1/012054](https://doi.org/10.1088/1757-899X/949/1/012054)
- Pietroni, E. (2012). An augmented experiences in cultural heritage through mobile devices: “Matera tales of a city” project. *International Conference on Virtual Systems and Multimedia*, (pp. 117–124). IEEE. doi:[10.1109/VSM.2012.6365915](https://doi.org/10.1109/VSM.2012.6365915)
- Popovici, D.-M., Iordache, D., Comes, R., Neamțu, C., & Băutu, E. (2022). Interactive Exploration of Virtual Heritage by Means of Natural Gestures. *Applied Sciences*, *12*. doi:<https://doi.org/10.3390/app12094452>

- Portalés, C., Lerma, J. L., & Pérez, C. (2009). Photogrammetry and augmented reality for cultural heritage applications. *The Photogrammetric Record*, 24(128), 316-331. doi:<https://doi.org/10.1111/j.1477-9730.2009.00549.x>
- Pouloupoulos, V., & Wallace, M. (2022). Digital Technologies and the Role of Data in Cultural Heritage: The Past, the Present, and the Future. (MDPI, Éd.) *Big Data Cogn. Comput*, 6(3). doi:<https://doi.org/10.3390/bdcc6030073>
- Putra, E. Y., Wahyudi, A. K., & Dumingan, C. (2016). A proposed combination of photogrammetry, Augmented Reality and Virtual Reality Headset for heritage visualisation. *International Conference on Informatics and Computing* (pp. 43-48). IEEE. doi:doi: 10.1109/IAC.2016.7905687
- RA, H., & SW, Z. (1983). On the foundations of relaxation labeling processes. *IEEE Trans Pattern Anal Mach Intell* (pp. 67-87). IEEE. doi:10.1109/tpami.1983.4767390. PMID: 21869111
- Rahaman, H., Champion, E., & Bekele, M. (2019). From photo to 3D to mixed reality: A complete workflow for cultural heritage visualisation and experience. *Digital Applications in Archaeology and Cultural Heritage*, 13. doi:<https://doi.org/10.1016/j.daach.2019.e00102>
- Rizvić, S., Bošković, D., Okanović, V., Kihic, I., Prazina, I., & Mijatović, B. (2021). Time Travel to the Past of Bosnia and Herzegovina through Virtual and Augmented Reality. *Applied Sciences* , 11(8). doi:<https://doi.org/10.3390/app11083711>
- Rodríguez-González, P., Jimenez Fernandez-Palacios, B., Muñoz-Nieto, Á., Arias-Sanchez, P., & Gonzalez-Aguilera, D. (2017). Mobile LiDAR system: New possibilities for the documentation and dissemination of large cultural heritage sites. *Remote Sensing*, 9(3). doi:<https://doi.org/10.3390/rs9030189>
- Rome Reborn*. (s.d.). Consulté le 12 10, 2022, sur [www.romereborn.org](http://www.romereborn.org): <https://www.romereborn.org/>
- Roy, S., & Cox, I. J. (1998). A maximum-flow formulation of the N-camera stereo correspondence problem. *Sixth International Conference on Computer Vision* (pp. 492-499). IEEE. doi:10.1109/ICCV.1998.710763
- Samain, E. (s.d.). *Télémétrie Laser*. Caussols.
- Sandip, K., & Bharat, B. (2014). A Survey on Markerless Augmented Reality. *International Journal of Engineering Trends and Technology*, 10(13), 639-641. doi:10.14445/22315381/IJETT-V10P328
- Sato, Y., Fukuda, T., Yabuki, N., Michikawa, T., & Motamedi, A. (2016). A Marker-less Augmented Reality System using Image Processing Techniques for Architecture and Urban Environment. *The 21st International Conference on Computer-Aided Architectural Design Research in Asia*.
- Scharstein, D., & Szeliski, R. (2003). High-accuracy stereo depth maps using structured light. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. I-I). IEEE. doi:10.1109/CVPR.2003.1211354
- Scianna, A., Gaglio, G., Grima, R., & La Guardia, M. (2020). The virtualization of CH for historical reconstruction: The AR fruition of the fountain of St. George square in Valletta (Malta). *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*, 143-149. doi:<https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-143-2020>
- Sebastiani, A. (2021). Digital Artifacts and Landscapes. Experimenting with Placemaking at the Impero Project. *Heritage*, 4(1), 281-303. doi:<https://doi.org/10.3390/heritage4010018>
- Séverinie, D. (2008). Plaquage de texture, Algorithmique graphique et modélisation.

