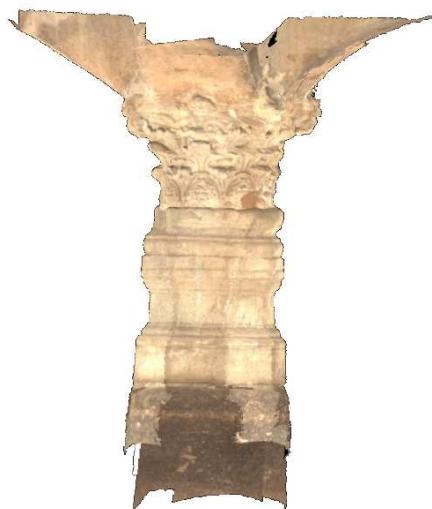




**Rapport de Projet Pluridisciplinaire
Mastère spécialisé Photogrammétrie, positionnement et mesures
de déformations (PPMD)**

**FUSION DE CARTES DE PROFONDEURS EN
PHOTOGRAMMETRIE ARCHITECTURALE**



GARDON Elzbieta

le 30 septembre 2009

Non confidentiel Confidential IGN Confidential Industrie jusqu'au

Jury

Président du jury

Michel Kasser, directeur de l' ENSG

Commanditaire :

Marc Pierrot Deseilligny,
MATIS/IGN
2, 4 Avenue Pasteur - 94165 Saint-Mandé

Encadrement de stage :

Marc Pierrot Deseilligny, Direction technique/IGN, maître de stage
Olivier Dissard, EL, ENSG/IGN, rapporteur principal

Responsable pédagogique de la formation PPMD (option photogrammétrie) :

Raphael Heno, Département Imagerie Aérienne et Spatiale

© ENSG

Stage Pluridisciplinaire du 27 avril au 18 septembre

Diffusion Web : Internet Intranet Polytechnicum Intranet ENSG

Situation du document :

rapport de stage pluridisciplinaire présenté en fin de la formation PPMD

Nombre de pages : 67 dont 7 d'annexes

Système hôte : Word 2003

MODIFICATIONS

Remerciements

Je voudrais d'abord exprimer mes remerciements à Marc Pierrot Deseilligny, mon maître de stage, pour m'avoir encadrée durant le stage. Il a su me guider tout au long de ces 4 mois et me donner des conseils pertinents.

Un grand merci à l'équipe du laboratoire LOEMI pour son accueil dans ses locaux et plus particulièrement à mon collègue de bureau Pierre.

Je remercie également les stagiaires rencontrés à l'occasion de ce stage Christophe et Sébastian, pour les discussions variées, les échanges à propos de nos stages.

Résumé

En photogrammétrie architecturale, la scène est représentée couramment comme un ensemble de cartes de profondeurs. Lorsque ces cartes se recouvrent largement, il est nécessaire d'effectuer un travail de fusion.

Dans le cadre des travaux de recherche du laboratoire MATIS un nouvel algorithme de fusion de cartes de profondeur a été élaboré. Le but de la méthode est de conserver les parties les plus fiables de chaque carte. Pour cela, plusieurs critères permettant de mesurer la fiabilité ont été définis : le critère de l'angle d'incidence, le critère de corrélation et enfin le critère d'étirement de triangle.

Dans ce contexte, l'étape préalable importante a été de déterminer la visibilité entre les différents nuages de points 3D. Pour ce faire, la solution retenue consiste à traiter à un moment donné tous les nuages dans la même géométrie avec la mise en place d'un Z-Buffer. Plusieurs exemples montrant l'efficacité et les limitations de la méthode sont présentés.

Mots-clés

fusion, carte de profondeur, reconstruction 3D, photogrammétrie architecturale

Abstract

In architectural photogrammetry, the scene is usually represented as a set of depth maps. When these maps are largely recovered, it is necessary to make a work of fusion.

Within the framework of the research works of the laboratory MATIS a new algorithm of fusion of depth maps was elaborated. The purpose of the method is to preserve the most reliable parts of each map. For that, several criteria allowing to measure the reliability were defined: the criterion of the angle of incidence, the criterion of correlation and finally the criterion of stretching of triangle.

In this context, the important preliminary stage was to determine the visibility between the different clouds of 3D points. The adopted solution consists in processing, at one moment, all the clouds in the same geometry with the implementation of Z-Buffer. Several examples showing the efficiency and the limitations of the method are presented.

Keywords

fusion, depth carte, reconstruction 3D, photogrammetry architectural

Table des matières

INTRODUCTION.....	11
1. CONTEXTE DE L'ETUDE	13
1.1. PRESENTATION DU MATIS	13
1.2. CONTEXTE DU STAGE	13
1.3. OBJECTIF DU STAGE	14
2. METHODES DE RECONSTRUCTION PHOTOGRAMMETRIQUE D'UNE SCENE ARCHITECTURALE	16
2.1. LE FLUX DE TRAITEMENTS DU MATIS	16
2.1.1. <i>Calibrage</i>	16
2.1.2. <i>Orientation et localisation</i>	18
2.1.3. <i>Mise en correspondance</i>	19
2.1.4. <i>Caractéristique des cartes de profondeur</i>	22
2.2. AUTRES APPROCHES.....	23
2.2.1. <i>Reconstruction à partir de silhouettes</i>	23
2.2.2. <i>Reconstruction à partir de la texture</i>	24
3. METHODES DE FUSION EXISTANTES	26
3.1. METHODES VOLUMIQUES.	26
3.2. METHODES SURFACIQUES.	26
4. METHODE DE FUSION PROPOSEE	27
4.1. DONNEES EN ENTREE.....	27
4.2. CRITERES	27
4.2.1. <i>Critère de visibilité</i>	28
4.2.2. <i>Critère d'angle d'incidence</i>	29
4.2.3. <i>Critère de corrélation</i>	32
4.2.4. <i>Critère d'étirement de triangle</i>	33
4.3. APPROCHE AVEC BASCULEMENT	34
4.3.1. <i>Stratégie générale</i>	34
4.3.2. <i>Préfiltrage</i>	37
4.3.3. <i>Filtrage morphologique</i>	37
4.3.4. <i>Reprojection d'un masque déjà calculé</i>	38
4.3.5. <i>Schéma de l'algorithme détaillé</i>	40
4.4. APPROCHE « DIRECTE ».....	42

4.4.1.	<i>Stratégie générale</i>	42
4.4.2.	<i>Critère de visibilité sans basculement</i>	42
4.5.	IMPLEMENTATION	45
4.5.1.	<i>Outils de développement</i>	45
4.5.2.	<i>Interface graphique et ses fonctionnalités</i>	47
4.5.3.	<i>Documentation</i>	49
5.	TESTS ET LEURS RESULTATS	51
5.1.	PILIER (ALEXANDRIE)	51
5.2.	NUBIENNE (LOUVRE)	52
5.3.	STALAGMITE (GROTTES DE CHORANCE)	54
5.4.	EGLISE DE ST- LAURENT (LATTES).....	55
5.5.	AIGUILLE DU MIDI (MASSIF DU MONT-BLANC)	57
5.6.	BILAN.....	58
	CONCLUSION	59
	BIBLIOGRAPHIE.....	60

Liste des figures

<i>Figure 1 : Images et cartes de profondeurs (en ombrage) sur trois points de vue d'un pilier de la citerne El-Nabih à Alexandrie.</i>	14
<i>Figure 2 : Flux de traitements (MATIS) de reconstruction 3D d'une scène architecturale.....</i>	16
<i>Figure 3 : Polygone et cible de calibrage.....</i>	17
<i>Figure 4 : Transformation rigide entre les coordonnées du monde et les coordonnées de la caméra.</i>	18
<i>Figure 5 : La mise en correspondance dans l'espace d'objet.....</i>	19
<i>Figure 6 : Les images du haut : deux vues d'un portrait dans la géométrie épipolaire. Les cercles désignent des points correspondants. L'image du bas: la carte de disparité.....</i>	20
<i>Figure 7 : Calcul de la profondeur. Cas de deux images dans la géométrie épipolaire.....</i>	21
<i>Figure 8 : Ligne de parcours de l'espace de l'image maîtresse à (c, l) fixée. Reprojection d'un point terrain dans les autres images.</i>	22
<i>Figure 9 : Ecarts entre deux cartes. En verts : les zones où l'écart entre deux surfaces est inférieur à 1 pixel (image gauche) et à 10 pixels (image droite).</i>	23
<i>Figure 10 : Enveloppe visuelle.</i>	24
<i>Figure 11 : « Space carving ».....</i>	24
<i>Figure 12 : Structure de données avant la fusion</i>	27
<i>Figure 13 : A gauche : Deux nuages 3D. A droite : résultat de la mise en place du Z-buffer (nuage B basculé dans la géométrie du nuage A).....</i>	28
<i>Figure 14: De gauche à droite: 1) nuage de points A ; 2) nuage de points B ; 3) basculement du nuage B dans la géométrie du nuage A ; 4) écart de profondeur entre les deux nuages.....</i>	29
<i>Figure 15 : Représentations de la même surface. A gauche : angle d'incidence plus fiable. A droite : valeur d'angle d'incidence plus important.....</i>	30
<i>Figure 16 : Cartes d'angle d'incidence. De gauche : sans lissage. Au centre : lissage de 5 pixels. De droite : lissage de 15 pixels.</i>	31
<i>Figure 17 : Critère d'angle d'incidence.</i>	31
<i>Figure 18 : L'image de gauche : carte d'angle. L'image de droite : « masque de définition » retenu.</i>	32
<i>Figure 19 : Exemple de la carte de corrélation.</i>	33
<i>Figure 20 : De gauche à droite: 1) nuage « maitresse » ; 2) nuage à basculer ; 3) carte d'étirement de triangle</i>	34
<i>Figure 21 : L'enchaînement de l'approche avec basculement.....</i>	35
<i>Figure 22 : Segmentation d'un “masque de définition”.....</i>	37
<i>Figure 23 : A gauche : « masque de définition » sans filtrage. Au centre : érosion géodésique appliquée à l'image (les artefacts dans le carré bleu ont disparus). A droite : érosion et dilatation appliqués à l'image (les « trous » dans le carré vert ont été bouchés).</i>	38

<i>Figure 24 : L'algorithme de reprojection.</i>	39
<i>Figure 25 : Critère de visibilité.</i>	43
<i>Figure 26 : Critère de visibilité. Deux points n'appartiennent pas à la « même » surface.</i>	44
<i>Figure 27 : Défaut du critère de visibilité.</i>	45
<i>Figure 28 : Interface “Viewer of 3D point clouds”.</i>	47
<i>Figure 29 : Fenêtre de l'export d'un modèle 3D.</i>	48
<i>Figure 30 : Fenêtre des nuages</i>	49
<i>Figure 31 : Documentation du code.</i>	50
<i>Figure 32 : Pilier d'Alexandrie. En haut : modèle avant et après la fusion. En bas : modèle après la fusion texturé</i>	51
<i>Figure 33 : Pilier. Défaut du résultat. Parties manquantes.</i>	52
<i>Figure 34 : Nubienne du Louvre. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturée</i>	53
<i>Figure 35 : Stalagmite en Alpes. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturée</i>	54
<i>Figure 36 : Stalagmite. Défaut du résultat « trous ».</i>	55
<i>Figure 37 : Eglise de St-Laurent de Lattes. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturé</i>	56
<i>Figure 38 : Aiguille de Midi en Alpes. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturé</i>	57
<i>Figure 39 : Tableau contenant les valeurs de redondance avant est après la fusion.</i>	58
<i>Figure 40 : Redondance aux raccords entre les cartes fusionnées (pixels rouge).</i>	58
<i>Figure 41 : Distance géodésique.</i>	67

Liste des annexes

<i>Annexe 1 « Formule d'image »</i>	61
<i>Annexe 2 Relèvement dans l'espace.....</i>	62
<i>Annexe 3 Orientation relative entre deux images</i>	64
<i>Annexe 4 Géométrie épipolaire.....</i>	65
<i>Annexe 5 Erosion et dilatation géodésique.....</i>	67

INTRODUCTION

La reconstruction 3D à partir d'images a été largement étudiée dans le domaine de la vision par ordinateur, d'où les nombreux algorithmes proposés. Leurs différences sont principalement liées au modèle utilisé pour représenter la scène. Les méthodes utilisant des modèles volumiques comme les voxels ou les ensembles de niveau, sont basées sur une discréétisation de l'espace en cellules. Leur but est de déterminer la frontière entre les cellules pleines et les cellules vides. Néanmoins, la discréétisation initiale limite la résolution finale des modèles. D'un autre côté, les méthodes utilisant des maillages peuvent, en théorie, adapter leur résolution pour mieux reconstruire les détails de la scène. Par contre, elles souffrent du problème lié à la difficulté de gestion des changements topologiques pendant le traitement.

Les développements récents en lasergrammétrie font de ce nouvel instrument un outil robuste de la reconstruction 3D. Cette technique permet d'acquérir les données avec une grande fiabilité. En revanche, l'image reste toujours nécessaire pour texturer et comprendre le modèle relevé. Cependant, le laser est une technique d'acquisition relativement lente, alors que l'acquisition photographique est quasi instantanée. De nombreux projets combinent les techniques de photogrammétrie et de lasergrammétrie, ce qui constitue une voie très prometteuse de la modélisation 3D du patrimoine.

Les outils actuels de photogrammétrie permettent de reconstruire des modèles 3D de scènes sous forme de cartes de profondeur. L'utilisation ces cartes élimine les problèmes de discréétisation et des changements topologiques.

Cependant, dans le cas d'un relief non plan, il est fréquent qu'une seule carte ne suffise pas pour représenter la totalité de la scène. Pour modéliser un objet de type sculpture, rencontré couramment en photogrammétrie architecturale, il faut une carte de profondeur pour ces principaux points de vue. Lorsque ces cartes se recouvrent largement, il est nécessaire d'effectuer un travail de fusion. Pour résoudre ce problème, plusieurs approches existent dans la littérature. D'une part, les méthodes volumiques (*Curless et Levoy [1], Hilton[2]*) et d'autre part les méthodes surfaciques (*Sousy et Larendeau [3], Turk et Levoy [4], Pito [5]*). Les méthodes volumiques sont basées sur le principe de construction d'une surface implicite dans l'espace 3D. Par contre, dans les méthodes surfaciques les images de profondeurs sont triangulées et superposées, puis localement fusionnées pour éliminer les données redondantes.

Le problème de fusion a été également étudié dans le cas des nuages de points 3D issus de scanner laser. Par contre, les méthodes proposées peuvent difficilement être adaptées en photogrammétrie car elles supposent l'absence d'erreurs liées au processus.

Ce document présente un travail de recherche sur l'algorithme de fusion de cartes de profondeur. Le but des traitements est de conserver l'information la plus fiable. Les mesures de fiabilité se font par vérification de plusieurs critères: d'angle d'incidence, de corrélation, d'étirement de triangle.

Le premier chapitre décrit le contexte et les objectifs du stage. Le deuxième chapitre présente les techniques de réalisation d'une carte de profondeur élaboré au laboratoire MATIS. Ensuite, la troisième partie concerne l'étude de l'existant en terme de fusion de carte de profondeur. Le quatrième chapitre contient la description détaillée de l'algorithme proposé. Dans la dernière partie du document, plusieurs exemples illustrent l'efficacité et les limitations de la méthode.

1. CONTEXTE DE L'ETUDE

1.1. Présentation du MATIS

Ce stage de fin de mastère à été réalisé au laboratoire des Méthodes d'Analyses pour le Traitement d'Images et la Stéréorestitution (MATIS) à l'Institut National Géographique (IGN) dans le cadre de l'axe de recherche PAMPA.

La mission du laboratoire est de mener des recherches contribuant à l'amélioration de la production des bases de données, actuelles et futures à partir des données images "brutes". Ses compétences portent sur le traitement d'images et la vision par ordinateur, la physique de l'image et les problèmes inverses, la reconnaissance de formes, l'algorithmie discrète, les mathématiques appliquées, la théorie des graphes... ainsi que l'informatique et la programmation.

Le laboratoire est structuré en 4 actions de recherche.

- ARCHI : Photogrammétrie Terrestre et Architecturale
- BRIDGE : Reconstruction du bâti pour la cartographie à grande échelle
- LASER : Laser aéroporté
- PAMPA : Photogrammétrie Architecturale

1.2. Contexte du stage

Les outils actuels de photogrammétrie (compensation de faisceaux, vision par ordinateur, corrélation automatique) permettent de construire automatiquement des modèles 3D de scènes sous forme de cartes de profondeur en général relativement fiables.

Dans ces cartes de profondeur, le relief est représenté sous forme d'une matrice $Z = f(X, Y)$. Lorsque la scène est globalement plane (bas-relief, terre vue du ciel ...), une seule de ces représentations permet une modélisation géométrique correcte. Par contre, avec les objets "réellement" 3D (sculptures, piliers), rencontrés couramment en photogrammétrie architecturale, il est nécessaire d'avoir plusieurs points de vue pour obtenir une description satisfaisante de l'ensemble de la scène.

La figure 1 représente trois points de vue (images et cartes calculées) sur un pilier de la citerne El-Nabih d'Alexandrie.

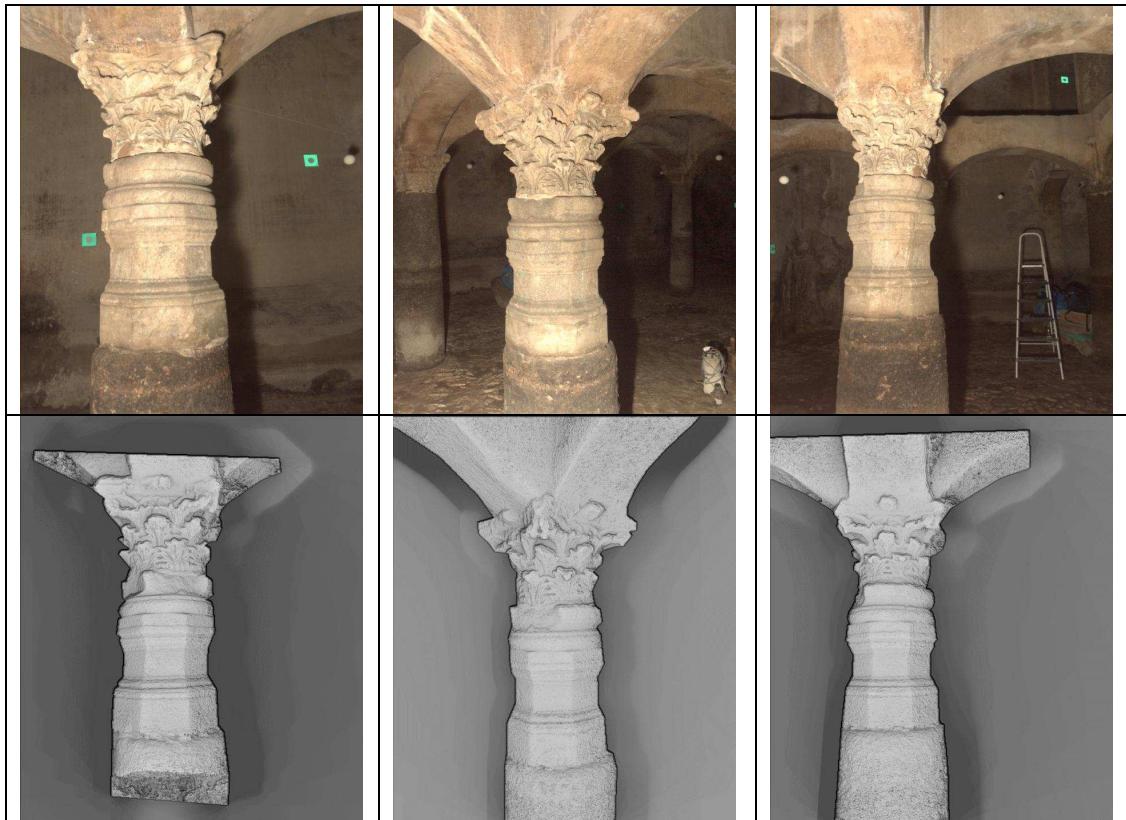


Figure 1 : Images et cartes de profondeurs (en ombrage) sur trois points de vue d'un pilier de la citerne El-Nabih à Alexandrie.

Les approches de reconstruction par les cartes de profondeur sont les plus fréquentes mais ce ne sont pas les seules. Il existe notamment des méthodes d'optimisation directe de surface 3D (« Level Sets ») qui ne nécessitent pas d'étape de fusion. Elles sont très efficaces en terme de qualité, mais très lourdes au niveau de temps de calcul et encombrement mémoire.

1.3. Objectif du stage

Un recouvrement large entre les cartes a été conservé en cours de traitements pour assurer l'existence d'information sur chaque point de l'objet avec une incidence raisonnable. Le fait d'avoir une redondance de données nécessite de les fusionner indépendamment depuis chaque point de vue.

Deux étapes de fusion peuvent être distinguées:

- Une première étape qui consiste à supprimer les redondances en ne conservant qu'une partie de l'information pour chaque carte. Cela revient au calcul d'un masque, appelé ensuite « *masque de définition* ».
- Une deuxième étape qui consiste à "gommer" les discontinuités qui subsistent aux raccords entre les cartes

Le stage porte prioritairement sur la première étape. Son objectif est d'imaginer et d'implémenter des algorithmes de traitement d'images permettant de calculer automatiquement pour chaque point de vue le « *masque de définition* ». Les critères à prendre en compte pour calculer un « bon » masque sont les suivants:

- les masques doivent assurer une partition de la scène 3D (ni recouvrement ni omission)
- lorsqu'il y a redondance, les masques doivent permettre de choisir l'information la plus fiable (en utilisant des différents critères)
- les masques doivent être assez réguliers

L'objectif du stage a été enrichi par l'élaboration d'une interface de visualisation des nuages 3D générés directement à partir des cartes de profondeurs.

Pour effectuer ce travail, les cartes de disparité et une bibliothèque permettant de faire tous les passages 2D vers 3D ont été fournis.

2. METHODES DE RECONSTRUCTION PHOTOGRAMMETRIQUE D'UNE SCENE ARCHITECTURALE

2.1. Le flux de traitements du MATIS.

Dans l'approche de reconstruction 3D d'une scène architecturale élaborée par le laboratoire, on peut distinguer 3 étapes principales : le calibrage, l'orientation et la localisation, enfin la mise en correspondance. La description détaillée de chaque étape est présentée dans les paragraphes suivants. Le résultat du traitement est l'ensemble de cartes de profondeur. Le recouvrement entre ces cartes est assez important donc il est nécessaire par la suite de rajouter une autre étape permettant de les fusionner. Le problème d'unification du modèle 3D est le sujet de ce stage.

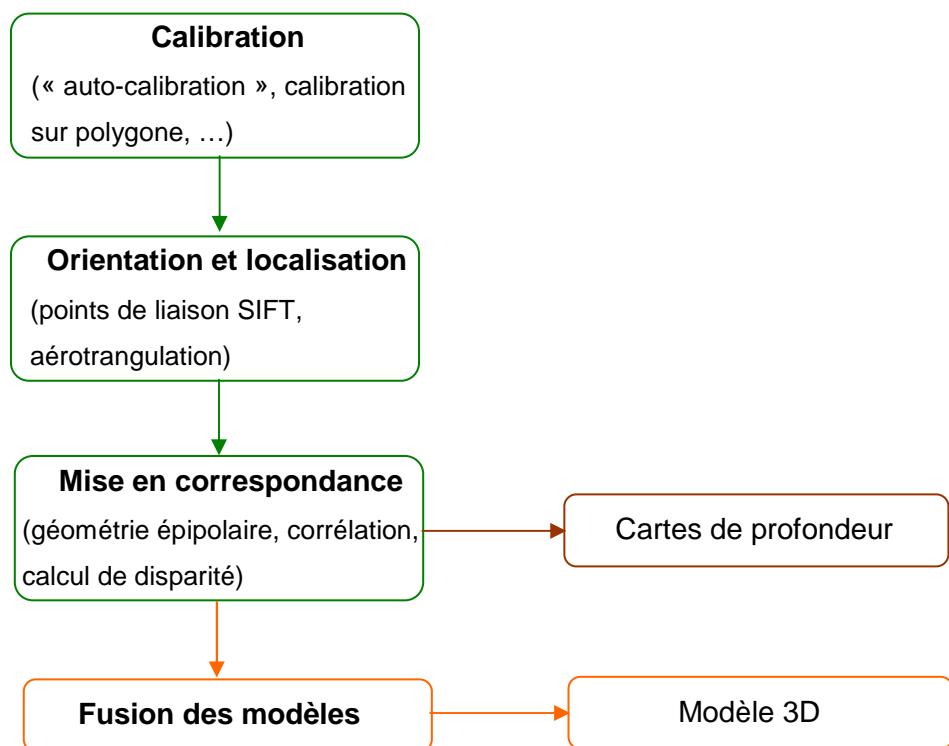


Figure 2 : Flux de traitements (MATIS) de reconstruction 3D d'une scène architecturale

2.1.1. Calibrage

Le processus de *calibrage* (*calibration*, *calibration intrinsèque*, *orientation interne*) consiste à déterminer les paramètres de caméra (par exemple : focale, point principal d'autocollimation, distorsion) à partir de points 3D connus ou d'objet de calibration.

Il existe plusieurs méthodes de calibrage :

- en laboratoire sur banc optique
- par « auto-calibration » d'une scène inconnue prise de différents points de vue
- par utilisation d'une scène à la géométrie maîtrisée (par exemple calibrage sur polygone)

Les caméras du laboratoire sont calibrées sur le polygone de calibrage qui se trouve auprès des locaux de l'IGN. Le principe de cette méthode est présenté ci-dessous.



Figure 3 : Polygone et cible de calibrage.

Le polygone est photographié depuis plusieurs points de vue. Chaque cible est connue en trois dimensions. Les coordonnées des cibles dans les images sont saisies par un opérateur où détectées automatiquement.

On note $(x_l \ y_l \ z_l)^T$ la position terrain de la $l^{\text{ième}}$ cible.

On note $(i_{l,k} \ j_{l,k})$ la projection dans la $k^{\text{ième}}$ image de la $l^{\text{ième}}$ cible.

Soit $k^{\text{ième}}$ la pose de camera définie par les paramètres O_k (coordonnées du sommet de prise de vue) et R_k (matrice de rotation de la camera).

Pour calculer les paramètres intrinsèques \mathfrak{I} (focale, point principal, distorsion), il faut résoudre le système d'équations donné par la formule d'image (annexe 1) pour tous les (l, k) tels que la cible l est détectée dans l'image k .

$$(i \ j) = \mathfrak{I} \left(\pi \left(R_k * \begin{pmatrix} x \\ y \\ z \end{pmatrix} - O_k \right) \right)$$

Les coordonnées terrain et coordonnées images des cibles sont les observations. Les paramètres intrinsèques et la pose de camera sont les inconnues.

Schéma général de calculs :

- calcul de valeurs initiales obtenues par les démarches décrit au-dessous
- calcul des valeurs initiales des paramètres de pose par relèvement dans l'espace (méthode de calcul inclus dans l'Annexe 2)
- compensation par les moindres carrés

2.1.2. Orientation et localisation

Cette étape consiste à trouver la relation entre le référentiel environnement et le référentiel caméra, c'est-à dire, position et orientation de la caméra par rapport à la scène. Les termes équivalents sont : la calibration extrinsèque, l'aéro-triangulation, l'orientation externe.

La calibration extrinsèque est une transformation rigide 3D. La translation T sert à faire correspondre les origines des systèmes de coordonnées. Par contre, la rotation R à faire correspondre les axes des systèmes.

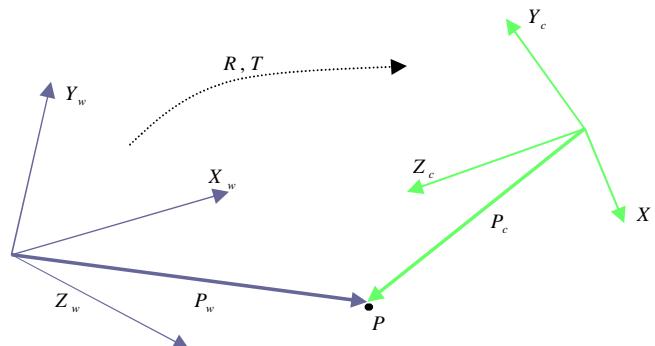


Figure 4 : Transformation rigide entre les coordonnées du monde et les coordonnées de la caméra.

Le but de l' «**orientation relative** » est de déterminer les orientations relatives des deux faisceaux perspectifs, telle que tous les rayons homologues se coupent. Les seules informations viennent de points homologues entre les images. Elle nécessite au minimum 5 points de liaison par couple d'images. Les détails des calculs d'orientation relative entre deux images sont inclus dans l'annexe 3.

Pour trouver des points de liaison, les différents détecteurs de points d'intérêt sont employés. De nombreuses méthodes existantes, les plus connues sont : le SIFT (*Scale Invariant Feature Transform*), le SUSAN, le détecteur de Haris, le détecteur de coins de Moravec, etc.

La « **géométrie épipolaire** » décrit également les contraintes existantes entre les deux projections d'un point dans l'espace des deux images. En connaissant la matrice essentielle, il est possible de calculer directement l'orientation relative d'un couple d'images.

2.1.3. Mise en correspondance

A la base de l'appariement dense de deux images se trouve la recherche d'homologues entre les images. Etant donné un pixel dans une image, on appelle homologue le pixel de la deuxième image qui correspond au même point de la scène.

La mise en correspondance peut se faire dans « *l'espace d'objet* » ou dans « *l'espace d'image* ».

Dans le premier cas, l'espace euclidien est discrétement à une résolution liée aux images. Pour trouver des points homologues, on recherche le meilleur Z (score de corrélation maximal) sur la ligne verticale.