- Shetty, M., Lasrado, V., & Mohammed, R. (2015). Marker Based Augmented Reality Using Android OS. *International Journal of Innovative Research in Computer and Communication Engineering*.
- Shih, N., Diao, P., & Chen, Y. (2019). ARTS, an AR tourism system, for the integration of 3D scanning and smartphone AR in cultural heritage tourism and pedagogy. *Sensors*, 19(17). doi:<https://doi.org/10.3390/s19173725>
- Silva, M., & Teixeira, L. (2020). Developing an eXtended Reality platform for Immersive and Interactive Experiences for Cultural Heritage: Serralves Museum and Coa Archeologic Park. *Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct* (pp. 300–302). IEEE. doi:10.1109/ISMAR-Adjunct51615.2020.00084
- Sites, C., & Elvet, O. (2013). Cultural heritage sites exploring dual-camera-based augmented reality for cultural heritage sites. *Thesis for the Degree of Master of Science by Research*.
- Siti Zuliana, S., & Abd Razak, B. (2022). Bibliometric and content analysis on publications in digitization technology implementation in cultural heritage for recent five years (2016–2021). *Digital Applications in Archaeology and Cultural Heritage*, 25. doi:<https://doi.org/10.1016/j.daach.2022.e00225>
- Slavec, A., Sajincic, N., & Starman, V. (2021). Use of smartphone cameras and other applications while traveling to sustain outdoor cultural heritage. *Sustainability*, 13(13). doi:<https://doi.org/10.3390/su13137312>
- Stéphane, B., & Bruno, T. (2009). Localisation and augmented reality for mobile applications in cultural heritage. *3rd International Workshop 3D ARCH 2009*.
- Tan, K., & Lim, C. (2017). Digital heritage gamification: An augmented-virtual walkthrough to learn and explore historical places. *AIP Conference Proceedings*, 1891. doi:<https://doi.org/10.1063/1.5005472>
- Teneketzis, A. (2020). Exploring the emerging digital scene in Art History and museum practice. *Esboços*, 27(45), 187–206. doi:<https://doi.org/10.5007/2175-7976.2020.e67954>
- The Ultimate Definition of Gamification (with 6 Real World Examples)*. (s.d.). Consulté le 12 11, 2022, sur [www.growthengineering.co.uk](http://www.growthengineering.co.uk): <https://www.growthengineering.co.uk/definition-of-gamification/>
- Tisserand, Y., Magnenat-Thalmann, N., Unzueta, L., Linaza, M., Ahmadi, A., O’connor, N., . . . Daras, P. (2017). Preservation and gamification of traditional sports. In *Mixed Reality and Gamification for Cultural Heritage*. Dans M. M.-T. Loannides. doi:[https://doi.org/10.1007/978-3-319-49607-8\\_17](https://doi.org/10.1007/978-3-319-49607-8_17)
- Tom Dieck, M., & Jung, T. (2017). Value of augmented reality at cultural heritage sites: A stakeholder approach. *Journal of Destination Marketing & Management*, 6(2), 110–117. doi:<https://doi.org/10.1016/j.jdmm.2017.03.002>
- Tom Dieck, M., Jung, T., & Tom Dieck, D. (2018). Enhancing art gallery visitors: learning experience using wearable augmented reality: Generic learning outcomes perspective. *Current Issues in Tourism*, 21(17), 2014–2034. doi:<https://doi.org/10.1080/13683500.2016.1224818>
- Trunfio, M., & Campana, S. (2020). A visitors’ experience model for mixed reality in the museum. *Current Issues in Tourism*, 23(9), 1053–1058. doi:<https://doi.org/10.1080/13683500.2019.1586847>
- Trunfio, M., Campana, S., & Magnelli, A. (2020). Measuring the impact of functional and experiential mixed reality elements on a museum visit. *Current Issues in Tourism*, 23(16), 1990–2008. doi:<https://doi.org/10.1080/13683500.2019.1703914>
- Tzima, S., Styliaras, G., & Bassounas, A. (2019). Augmented reality applications in education: Teachers point of view. *Education Sciences*, 9(2). doi:<https://doi.org/10.3390/educsci9020099>