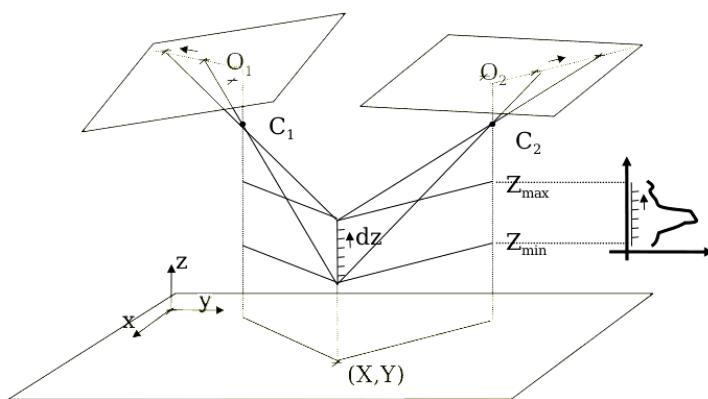


Figure 5 : La mise en correspondance dans l'espace d'objet.

Dans la géométrie épipolaire (« espace image »), pour tout point dans une image, le point correspondant se trouve sur la même ligne dans l'autre image. Les correspondances peuvent servir comme les points de liaison pour une nouvelle estimation de pose.

La « **disparité** » d est l'écart entre deux points correspondants. Elle va permettre de calculer la profondeur de chaque pixel.

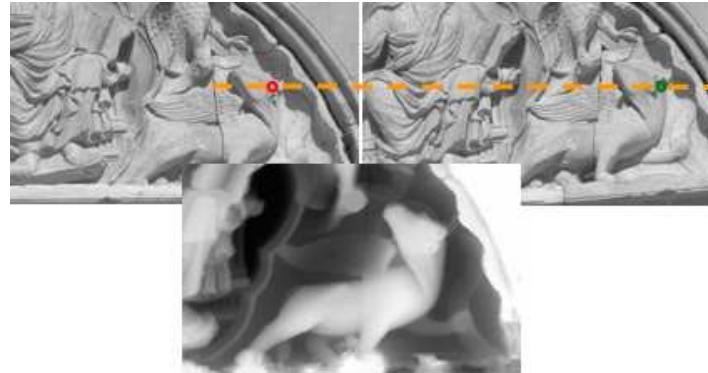


Figure 6 : Les images du haut : deux vues d'un portrait dans la géométrie épipolaire. Les cercles désignent des points correspondants. L'image du bas: la carte de disparité.

Les méthodes d'appariement dense sont basées sur la **mesure de similarité**.

Pour un pixel (x_1, y_1) de la première image, on estime le degré de la « ressemblance » à un pixel (x_2, y_2) de la deuxième image. Pour cela, on compare les valeurs des pixels autour de (x_1, y_1) et (x_2, y_2) .

Il existe de nombreuses fonctions de ressemblance de voisinage. La mesure la plus courante est le coefficient de corrélation normalisé et centré. Ce coefficient est invariant par translation, par homothétie des radiométries et est rapide à calculer.

$$C(I_1 I_2) = \frac{E(I_1 I_2) - E(I_1)E(I_2)}{\sqrt{(E(I_1^2) - E^2(I_1))(E(I_2^2) - E^2(I_2))}} \quad E(I) = \frac{\sum_{x=-N}^{x=+N} \sum_{y=-N}^{y=+N} I(X+x, Y+y)}{(2 * N + 1)^2}$$

- Approches énergétiques

Ces méthodes consistent à trouver une parallaxe $\Gamma(x, y)$ qui minimalise une énergie $E(\Gamma)$.

$$E(\Gamma) = \underbrace{\iint A(I_K(\pi_k(x, y, \Gamma(x, y))))}_{\text{attache aux données}} + \underbrace{\alpha F(\partial\Gamma)}_{\text{régularisation}}$$

Le terme d'attache aux données est le coefficient de corrélation normalisé. La régularisation de la surface est un terme à priori.

Cette approche est une méthode qui permet de combiner les observations (quantifiées par les mesures de similarités) et les connaissances a priori sur la scène (régularité). La difficulté est de trouver une méthode de calcul du minimum. Parfois, on ne peut pas calculer sa valeur exacte et on se contente d'un « pseudo optimum ».

« Programmation dynamique»

C'est une méthode de calcul approchée du minimum, mais rapide et très générale. Elle est équivalente à la recherche d'un chemin optimal dans un graphe, par la minimisation d'une fonction de coût. D'abord, l'espace est balayé selon une direction quelconque. Chaque ligne est traitée indépendamment. Pour chaque hauteur de la $k^{ième}$ colonne, le plus court chemin qui y mène depuis le début de la ligne est calculé. Lorsque les lignes sont traitées indépendamment, les artefacts suivant les lignes apparaissent. Pour lisser ces défauts, on combine les résultats suivant de multiples directions.

« Flot de graphe »

C'est une méthode de calcul exacte du minimum, mais elle ne fonctionne qu'avec certaines énergies et est assez consommatrice de ressources (CPU et mémoire). Le but de cette approche est de trouver la coupure d'un graphe qui possède un coût minimum parmi toutes les coupures possibles, en minimisant une fonction d'énergie. Les terminaux du graphe représentent les disparités pouvant être associées à un pixel de l'image. Pour définir la fonction d'énergie, on introduit une fonction de coût basée sur un critère de photoconsistance (mesure de similarité) appelé le terme de données (data term). Un second terme, appelé terme d'homogénéité ou d'uniformité spatiale (smoothness term)

Tant que la disparité est calculée, la profondeur peut être estimée.

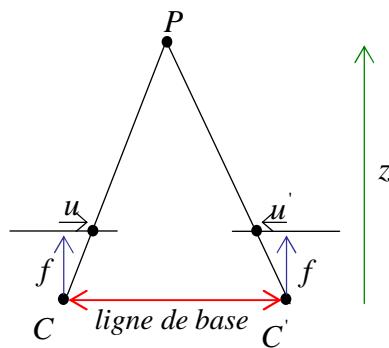


Figure 7 : Calcul de la profondeur. Cas de deux images dans la géométrie épipolaire.

$$disparité = u - u' = \frac{\text{ligne de base} * f}{Z} \Rightarrow Z = \frac{\text{ligne de base} * f}{disparité}$$

Si la ligne de base est trop petite une grande erreur sur la profondeur se reproduira. Par contre, si la ligne de base est grande, la mise en correspondance sera plus difficile.

2.1.4. Caractéristique des cartes de profondeur

Les cartes de profondeur fournies ont été obtenues à partir d'images multiples. Le passage de deux images à un nombre supérieur peut permettre d'améliorer les résultats. En effet, à deux images, la contrainte épipolaire est à une dimension. De nombreux auteurs proposent de résoudre le problème de reconstruction à partir d'images multiples par le calcul des cartes de profondeurs. Dans la chaîne de traitement élaboré au MATIS, les calculs se déroulent dans la géométrie d'image. D'abord, une image maîtresse est choisie. Pour chacun de ses pixels (colonne, ligne) la profondeur à laquelle se situe l'image réelle est cherchée. Un point M de l'espace en trois dimensions est repéré par ces coordonnées (c, l, p) liées à l'image maîtresse : colonne, ligne profondeur. Le principe est le suivant : à c et l fixés, le meilleur » profondeur p qui maximise le score de corrélation est recherché entre deux bornes. Le coefficient de corrélation mesure la similarité de chaque vignette avec la vignette de l'image maîtresse.

La figure 8 montre la stratégie de parcours de l'espace en géométrie image.

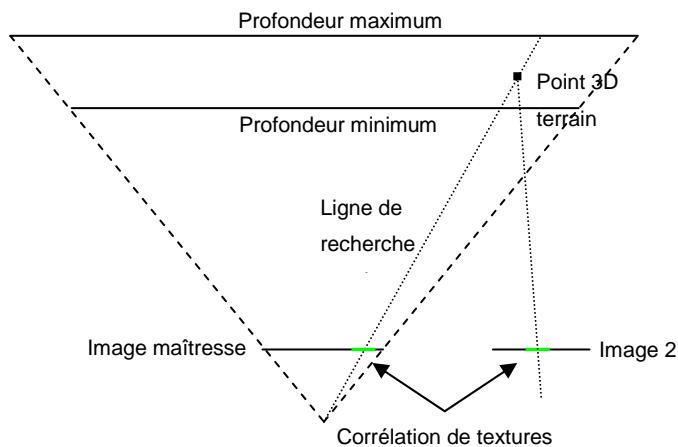


Figure 8 : Ligne de parcours de l'espace de l'image maîtresse à (c, l) fixée. Reprojection d'un point terrain dans les autres images.

Les pas de discréétisation en colonne et en ligne sont de un pixel, suivant la grille image. Un pas constant de profondeur n'est pas adapté car il mène à une description trop fine aux grandes profondeurs. Elle est quantifiée en fonction du z (un $\text{pas}1/z$) où z est la distance entre l'objet et le centre de prise de vue.

La résolution d'une carte de profondeur vaut environ un pixel en profondeur. Pour vérifier que l'erreur est de l'ordre de grandeur du pixel, on compare deux cartes issues d'images différentes (en absence de référence absolue). La figure 9 illustre le résultat de la comparaison.

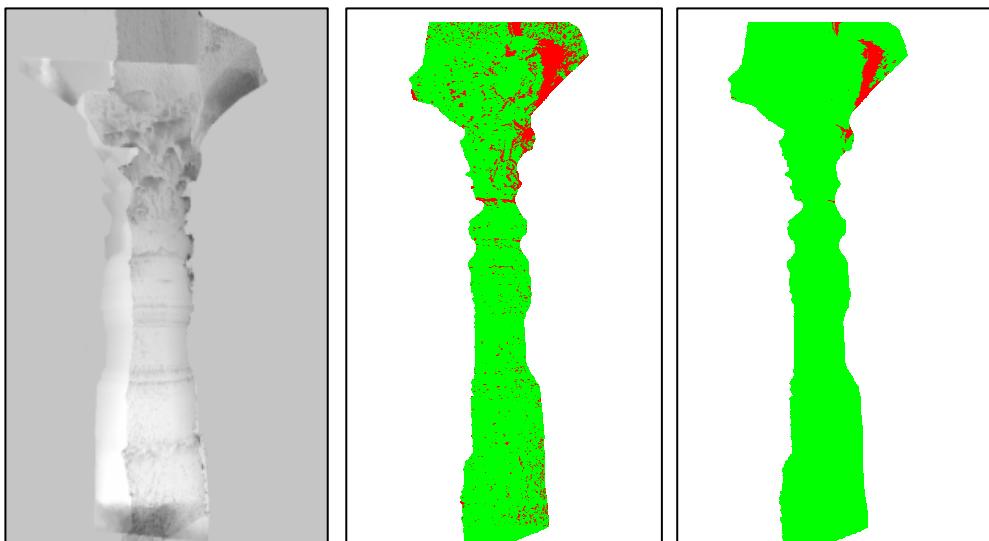


Figure 9 : Ecarts entre deux cartes. En verts : les zones où l'écart entre deux surfaces est inférieur à 1 pixel (image gauche) et à 10 pixels (image droite).

2.2. Autres approches

2.2.1. Reconstruction à partir de silhouettes

Cette approche utilise les contours occultants, correspondant à la silhouette de l'objet dans l'image. L'idée est de déliminer un volume dans l'espace se rapprochant le plus possible de l'objet observé. Cette approximation est appelée "enveloppe visuelle". Il existe deux manières de calculer cette enveloppe visuelle : soit par intersection des "cônes d'observation" (volume dans l'espace délimité par les contours dans l'image), soit en "sculptant" un volume de départ, en sélectionnant les voxels suivant leur reprojecion dans les images, à l'intérieur ou à l'extérieur du contour.

Ces techniques sont intéressantes sur des structures de petites ou moyennes dimensions, pour lesquelles il est possible d'utiliser une table tournante ou de déplacer une caméra autour de l'objet, et de définir un fond pouvant être facilement soustrait dans les images afin de mieux définir les contours de l'objet.

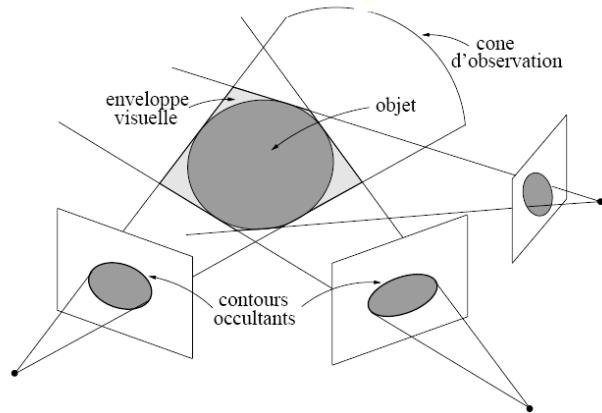


Figure 10 : Enveloppe visuelle.

L'enveloppe visuelle d'un objet observé à partir de plusieurs caméras est, par définition, le volume maximal dans l'espace compatible avec les silhouettes de l'objet dans les images. Son estimation repose donc sur le calcul de ce volume, intersection des cônes d'observation issus des silhouettes. Avec cette méthode, on ne peut pas reconstruire les concavités des objets.

2.2.2. Reconstruction à partir de la texture

Les méthodes de reconstruction à partir de la texture sont des méthodes voxéliques de « sculptage » d'un volume (ou « **space carving** ») dans l'espace suivant un critère de photoconsistance. Chaque voxel de la surface est projeté dans toutes les images où il est visible. Un voxel disparaît si ses projections sur les images ne sont pas de la même couleur. L'opération est répétée jusqu'à ce que le creusement converge vers la forme non vide.

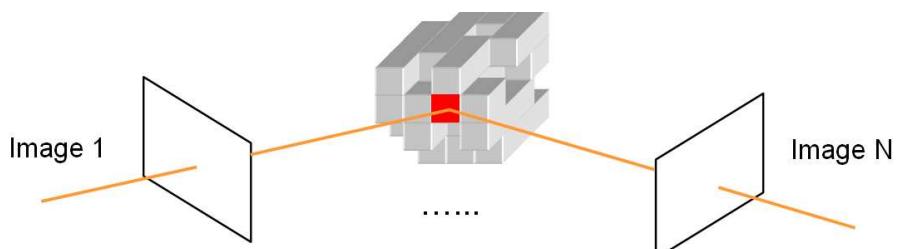


Figure 11 : « Space carving ».

Les avantages de ce genre de méthodes sont la simplicité des calculs et le texturage immédiat du modèle 3D. Mais il y a également les inconvénients comme la sensibilité aux changements d'éclairage et le temps de calcul élevé pour obtenir une bonne définition. D'autres méthodes voxéliques se basent sur l'évolution de surface par EDF (exemple « Level Sets » pratiqué au laboratoire CERTIS de l'Ecole Nationale des Ponts et de Chaussées). Dans cette méthode, l'espace est balayé avec une frontière mouvante représentée par une équation aux dérivées partielles (EDP). L'évolution de cette frontière est pilotée par la mesure de consistance de couleur. Cette approche est beaucoup plus sophistiquée, de point de vue mathématique, mais moins complexe à mettre en œuvre.