- Tzima, S., Styliaras, G., & Bassounas, A. (2021). Revealing hidden local cultural heritage through a serious escape game in outdoor settings. *Information*, 12(1). doi:<https://doi.org/10.3390/info12010010>
- UNESCO. (2022, 12 10). *Concept of Digital Heritage*. Consulté le 12 10, 2022, sur [en.unesco.org: https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-heritage](https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-heritage)
- Van Nguyen, S., Le, S., Tran, M., & Tran, H. (2022). Reconstruction of 3D digital heritage objects for VR and AR applications. *Journal of Information and Telecommunication*, 6(3), 254–269. doi:[10.1080/24751839.2021.2008133](https://doi.org/10.1080/24751839.2021.2008133)
- Van Ruymbek, M., Nofal, E., & Billen, R. (2022). 3D Digital Heritage and Historical Storytelling: Outcomes from the Interreg EMR Terra Mosana Project. (Springer, Éd.) *Computer Science*, 13324. doi:[https://doi.org/10.1007/978-3-031-05434-1\\_17](https://doi.org/10.1007/978-3-031-05434-1_17)
- Vert, S., Andone, D., Ternauciuc, A., Mihaescu, V., Rotaru, O., Mocofan, M., . . . Vasiu, R. (2021). User evaluation of a multi-platform digital storytelling concept for cultural heritage. *Mathematics*, 9(21). doi:<https://doi.org/10.3390/math9212678>
- Viinikkala, L., Yli-Seppälä, L., Heimo, O., Helle, S., Härkänen, L., Jokela, S., & Lehtonen, T. (2016). Reforming the representation of the reformation: Mixed reality narratives in communicating tangible and intangible heritage of the protestant reformation in Finland. *Proceedings of the 22nd International Conference on Virtual System & Multimedia*, (pp. 1–9). doi:[10.1109/VSMM.2016.7863203](https://doi.org/10.1109/VSMM.2016.7863203)
- Vlahakis, V., Ioannidis, M., Karigiannis, J., Tstros, M., Gounaris, M., Stricker, D., . . . Almeida, L. (2002). Archeoguide: An augmented reality guide for archaeolog sites. *IEEE Computer Graphics and Applications*, 22(5), 52–60. doi:[10.1109/MCG.2002.1028726](https://doi.org/10.1109/MCG.2002.1028726)
- Vlizos, S., Sharamyeva, J.-A., & Kotsopoulos, k. (2021). Interdisciplinary Design of an Educational Applications Development Platform in a 3D Environment Focused on Cultural Heritage Tourism., (pp. 79–96).
- Wen, Y., & Chen, J. (2016). Intangible cultural heritage display using augmented reality technology of Xtion PRO interaction. *International Journal of Simulation: Systems, Science & Technology*, 17. doi:[10.5013/IJSSST.a.17.40.29](https://doi.org/10.5013/IJSSST.a.17.40.29)
- What is Intangible Cultural Heritage*. (s.d.). Consulté le 12 11, 2022, sur [ich.unesco.org: https://ich.unesco.org/en/what-is-intangible-heritage-00003](https://ich.unesco.org/en/what-is-intangible-heritage-00003)
- Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). real-time tracking of the human body. *Transactions on Pattern Analysis and Machine Intelligence*. 19, pp. 780-785. IEEE. doi:[10.1109/34.598236](https://doi.org/10.1109/34.598236)
- Xie, X., & Tang, X. (2018). The application of augmented reality technology in digital display for intangible cultural heritage: The case of cantonese furniture. *International Conference on Human-Computer Interaction*, 10902, pp. 334–343. doi:[https://doi.org/10.1007/978-3-319-91244-8\\_27](https://doi.org/10.1007/978-3-319-91244-8_27)
- Xu, F., Buhalis, D., & Weber, J. (2017). Serious games and the gamification of tourism. *Tourism Management*, 60, 244–256. doi:<https://doi.org/10.1016/j.tourman.2016.11.020>
- Yang, T., & Zhao, R. (2017). Research on Combination of Intangible Cultural Heritage and Augmented Reality. *Proceedings of the 2nd International Conference on Contemporary Education*, (pp. 536–538).