3. METHODES DE FUSION EXISTANTES

3.1. Méthodes volumiques.

Ces méthodes ont été appelées volumiques car les traitements qu'elles contiennent se déroulent directement dans l'espace 3D. Les deux méthodes principales sont décrites dans Curless et Levoy [1] et Hilton [2]. Elles sont basées sur le même principe de construction d'une surface implicite dans l'espace 3D discrétré puis de l'extraction d'une iso-surface. Une fonction f est définie telle que sa courbe de niveau 0 contenant les points de nuage. Cette opération est réalisée pour chacun de nuages à fusionner et donne donc un ensemble de fonctions f_i . Il reste ensuite à définir une fonction f , « fusion » des f_i puis à extraire la surface recherchée.

Hilton [2] se sert d'une fonction de type distance signée en ajoutant une fonction binaire pour prendre en compte les discontinuités des nappes à fusionner. Un ensemble de règles leur permettent ensuite d'obtenir une fonction f unique, calculée en tout point de l'espace.

3.2. Méthodes surfaciques.

Dans ces méthodes, les images de profondeurs sont triangulées et superposées, puis localement fusionnées pour éliminer les données redondantes.

Les méthodes de : Soucy et Laurendeau [3], Turk et Levoy [4] et Pito [5] travaillent à partir de surfaces issues de nuages initiaux triangulés indépendamment chacun avec leur topologie. En général, elles respectent le même schéma algorithmique: d'abord la détection des zones en recouvrement, ensuite la suppression des données redondantes, puis raccordement des surfaces pour créer une surface unique.

4. METHODE DE FUSION PROPOSEE

4.1. Données en entrée

En connaissant les paramètres intrinsèques et extrinsèques de la camera, on peut calculer pour chaque pixel de l'image le rayon perspectif correspondant dans l'espace terrain. L'information sur sa profondeur est contenue dans la carte de profondeur. Par conséquence, pour chaque pixel de l'image on peut déterminer les coordonnées du point terrain correspondant. L'illustration ci-dessous représente la structure de données avant la fusion.

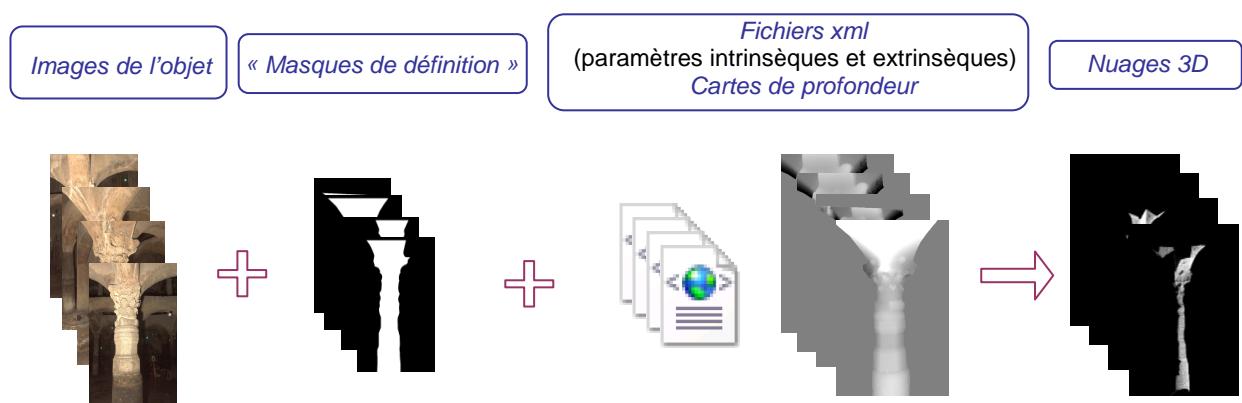


Figure 12 : Structure de données avant la fusion

Pour chaque prise de vue, on dispose :

- « Masque de définition » initial: la zone de l'image dans laquelle se trouve l'objet d'intérêt.
- Fichier xml contenant tous les paramètres de calibration de la camera
- Carte de profondeur

4.2. Critères

Le but de la méthode est de conserver l'information la plus fiable de chaque carte de profondeur. Les mesures de fiabilité se font par la vérification des critères: d'angle d'incidence, de corrélation, d'étirement de triangle. Premièrement, il est nécessaire de déterminer pour chaque point de vue la partie visible de la surface reconstruite à partir de tout point de vue voisin. Cela nous permet de dire si deux points image provenant de deux points de vue différents se correspondent, c'est-à-dire, s'ils appartiennent à la même surface en terme du modèle. Enfin, on pourra appliquer les critères de fiabilité.

4.2.1. Critère de visibilité

La première approche de ce critère s'est appuyé sur le calcul de la distance euclidienne entre deux points 3D (description détaillée incluse dans le paragraphe « Approche directe »). Elle n'a pas été efficace pour tous les objets testés et une nouvelle solution basée sur un Z-buffer a été proposée. On bascule un nuage de points 3D dans la géométrie d'un autre nuage (« nuage maîtresse ») en ne conservant que les points visibles à partir du point de vue du « nuage maîtresse ».

L'image de gauche de la figure 13 représente deux nuages de points 3D: A et B dans leurs géométries d'origine. L'image de droite illustre le résultat de la mise en place du Z-buffer.

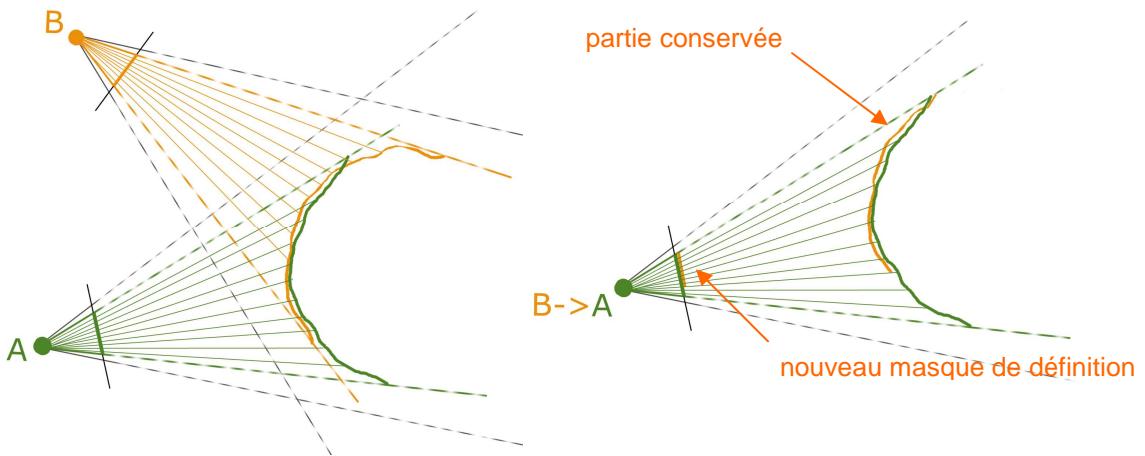


Figure 13 : A gauche : Deux nuages 3D. A droite : résultat de la mise en place du Z-buffer (nuage B basculé dans la géométrie du nuage A).

Le basculement se fait tout d'abord par une triangulation du « nuage à basculer» et ensuite par la reprojecion des sommets du triangle obtenus dans la géométrie du « nuage maitresse ». Enfin, les nouvelles coordonnées des sommets sont interpolées.

La détermination des parties visibles et des parties cachées est gérée par le Z-buffer. L'idée générale est de conserver pour chaque pixel d'un « masque de définition » du « nuage maîtresse » les points les moins profonds du « nuage à basculer ». Une condition additionnelle est l'élimination des faces arrière. Il repose sur le calcul de la normale à un petit élément de surface et leur mise en correspondance avec le vecteur issu du point d'observation se dirigeant vers la surface. En calculant le produit scalaire de ces vecteurs on détermine un critère de visibilité (positif->face cachée, négatif-> face visible).

La figure ci-dessous illustre un exemple de l'image issue du basculement.

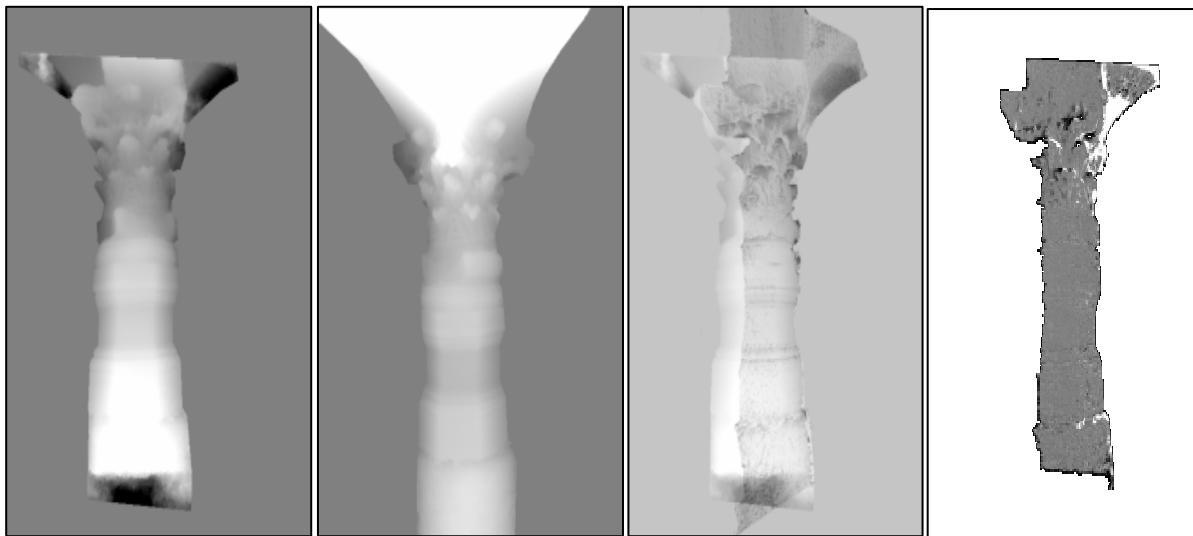


Figure 14: De gauche à droite: 1) nuage de points A ; 2) nuage de points B ; 3) basculement du nuage B dans la géométrie du nuage A ; 4) écart de profondeur entre les deux nuages

Le basculement nous permet d'une part, de traiter à un moment donné tous les nuages dans la même géométrie, d'autre part, de gérer une fois pour toute de manière rigoureuse les problèmes d'occlusion.

Dans la suite du document, on va appeler les points homologues, deux points de points de vue différents appartenant à la même surface au niveau du modèle complet et en respectant le critère de visibilité.

4.2.2. Critère d'angle d'incidence

L'angle d'incidence est défini comme l'angle que fait un rayon incident avec la normale à la surface au point d'incidence. Le processus de corrélation donne les meilleurs résultats tant que sa valeur est faible (vignette « horizontale »).

La figure 15 montre un exemple de l'influence de l'incidence sur la reconstruction. Les images de haut représentent les valeurs d'angle d'incidence calculées pour deux cartes. La couleur plus claire signifie la valeur d'angle d'incidence moins importante. Le contour rouge indique la zone de recouvrement de ces deux cartes.

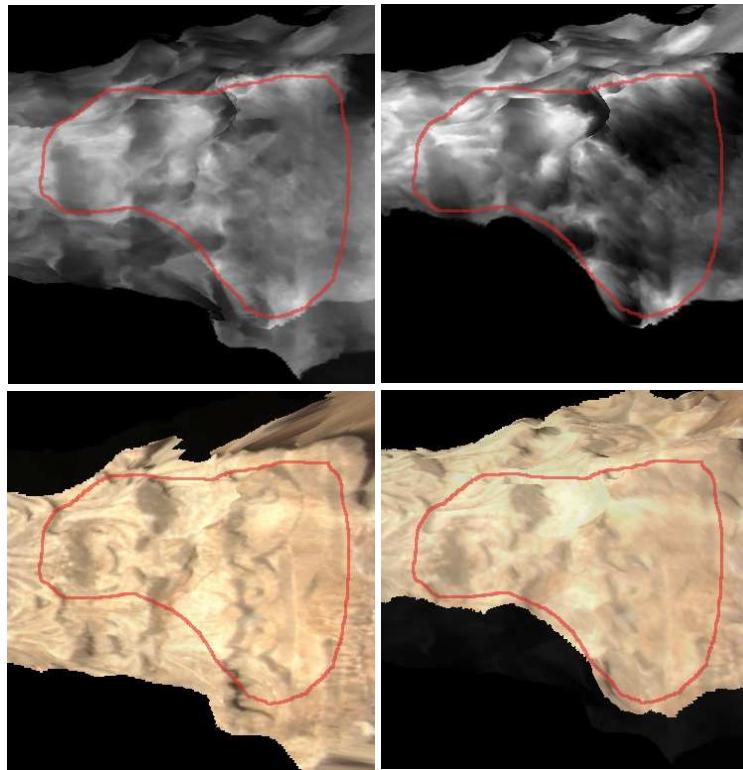


Figure 15 : Représentations de la même surface. A gauche : angle d'incidence plus fiable. A droite : valeur d'angle d'incidence plus importante.

La surface de l'objet reconstruit caractérise des petites variations (bruit). La méthode de lissage de la surface a été implémentée directement dans le calcul de l'angle d'incidence. La valeur de « lissage » est un des paramètres de la fusion. Plus sa valeur est élevée, plus la surface est lissée.

Calcul d'une carte des angles d'incidence :

Lorsque le calcul se fait dans la géométrie de l'image, on a une information directe sur le voisinage d'un pixel p . Cela apporte un grand avantage par rapport à la géométrie objet pour laquelle toutes les mesures du voisinage sont difficiles à mettre en place.

Pour chaque pixel p du « masque de définition »

- on projette $p(x, y)$ dans l'espace d'objet, on obtient un point $P(X, Y, Z)$
- on projette les points voisins $p'(x + "lissage", y)$ et $p''(x, y + "lissage")$ du point p , on obtient les points $P'(X', Y', Z')$ et $P''(X'', Y'', Z'')$
- on calcule la normale au plan $PP'P''$
- on calcule l'angle entre le rayon incident et la normale

La figure 16 illustre les cartes d'angle d'incidence obtenues pour différentes valeurs du lissage.

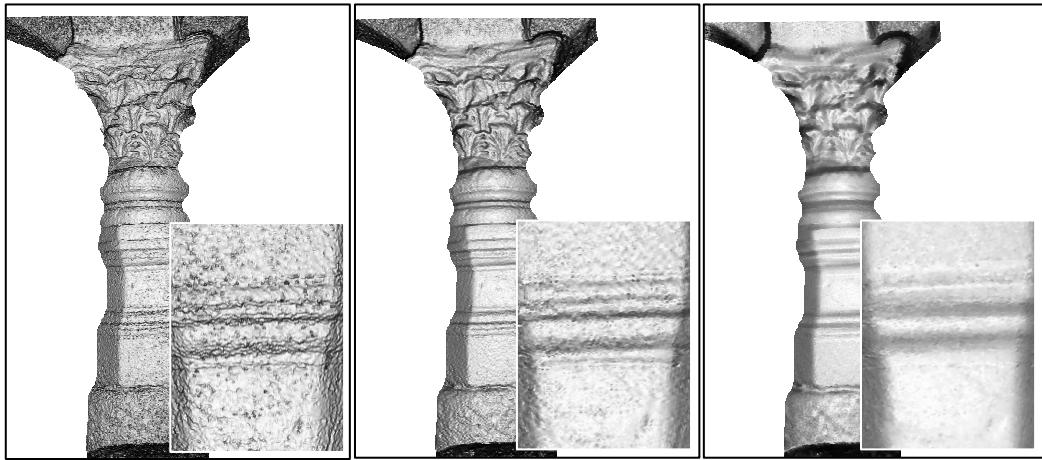


Figure 16 : Cartes d'angle d'incidence. De gauche : sans lissage. Au centre : lissage de 5 pixels. De droite : lissage de 15 pixels.

En comparant les valeurs de l'angle d'incidence pour deux points homologues et en éliminant le point pour lequel l'angle d'incidence est le plus significant, la partie rouge de l'objet à été attribuée au point de vue A et la partie verte au point de vue B (figure 17).

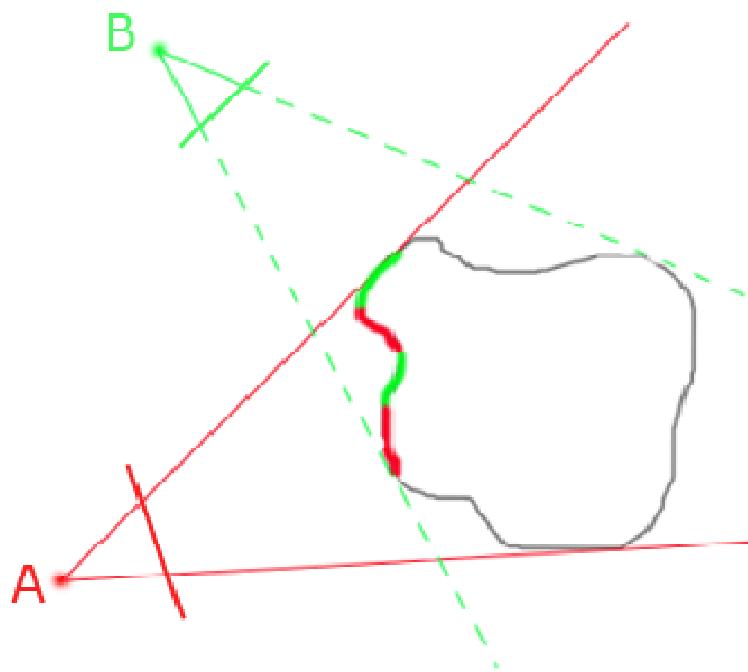


Figure 17 : Critère d'angle d'incidence.

La figure 18 montre le résultat de la mise en place du critère d'angle d'incidence. L'image gauche représente les valeurs de l'angle d'incidence pour un point de vue donné. L'image droite illustre le « masque de définition » ne contentant que les points retenus. On peut observer que la partie centrale de la colonne où l'angle d'incidence est proche de zéro (les valeurs le plus claires sur l'image gauche de la figure 18) a été conservée.

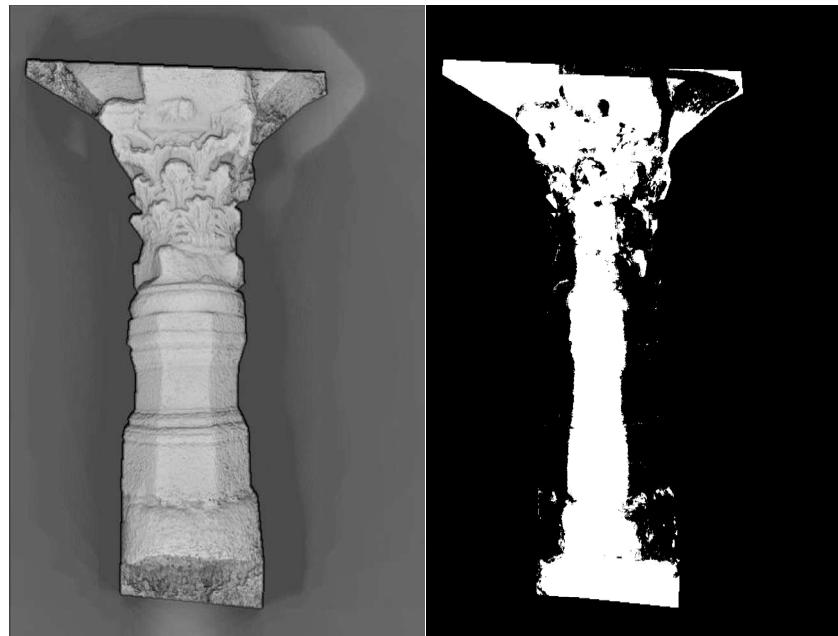


Figure 18 : L'image de gauche : carte d'angle. L'image de droite : « masque de définition » retenu.

4.2.3. Critère de corrélation

La valeur de coefficient de corrélation est une autre mesure de fiabilité d'un point. On considère que ce critère est moins significatif que le critère d'angle d'incidence car le score de corrélation a déjà été optimisé dans le processus de la réalisation d'une carte de profondeur. Ce critère n'est utilisé que dans le cas où la distance entre les deux points homologues est supérieure à la précision de reconstruction (seuil de « même » surface). Cela signifie que la position d'un des points est erronée. En général, les erreurs locales du modèle sont dues aux défauts de corrélation : l'absence de texture significative, les fortes discontinuités de la surface. Par ailleurs, il est quasiment uniforme pour tous les points de vue car le fait qu'il n'y ait pas de texture en un point n'est pas lié au point de vue.

Pour chaque prise de vue on dispose d'une carte de corrélation. Les valeurs possibles du coefficient de corrélation (-1 ;1) sont étalonnées sur les 256 valeurs d'une image de niveau de gris. Parmi les deux points homologues, on ne conserve que celui ayant le coefficient de corrélation plus élevé.

La figure 19 illustre un exemple d'une carte de corrélation.

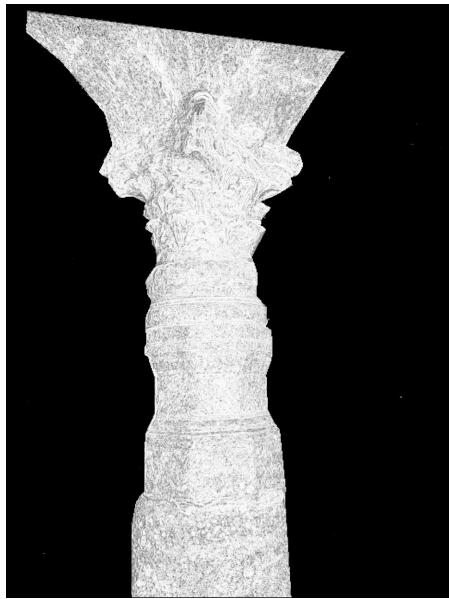


Figure 19 : Exemple de la carte de corrélation.

Les zones de mauvais score de corrélation se centralisent vers les parties où la surface n'est pas assez texturée ou bien il existe des grandes discontinuités du modèle.

4.2.4. Critère d'étirement de triangle

Le critère d'étirement de triangle est relié au basculement. Ce critère a pour but de gérer encore mieux les discontinuités de la surface.

L'idée générale est de quantifier la déformation d'un cercle qui est soumis à une transformation affine. Un cercle est associé à chaque triangle du nuage « à basculer » ensuite il se projette comme une ellipse dans la géométrie du nuage « maîtresse ». L'image de droite de la figure ci-dessous présente une carte d'étirement de triangle. Les zones plus claires signifient les valeurs d'étirement importantes donc la présence des triangles de discontinuités.

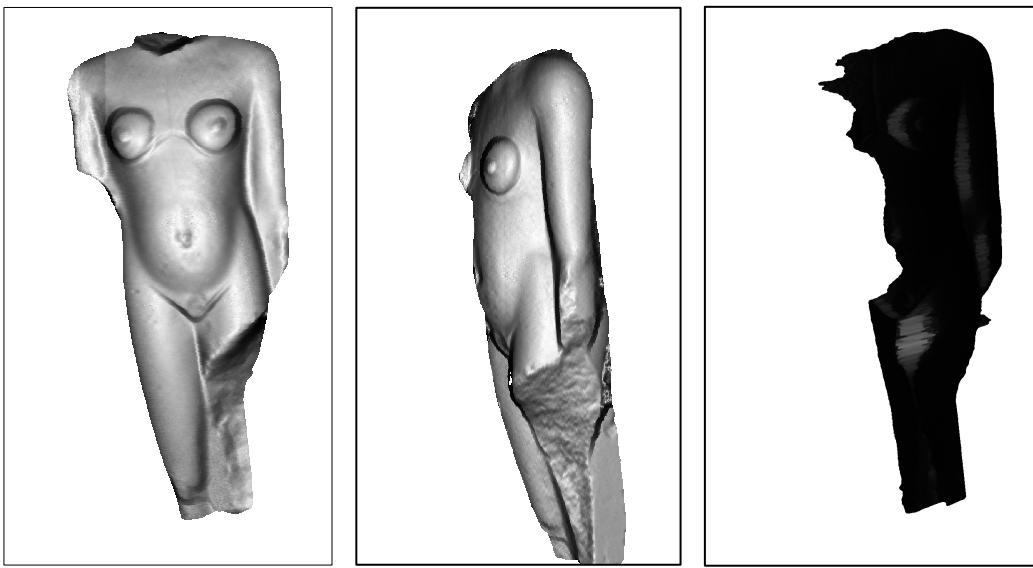


Figure 20 : De gauche à droite: 1) nuage « maîtresse » ; 2) nuage à basculer ; 3) carte d'étirement de triangle

4.3. Approche avec basculement

Cette approche a été appelée « avec basculement » car le problème de visibilité a été résolu par la mise en place du basculement des nuages. L'avantage de cela est le fait qu'on traite à un moment donné tous les nuages dans la même géométrie et qu'on gère ainsi une fois pour toute et de manière rigoureuse les problèmes d'occlusion.

4.3.1. Stratégie générale

Le traitement se décompose en trois étapes principales :

- Préfiltrage
- Application des critères
- Filtrage morphologique

Le préfiltrage a pour but de supprimer l'information la plus faible de chaque carte. Ensuite, on applique les critères pour sélectionner les données les plus fiables. Enfin, on rend les masques contenant les pixels conservés plus régulières par le filtrage morphologique.

La figure ci-dessous présente l'enchaînement du traitement simplifié

M_1, M_2, \dots, M_k : « masques de définition »

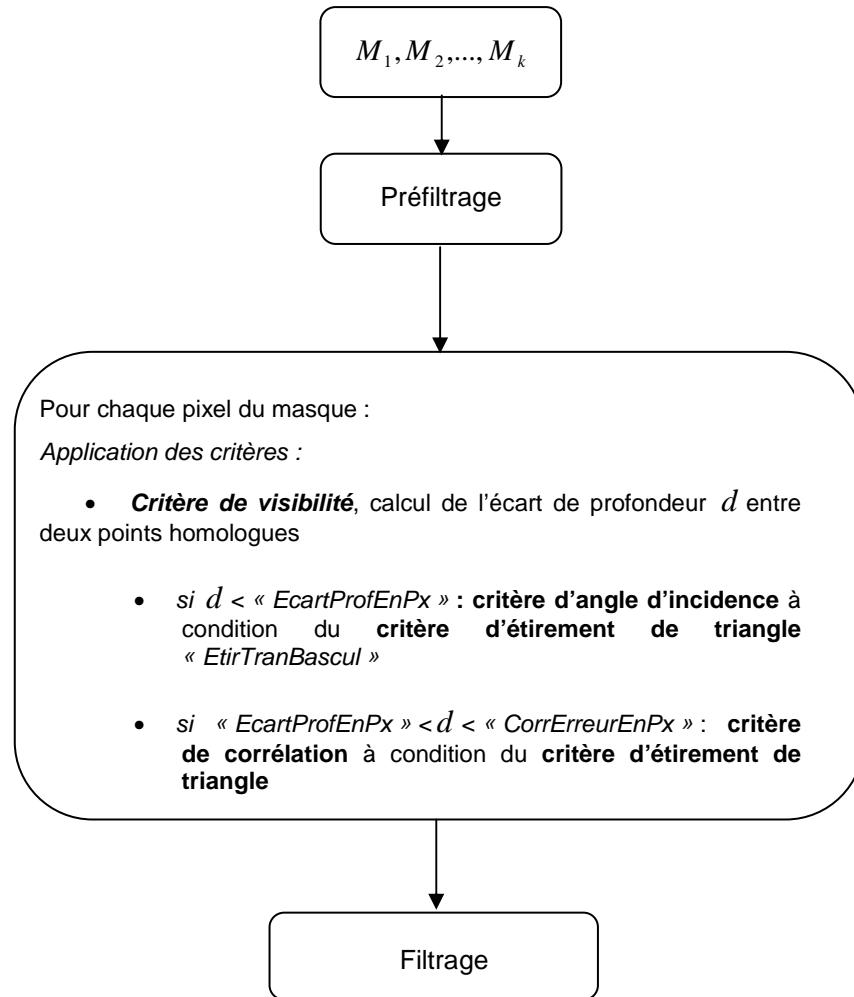


Figure 21 : L'enchaînement de l'approche avec basculement.

Description des étapes :

Etape 1 : Préfiltrage de chaque « masque de définition »

Cette étape consiste à supprimer de l'information la plus faible. La description détaillée est inclue dans le paragraphe suivant.

Etape 2 : Pour chaque « masque de définition »

Pour tout pixel du masque :

- vérification du **critère de visibilité** :
 - basculement des nuages, recherche des pixels correspondants dans les nuages basculés

- calcul de l'écart de profondeur d entre deux points homologues
 - si $d < \text{« EcartProfEnPx »}$ (les deux points appartiennent à la même surface) vérification du ***critère d'angle d'incidence*** (point avec plus petit d'angle d'incidence est gardé) à condition du ***critère d'étirement de triangle*** (valeur faible, inférieure à « EtirTrianBascul »)
 - si « EcartProfEnPx » $< d < \text{« CorrErreurEnPx »}$ les deux points appartiennent à la même surface, mais leurs positions sont erronées à cause d'erreurs de corrélation) vérification du ***critère de corrélation*** (point avec plus grande valeur de coefficient de corrélation est gardé) à condition du ***critère d'étirement de triangle***
 - si $d > \text{« CorrErreurEnPx »}$ (les deux points n'appartiennent pas à la même surface donc les deux points sont gardés)

Etape 3 : Filtrage morphologique (régularisation des masques)

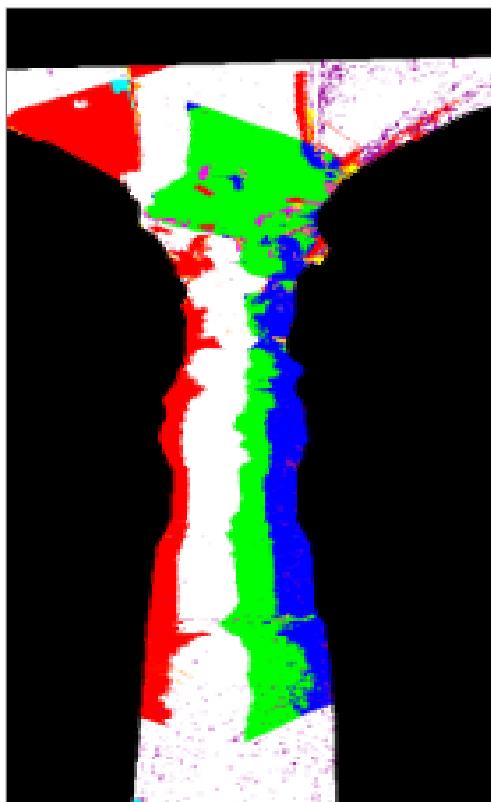
- Erosion géodésique : enlever des petits artefacts
- Dilatation géodésique : boucher les petits trous

La taille des trous bouchés et des artefacts supprimés dépendent de la taille et de la forme de l'élément structurant du filtre (« *ElemStructAngle* »).