- Zhang, C., & Chen, T. (2004). A survey on image-based rendering-representation, sampling and compression. *Signal Processing: Image Communication*, 1-28.
- Zhang, R., Tsai, P.-S., Cryer, J., & Shah, M. (1999). Shape-from-shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (pp. 690-706).

- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 12. IEEE. doi:10.1109/34.888718
- Zhao, Z. (2017). Digital protection method of intangible cultural heritage based on augmented reality technology. In *Proceedings of the 2017 International Conference on Robots & Intelligent System*, (pp. 135–138). doi:10.1109/ICRIS.2017.41
- Zhao, Z., & Ma, X. (2020). ShadowPlay2.5D: A 360-degree video authoring tool for immersive appreciation of classical Chinese poetry. *ACM J. Comput. Cult. Herit. Journal on Computing and Cultural Heritage*, 13(1), 1–20. doi:https://doi.org/10.1145/3352590
- Ziagkas, E., Stylianidis, P., Loukovitis, A., Zilidou, V., Lilou, O., Mavropoulou, A., & Douka, S. (2020). Greek traditional dances 3d motion capturing and a proposed method for identification through rhythm pattern analyses (terpsichore project). *Strategic Innovative Marketing and Tourism: Springer Proceedings in Business and Economics*, (pp. 657–665). doi:https://doi.org/10.1007/978-3-030-36126-6\_73
- Zoellner, M., Keil, J., Drevensek, T., & Wuest, H. (2009). Cultural Heritage Layers: Integrating Historic Media in Augmented Reality., (pp. 193-196). doi:10.1109/VSM.2009.35

## **Références Article de** (Djaber & Nour Eddine, 2022)

- Aharchi, M. and Ait Kbir, M. (2020) 'A review on 3D reconstruction techniques from 2D images', in Ben Ahmed, M., Boudhir, A., Santos, D., El Aroussi, M. and Karas, I. (Eds?): *Innovations in Smart Cities Applications Edition 3*. SCA 2019. Lecture Notes in Intelligent Transportation and Infrastructure, Springer, Cham [online] <https://doi.org/10.1007/978-3-030-37629-137>.
- Banks, M.S., Gepshtein, S. and Michael, S. (2004) 'Why is spatial stereoresolution so low', *Landy Journal of Neuroscience*, 3 March, Vol. 24, No. 9, pp.2077–2089 [online] <https://doi.org/10.1523/JNEUROSCI.3852-02.2004>.
- Bebeselea-Sterp, E., Brad, R. and Brad, R. (2017) 'A comparative study of stereovision algorithms', *International Journal of Advanced Computing Sciences and Application*, Vol. 8, pp.1–17 [online] <https://dx.doi.org/10.14569/IJACSA.2017.081144>.
- Brown, M.Z., Burschka, D. and Hager, G.D. (2003) 'Advances in computational stereo', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, August, Vol. 25, No. 8, pp.993–1008 [online] <https://doi.org/10.1109/TPAMI.2003.1217603>.
- Buelthoff, H.H. and Yuille, A.L. (1991) 'Shape-from-X: psychophysics and computation', *Proc. SPIE 1383, Sensor Fusion III: 3D Perception and Recognition*, 1 April [online] <https://doi.org/10.1117/12.25259>.
- Djaber, R. (2012) 'Combining the stereo-vision and visual hull methods for increasing the 3D reconstruction speed', *2012 International Conference on Information Technology and e-Services*, pp.1–6 [online] <https://doi.org/10.1109/ICITeS.2012.6216627>.
- Durou, J.D., Falcone, M., Quéau Y. and Tozza, S. (2020) 'A comprehensive introduction to photometric 3d-reconstruction', in Durou, J.D., Falcone, M., Quéau, Y. and Tozza, S. (Eds.): *Advances in Photometric 3D-Reconstruction. Advances in Computer Vision and Pattern Recognition*, Springer, Cham [online] <https://doi.org/10.1007/978-3-030-51866-01>.