Le choix des paramètres « *EcartProfEnPx* » et « *CorrErreurEnPx* » est dicté par la précision des cartes de profondeur. Sachant que la résolution d'une carte de profondeur vaut un pixel. On considère que les deux points homologues appartiennent à la même surface si leur écart de profondeur (« *EcartProfEnPx* ») est inférieur à 3 pixels avec le lissage de 10 pixels. L'erreur de corrélation est estimée comme valeur inférieure à 50 pixels.

Pour aider, d'une part, l'interprétation des résultats et l'affinement des paramètres de la fusion, d'autre part, la validation de différentes étapes de la programmation, le « masque de définition » a été classifié suivant les critères présentés dans les parties précédentes.

La figure ci-dessous présente le résultat de la classification d'un « masque de définition » suivant les critères. Les pixels : blancs, verts, verts clair et magentas correspondent aux parties conservés du masque.



Légende

Préfiltrage

- [cyan] valeur de l'angle d'incidence supérieur à « AngleIncidMax »
- [purple] valeur du coefficient de corrélation inférieur à « CorrelCoeffMin »

Fusion

- [white] pixels n'appartenant qu'à ce « masque de définition »
- [red] pixels reprojetés des masques « calculés »
- [green] pixels des meilleurs angles
- [blue] pixels des mauvais angles
- [magenta] pixels d'erreur de corrélation
- [light green] pixels des mauvais angles mais étirement de triangle inférieur à « EtirTrianBascul »
- [yellow] pixels d'erreur de corrélation angles mais étirement de triangle inférieur à « EtirTrianBascul »

Figure 22 : Segmentation d'un “masque de définition”.

4.3.2. Préfiltrage

Le préfiltrage a été appliqué avant le lancement de l'algorithme de la fusion. Il a pour but d'enlever les pixels pour lesquels la fiabilité est faible. De nouveaux paramètres ont été introduits : la valeur minimale du coefficient de corrélation (« CorrelCoeffMin ») et la valeur maximale de l'angle d'incidence (« AngleIncidMax »). Tous les pixels ayant le score de corrélation inférieur à « CorrelCoeffMin » ou l'angle d'incidence supérieur à « AngleIncidMax » sont supprimés du « masque de définition ».

4.3.3. Filtrage morphologique

Lorsque, le calcul d'un nouveau masque se fait pixel par pixel, le masque obtenu n'est pas régulier. Pour rendre les nouveaux « masques de définition » plus homogènes, la reconstruction géodésique a été appliquée. Les deux phases du filtrage sont les suivantes :

1. Erosion géodésique qui a pour effet de supprimer des petits artefacts
2. Dilatation géodésique qui a pour effet de boucher des petits trous

La réalisation d'une reconstruction géodésique exige dans un premier temps, la construction de l'image des marqueurs et dans un deuxième temps l'intégration de cette information dans l'image originale. L'érosion géodésique consiste à effectuer une succession d'érosions sur l'image marqueur et de combiner le résultat de ces érosions avec l'image originale. Le résultat de la reconstruction géodésique correspond à l'image du marqueur érodée et conditionnée par l'image originale. La taille des structures est soumise à des modifications dépendant de la forme et de la taille d'élément structurant de l'érosion. Le principe de la dilatation géodésique est le même, sauf qu'à la place de l'érosion on applique la dilatation. Plus d'information théorique sur l'érosion et la dilatation géodésique sont contenus dans l'annexe 5.

La figure 23 montre un résultat du filtrage morphologique.

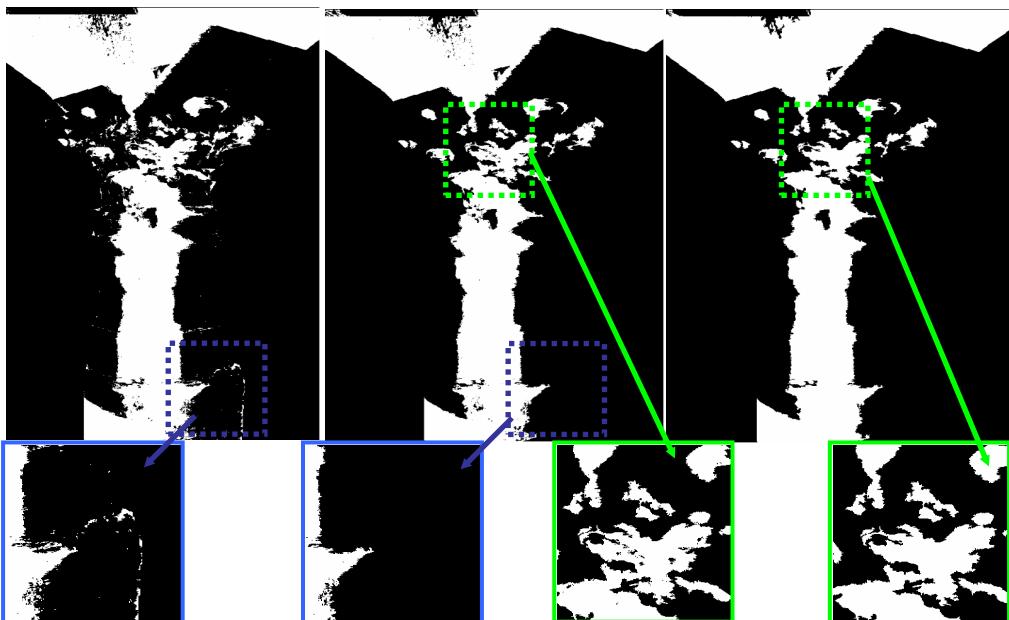


Figure 23 : A gauche : « masque de définition » sans filtrage. Au centre : érosion géodésique appliquée à l'image (les artefacts dans le carré bleu ont disparus). A droite : érosion et dilatation appliqués à l'image (les « trous » dans le carré vert ont été bouchés).

4.3.4. Reprojection d'un masque déjà calculé

L'objectif général de la méthode est la non redondance. Pour cela, on considère que si un point est conservé dans une image, les points homologues dans d'autres nuages seront interdits.

La stratégie suivie est la suivante. D'abord on étiquette les masques. Au départ, tous les masques possèdent l'étiquette «non calculé » (contour vert - Figure 24). Ensuite, pour chaque pixel d'un masque on cherche ces correspondants (critère de visibilité) puis on applique nos critères. On obtient un nouveau masque avec les points conservés et on l'étiquette « masque calculé » (contour rouge). Par la suite, on calcule le masque suivant, on cherche les points homologues pour ces pixels. Si le point correspondant appartient au masque déjà calculé on le supprime du masque courant, si non on applique les critères de qualité. On procède de la même manière avec le reste des masques.

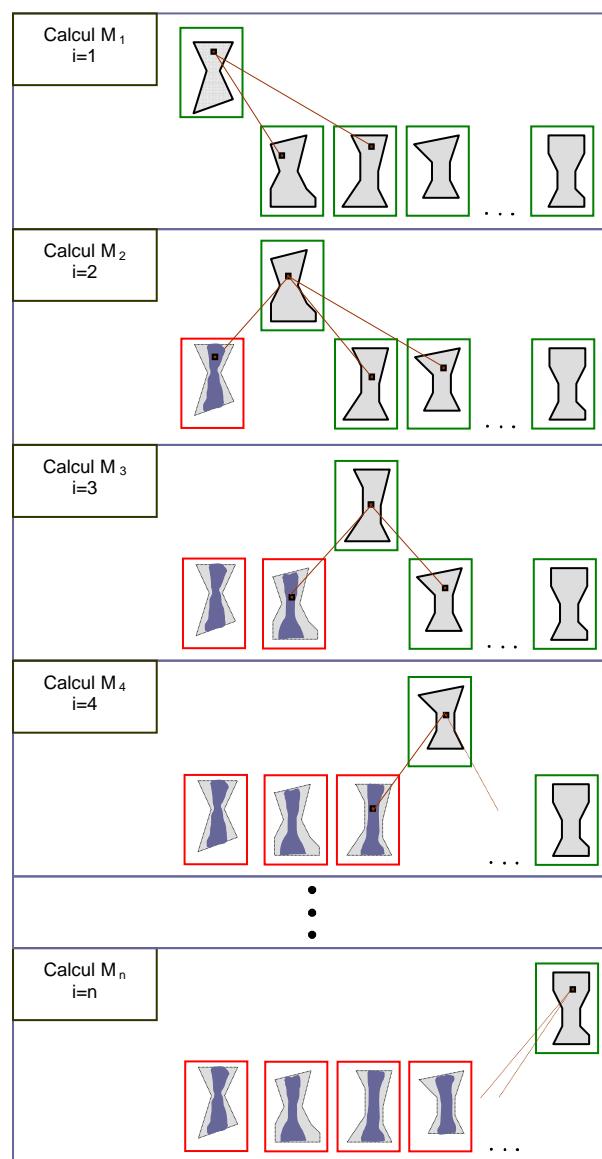


Figure 24 : L'algorithme de reprojection.

4.3.5. Schéma de l'algorithme détaillé

Algorithme, solution retenue :

M_1, \dots, M_n : ensemble de « masques de définition »

C_1, \dots, C_n : ensemble de cartes de corrélation

$i = 1, \dots, n$;

$k = 1, \dots, n$; où n est un nombre de « masque de définition »

$p = 1, \dots, m$; où m est un nombre de pixels dans un « masque de définition »

$C_{i,p}$: cette écriture signifie la valeur de $p^{ième}$ pixel appartenant à la $i^{ième}$ « carte de corrélation »

Premier étape : Calcul des cartes d'angle d'incidence.

Paramètres :

« lissage »: lissage de la surface en pixels

$\left\{ \begin{array}{c} n \\ i=1 \end{array} \right\} \left\{ \begin{array}{c} m \\ p=1 \end{array} \right\}$ Pour chaque pixel $p \in M_i$, calcul de l'angle d'incidence $\left. \right\} \left. \right\}_i \left. \right\}_p$

Résultats : A_1, \dots, A_n (cartes d'angle) qui correspondent à chaque masque de définition

Deuxième étape : Préfiltrage.

Paramètres :

« CorrelCoeffMin » : valeur admissible du coefficient de corrélation

« AngleIncidMax » : valeur admissible de l'angle d'incidence

$\left\{ \begin{array}{c} n \\ i=1 \end{array} \right\} \left\{ \begin{array}{c} m \\ p=1 \end{array} \right\}$ Pour chaque pixel $p \in M_i$: $\left. \right\} \left. \right\}_i \left. \right\}_p$

- ✓ si $A_{i,p} > \text{« AngleIncidMax »}$ ou $C_{i,p} < \text{« CorrelCoeffMin »}$: suppression p de M_i
- ✗ si non : $p+1$

Résultats : « masques de définition » M_1, \dots, M_n : filtrés

Troisième étape : Calcul des nouveaux « masques de définition ».

Paramètres :

« EcartProfEnPx » : écart acceptable entre les points appartenant à la même surface (en pixels)

« CorrErreurEnPx » : écart acceptable entre les points appartenant à la même surface mais avec une erreur de corrélation (en pixels)

« EtirTrianBascul » : valeur acceptable de l'étirement de triangle

« ElemStructAngle » : taille de l'élément structurant pour le filtrage morphologique de la zone conservée

« ElemStructZoneReproj » : taille de l'élément structurant pour le filtrage morphologique de la zone reprojeté

$\left\{ \begin{array}{l} \text{Etiquetage du } M_i, \text{ « masque non calculé »} \\ i=1 \end{array} \right.$

$\left\{ \begin{array}{l} \text{Pour chaque couple } M_i \text{ et } M_j \text{ où } i \neq j : \\ j=1 \end{array} \right.$

- Basculement M_j vers M_i
- Calcul d'une carte d'étirement triangle E_{ij}

$\left\{ \begin{array}{l} \text{Pour chaque } p \in M_i \\ p=1 \end{array} \right.$

✓ si $p \in M_j$: critère de visibilité, calcul de d (écart entre les profondeurs de $p \in M_i$ et de $p \in M_j$)

- ✓ si $d < « EcartProfEnPx »$
- ✓ si M_j est déjà « calculé » : suppression p de M_i
- ✗ si non : critère d'angle d'incidence :
 - ✓ si $A_{i,p} > A_{j,p}$ et $E_{ij,p} < « EtirTrianBascul »$: suppression p de M_i
 - ✗ si non : pixel conservé
- ✓ si $« EcartProfEnPx » < d < « CorrErreurEnPx »$: critère de corrélation :
 - ✓ si $C_{i,p} < C_{j,p}$ et $E_{ij,p} < « EtirTrianBascul »$: suppression p de M_i
 - ✗ si non : pixel conservé
- ✗ si non : $p+1$

$\left. \begin{array}{l} \} \\ \} \\ j \end{array} \right\} p$

Filtrage morphologique du M_i .

Etiquetage du M_i , « masque calculé »

$\left. \begin{array}{l} \} \\ i \end{array} \right\}$

Résultats : « masques de définition » M_1, \dots, M_n : après la fusion

4.4. Approche « directe »

4.4.1. Stratégie générale

La différence essentielle entre les deux méthodes : « avec basculement » et « directe » est la façon dans laquelle le problème de visibilité est traité. Dans cette approche, l'écart entre deux points homologues est calculé comme la distance euclidienne dans l'espace 3D. Si cet écart est inférieur à la précision de la reconstruction, on considère que les points appartiennent à la même surface de point de vue du modèle. Dans la solution « avec basculement », les nuages ont été mis à un moment donné dans la même géométrie et la différence entre les points correspondants a été calculée comme un écart de leur profondeur en pixels. Dans la méthode « directe » le préfiltrage n'a pas été appliqué. Le fait que les masques ont été saisis par un opérateur on peut avoir des petits écarts entre le contour saisi et le contour de l'objet. Les informations sur la profondeur en-dehors de l'objet sont fausses pour cela chaque masque de définition a été érodé de quelques pixels avant le traitement.

L'approche « directe » est le résultat d'un mois de travail. Elle a permis d'obtenir un résultat satisfaisant pour les objets bien texturés, avec les points de vue bien répartis (exemple du pilier d'Alexandrie).