- Gwet, D.L.L., Ottesteanu, M., Libouga, I.O., Bitjoka, L. and Popa, G.D. (2018) 'A review on image segmentation techniques and performance measures', *International Journal of Information, Control and Computer Sciences*, Vol. 12, No. 12 [online] <https://doi.org/10.5281/zenodo.2571650>.
- Ham, H. et al. (2019) 'Computer vision based 3D reconstruction: a review', *International Journal of Electrical and Computer Engineering (IJECE)* [online] <http://doi.org/10.11591/ijece.v9i4.pp2394-2402>.
- Hamzah, R.A. and Ibrahim, H. (2016) 'Literature survey on stereo vision disparity map algorithms', *Journal of Sensors* [online] <https://doi.org/10.1155/2016/8742920>.



- Han, X-F., Laga, H. and Bennamoun, M. (2021) 'Image-based 3D Object reconstruction: state-of-the-art and trends in the deep learning era', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1 May, Vol. 43, No. 5, pp.1578–1604, doi: 10.1109/TPAMI.2019. 2954885.
- Hartley, R. and Zisserman, A. (2004) *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge [online] <https://doi.org/10.1017/CBO9780511811685>.
- Hirschmuller, H. (2005) 'Accurate and efficient stereo processing by semi-global matching and mutual information', 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2, pp.807–814, doi: 10.1109/CVPR.2005.56.
- Hou, G., Zhang, W., Wu, B. and He, R. (2022) '3D reconstruction and positioning of surface features based on a monocular camera and geometric constraints', *Appl. Opt.*, Vol. 61, pp.C27–C36 [online] <https://doi.org/10.1364/AO.436234>.
- Kordelas, G., Agapito, J.D.P-M., Vegas, J.M. and Daras, H.P. (2010) 'State-of-the-art algorithms for complete 3D model reconstruction', Summer School ENGAGE – Immersive and Engaging Interaction with VH on Internet, September.
- Lane, R. and Thacker, N. (1996) *Stereo Vision Research: An Algorithm Survey*.
- Lin, H. and Wu, J. (2008) '3D reconstruction by combining shape from silhouette with stereo', 2008 19th International Conference on Pattern Recognition, pp.1–4 [online] <https://doi.org/10.1109/ICPR.2008.4761016>.
- Lin, H-Y., Chang, C-C. and Wu, J-R. (2014) 'A novel 3D reconstruction approach', *International Journal of Innovative Computing, Information Control, IJICIC*, IJICIC-13-12018.dvi.
- Morales, A.M.E. and Lopez, J.C.M. (2016) 'Embedded real-time stereo estimation via semi-global matching on the GPU', *Procedia Computer Science*, Vol. 80, pp.143–153 [online] <https://doi.org/10.1016/j.procs.2016.05.305>.
- Perkins, D.N. (1970) *Computer Stereo Vision: A Combinatorial Theory with Implementation*, PhD thesis, Massachusetts Institute of Technology, Department of Mathematics.
- Scharstein, D. and Szeliski, R. (2002) 'A taxonomy and evaluation of dense two-frame stereo correspondence algorithms', *International Journal of Computer Vision*, Vol. 47, No. 1, pp.7–42 [online] <https://doi.org/10.1023/A:1014573219977>.
- Seitz, S.M., Curless, B., Diebel, J., Scharstein, D. and Szeliski, R. (2006) 'A comparison and evaluation of multi-view stereo reconstruction algorithms', 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), pp.519–528 [online] <https://doi.org/10.1109/CVPR.2006.19>.
- Siudak, M. and Rokita, P. (2014) 'A survey of passive 3D reconstruction methods on the basis of more than one image', *Machine Graphics Vision*, pp.57–117, MG233 4057 117.pdf.
- Sun, X., Mei, X., Jiao, S., Zhou, M. and Wang, H. (2011) 'Stereo matching with reliable disparity propagation', 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, pp.132–139 [online] <https://doi.org/10.1109/3DIMPVT. 2011.24>.
- Zollhöfer, M., Stotko, P., Görlitz, A., Theobalt, C., Nießner, M., Klein, R. and Kolb, A. (2018) 'State of the art on 3D reconstruction with RGB-D cameras', *Computer Graphics Forum*, Vol. 37, pp.625–652 [online] <https://doi.org/10.1111/cgf.13386>.