4.4.2. Critère de visibilité sans basculement

Le critère de visibilité a pour le but de trouver des points homologues dans les différents nuages 3D pour ensuite pouvoir estimer leurs qualités de point de vue de modèle 3D complet. On peut caractériser un modèle 3D issu des cartes de profondeur comme un ensemble de surfaces très proches. Dans les endroits où les nuages se superposent on peut apercevoir un écart entre elles. Pour fixer un écart acceptable, il est nécessaire de déterminer les valeurs des précisions planimétrique et altimétrique du nuage. Donc :

- seuil de « même » surface de points de vue du modèle (seuilMS) :
Précision altimétrique : la valeur qui correspond à l'écart d'un pixel de la carte de profondeur, elle a été calculée globalement pour chaque modèle
Précision planimétrique : résolution de la carte de profondeur
- seuil d'erreur de corrélation (seuilCE) : dix fois la valeur du seuil de « même » surface

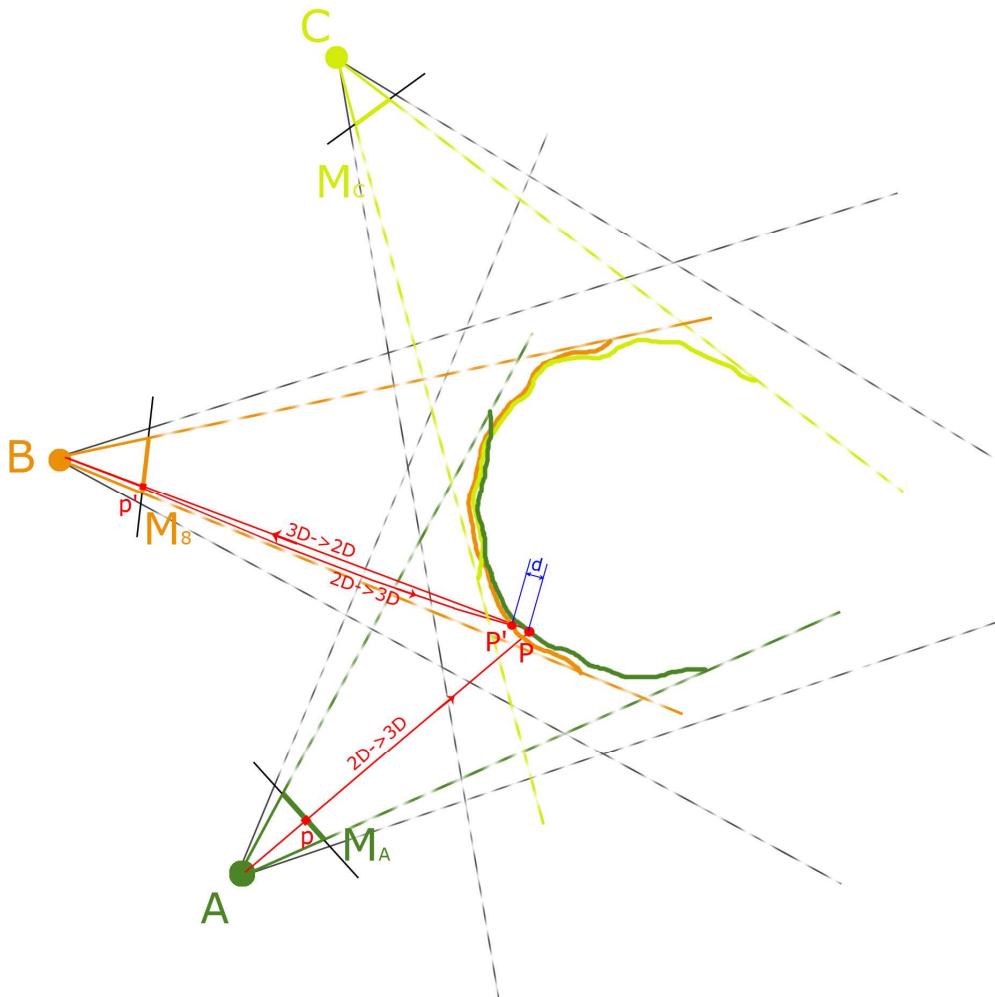


Figure 25 : Critère de visibilité.

Soit un pixel p appartenant au masque M_A . On calcule ses coordonnées P dans le nuage de points 3D correspondant à la prise de vue A (partie verte de l'objet). Ensuite, le point P est reprojeté dans la géométrie de la prise de vue B et on obtient en les coordonnées P' . Si le point P' se trouve dans le masque M_B , on le projette dans la géométrie de prise de vue B donc on obtient le point 3D P' . Enfin, on calcule la distance euclidienne d entre les points P et P' . Si la distance d est inférieure au seuil de la même surface (seuilMS) on passe au critère d'angle d'incidence. Par contre, si la distance d est supérieure au seuilMS et inférieure au seuil d'erreur de corrélation (seuilCE) on passe au critère de corrélation. Si la distance d est supérieure au seuilCE, on suppose que les deux points P et P' n'appartiennent pas à la même surface.

La figure ci-dessous illustre le cas où la distance d est supérieure au seuil d'erreur de corrélation donc deux points n'appartiennent pas à la « même » surface.

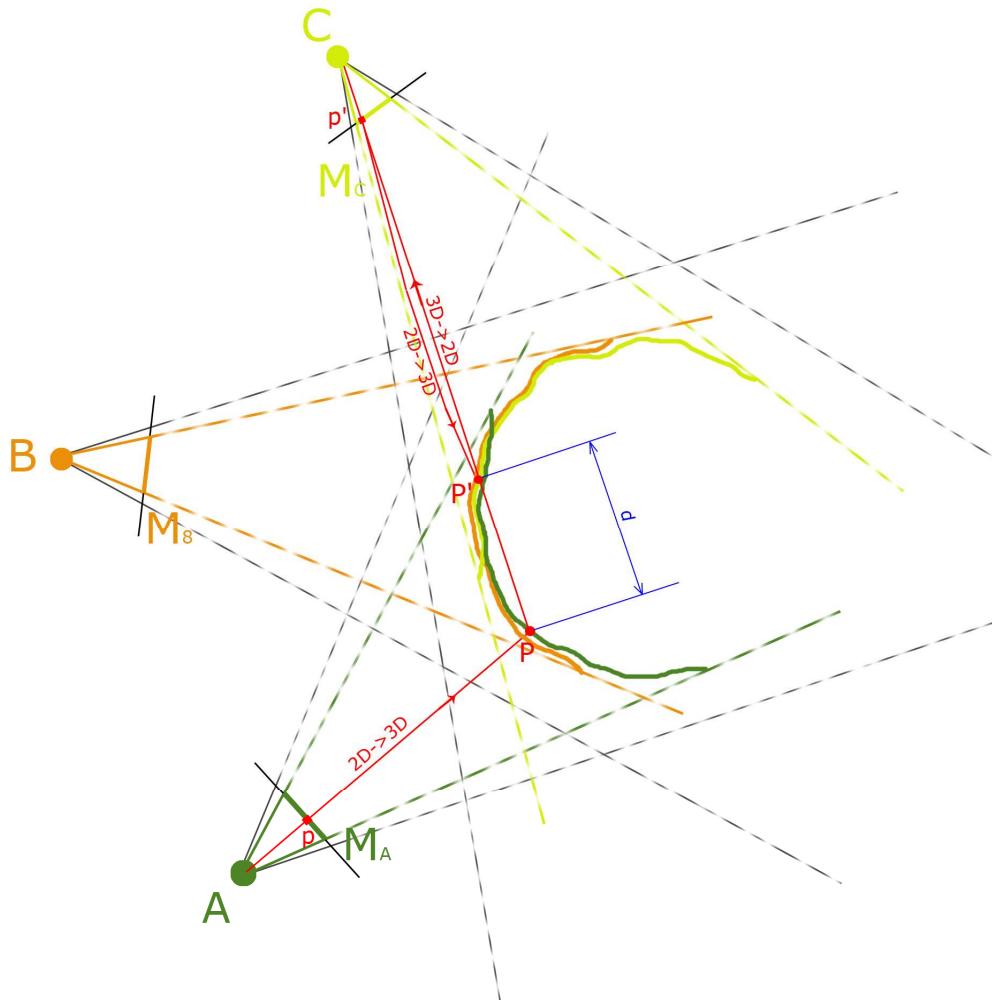


Figure 26 : Critère de visibilité. Deux points n'appartiennent pas à la « même » surface.

La figure ci-dessous montre un des cas où ce critère de visibilité n'est plus efficace. Sachant que la différence entre les points P et P' est inférieure au seuil d'erreur de corrélation, on s'aperçoit qu'on ne peut pas les classifier définitivement comme les points erronés appartenant à la même surface. Les points P et P' peuvent également être classifiés comme les points appartenant à deux surfaces différentes.

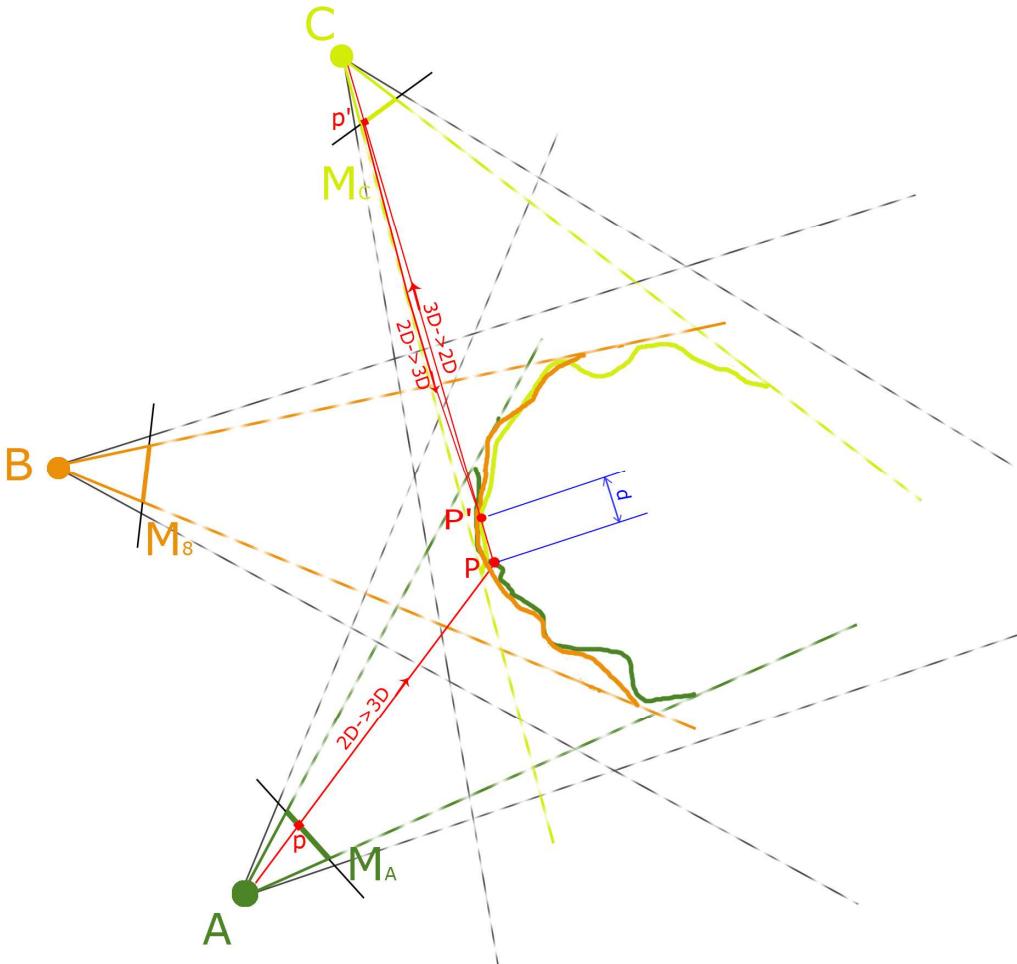


Figure 27 : Défaut du critère de visibilité.

4.5. Implémentation

4.5.1. Outils de développement

L'algorithme a été développé sous Linux en langage C++, associé par les bibliothèques :

- MicMac :

Cette bibliothèque a été élaborée par Marc Pierrot Deseilligny. Elle permet la mise en correspondance automatique (approche multi-résolution, approche par minimisation d'une fonction d'énergie).

Dans le cadre du stage, une seule classe a été utilisée. Les méthodes de cette classe assurent le passage entre l'espace 2D de l'image vers l'espace 3D de l'objet, le basculement et enfin le calcul d'«étirement de triangle».



ITK (Insight ToolKit) est une bibliothèque «open source» de traitement d'image. Elle contient des algorithmes de recalage d'image, de filtrage et de segmentation. Elle a été créée dans le but d'analyser toute sorte d'images médicales.

Cette bibliothèque a été employée dans tous les filtrages morphologiques des images : érosion, dilatation, reconstruction géodésique (géodilatation, géoerosion), ainsi, dans la totalité d'opérations sur un pixel de l'image.

Le développement de l'interface de visualisation 3D a été accompagné par les bibliothèques Qt et OpenGl.



Qt est une bibliothèque logicielle orientée objet et développée en C++ par la société Qt Software. Elle offre des composants d'interface graphique (widgets), d'accès aux données, de connexions réseaux, de gestion des fils d'exécution, d'analyse xml, etc.

Tous les widgets de l'interface et le fichier xml avec les paramètres de l'algorithme ont été développée à l'aide de cette bibliothèque.



OpenGL est une bibliothèque de rendu graphique 2D et 3D. Elle se compose de plusieurs centaines de fonctions qui permettent au programmeur la création d'images de haute qualité. OpenGL est utilisé dans de nombreux domaines, aussi bien dans les moteurs de jeux vidéo, dans des logiciels de modélisation 3D ou encore dans l'imagerie médicale.

Ces fonctions ont permis le rendu des nuages 3D, ainsi que le placage de la texture et la coupe.

4.5.2. Interface graphique et ses fonctionnalités

Au cours du stage, une interface de visualisation 3D a été élaborée. Les modèles 3D sont générés et ensuite affichées directement à partir de fichiers xml contenant les caractéristiques de chaque point de vue. L'interface (version Linux) se trouve sur le CD joint.

L'affichage des modèles fusionnés et de leurs coupes a aidé à l'interprétation et à la validation des résultats de la fusion. Une autre utilité de l'interface est la possibilité de mettre la texture sur le modèle, ainsi, d'exporter le modèle sous la forme du fichier standard (obj).

La figure 28 illustre un imprimé écran de l'interface.

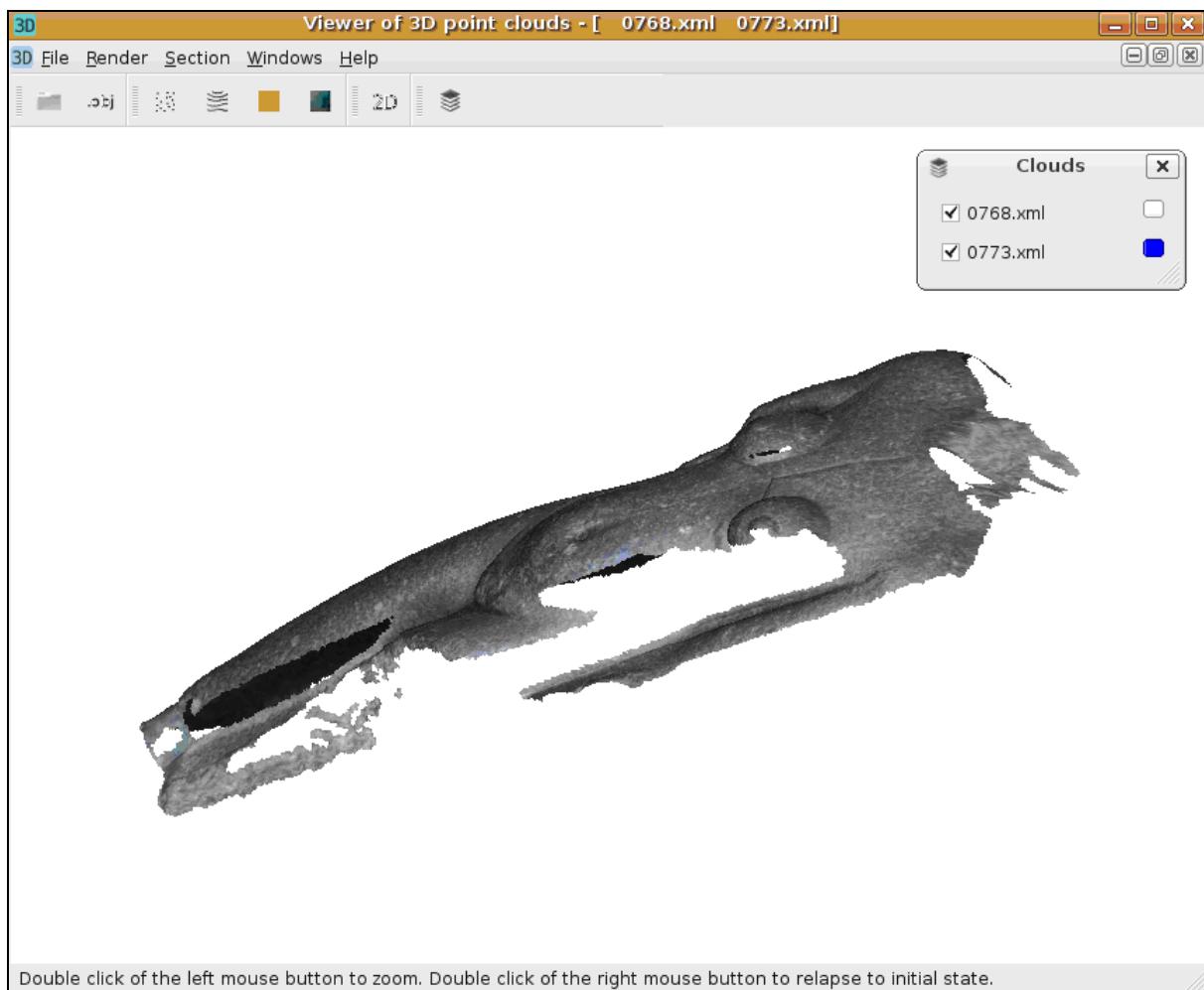


Figure 28 : Interface “Viewer of 3D point clouds”.

Description de l'interface 3D :

- Menu « *File* »
 - Load Cloud : charger un (des) nuage(s) de points 3D à partir de fichier xml. La fonction peut être également appelée par le raccourci clavier Ctrl+ 0.
 - Export : afficher la fenêtre de l'export (fichier .obj). Le raccourci clavier Ctrl+ E.



Figure 29 : Fenêtre de l'export d'un modèle 3D.

A partir de la fenêtre il est possible de choisir la forme du modèle (nuage de point 3D, modèle texturé), le répertoire et le nom de fichier de sortie. Dans le cas de l'export du modèle texturé un fichier des matériaux .mtl est généré également.

- Exit : fermer l'application. Le raccourci clavier pour cette fonction est Ctrl+ Q.
- Menu « *Render* »
 - Points : afficher le modèle sous la forme d'un nuage de points 3D
 - Lines : afficher le modèle sous la forme d'un ensemble de lignes 3D
 - Quads : afficher le modèle sous la forme d'une surface 3D
 - Texture : afficher le modèle avec la texture

- Menu « *Section* »
 - **2D** Section : afficher la coupe passante par deux points cliqués précédemment, perpendiculaire à la surface dans ces points
- Menu « *Windows* »
 -  Clouds : afficher la fenêtre contenant la liste des nuages



Figure 30 : Fenêtre des nuages

En cochant et décochant les cases, il est possible d'afficher et de masquer les nuages.

- Menu « *Help* »

Fonctions de navigation :

- Rotation : mouvement de la souris en appuyant sur son bouton gauche
- Changement d'échelle : mouvement de la molette
- Changement du centre de rotation : double clic de bouton gauche de la souris (double clic de bouton droit de la souris pour revenir à l'état initial)

4.5.3. Documentation

La documentation a été générée par le logiciel libre « Doxygen » qui permet de la créer à partir du code source du programme. Pour cela, il tient compte de la grammaire du langage dans lequel est écrit le code source, ainsi que des commentaires s'ils sont écrits dans un format particulier. La documentation est sous la forme des fichiers html. Elle contient la description détaillée de toutes les classes, leurs méthodes, attributs et les fichiers de code.

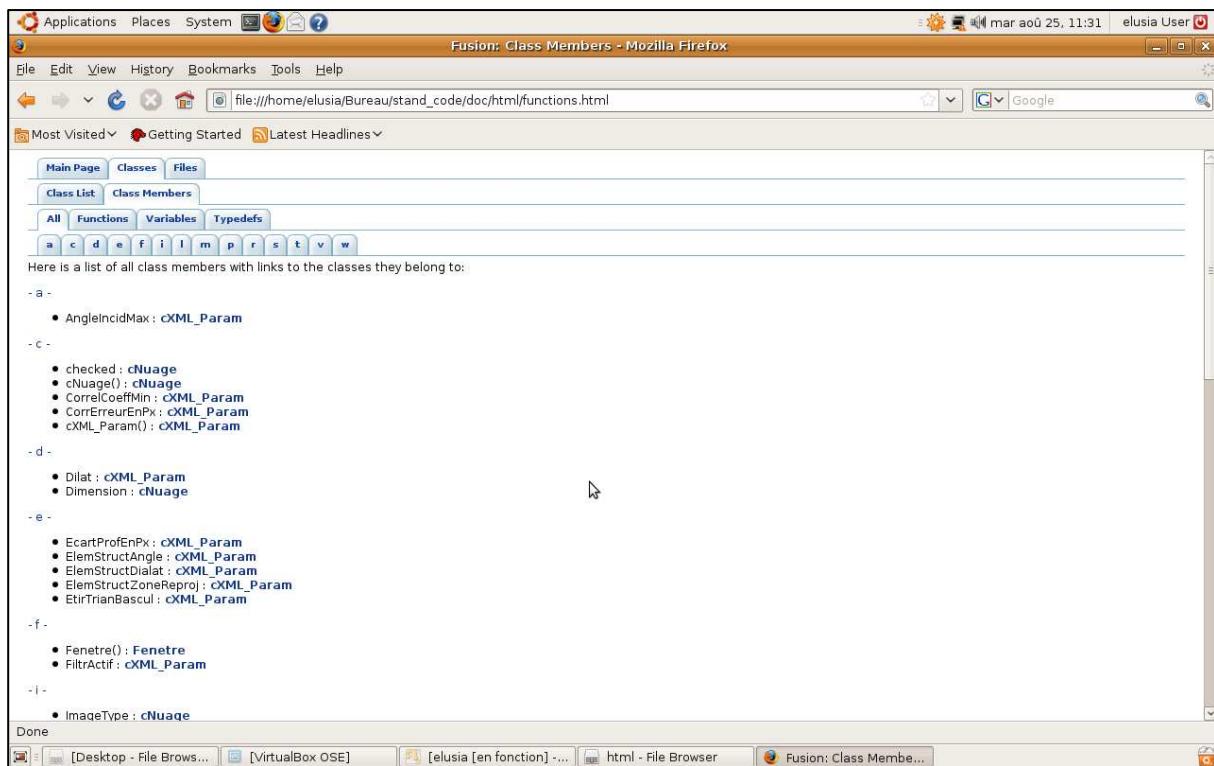


Figure 31 : Documentation du code.

La documentation se trouve sur le CD joint.

5. TESTS ET LEURS RESULTATS

Dans ce chapitre quelques exemples des résultats de la fusion ont été présentés. Le fichier standard .obj de chaque modèle texturé se trouve sur le CD joint ainsi le fichier de paramètres choisis.

5.1. Pilier (Alexandrie)

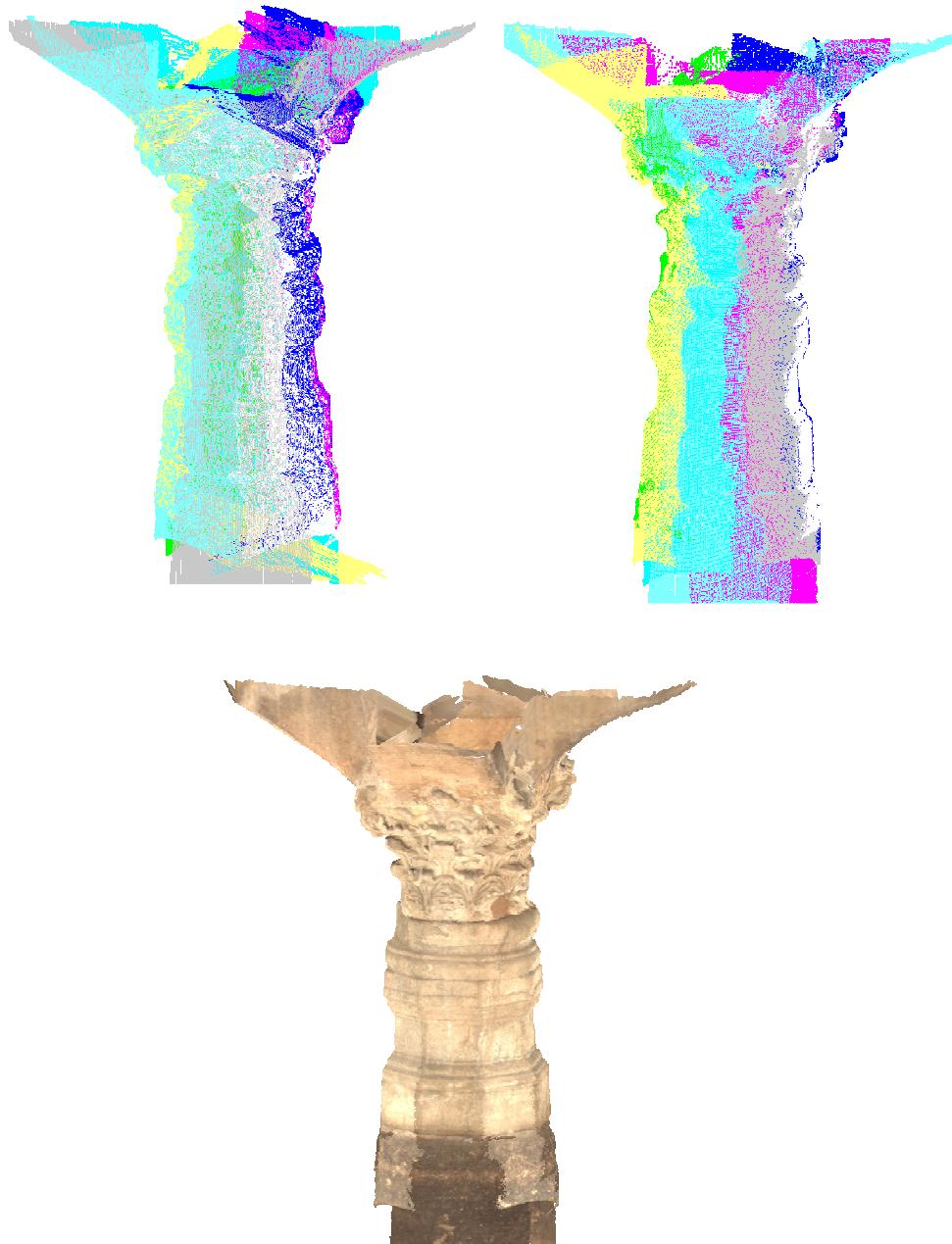


Figure 32 : Pilier d'Alexandrie. En haut : modèle avant et après la fusion. En bas : modèle après la fusion texturé

Caractéristique de l'objet : bien texturé, points de vue bien repartis, sans grande variations de surface

En général, le résultat de la fusion est satisfaisant. Cependant, on peut rencontrer des petits défauts sur la partie décorative de l'objet où les variations du vecteur normal à la surface sont importantes.

La figure ci-dessous montre un exemple de défaut retrouvé dans le résultat obtenu.

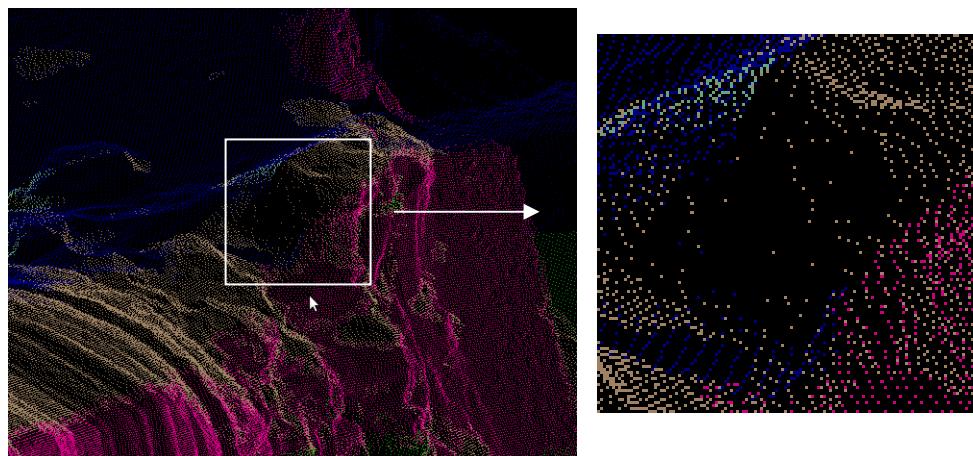


Figure 33 : Pilier. Défaut du résultat. Parties manquantes.

La source de cette erreur est issue du préfiltrage. Les points avec des valeurs importantes de l'angle d'incidence ont été rejetés donc on peut retrouver des petits « trous » dans le modèle.

5.2. Nubienne (Louvre)

Caractéristique de l'objet : bien texturé, points de vue bien repartis, sans variations importantes de la surface

Le modèle de la Nubienne obtenu après la fusion a été très « propre ». Aucun défaut remarquable n'a été trouvé.



Figure 34 : Nubienne du Louvre. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturée

5.3. Stalagmite (Grottes de Chorance)

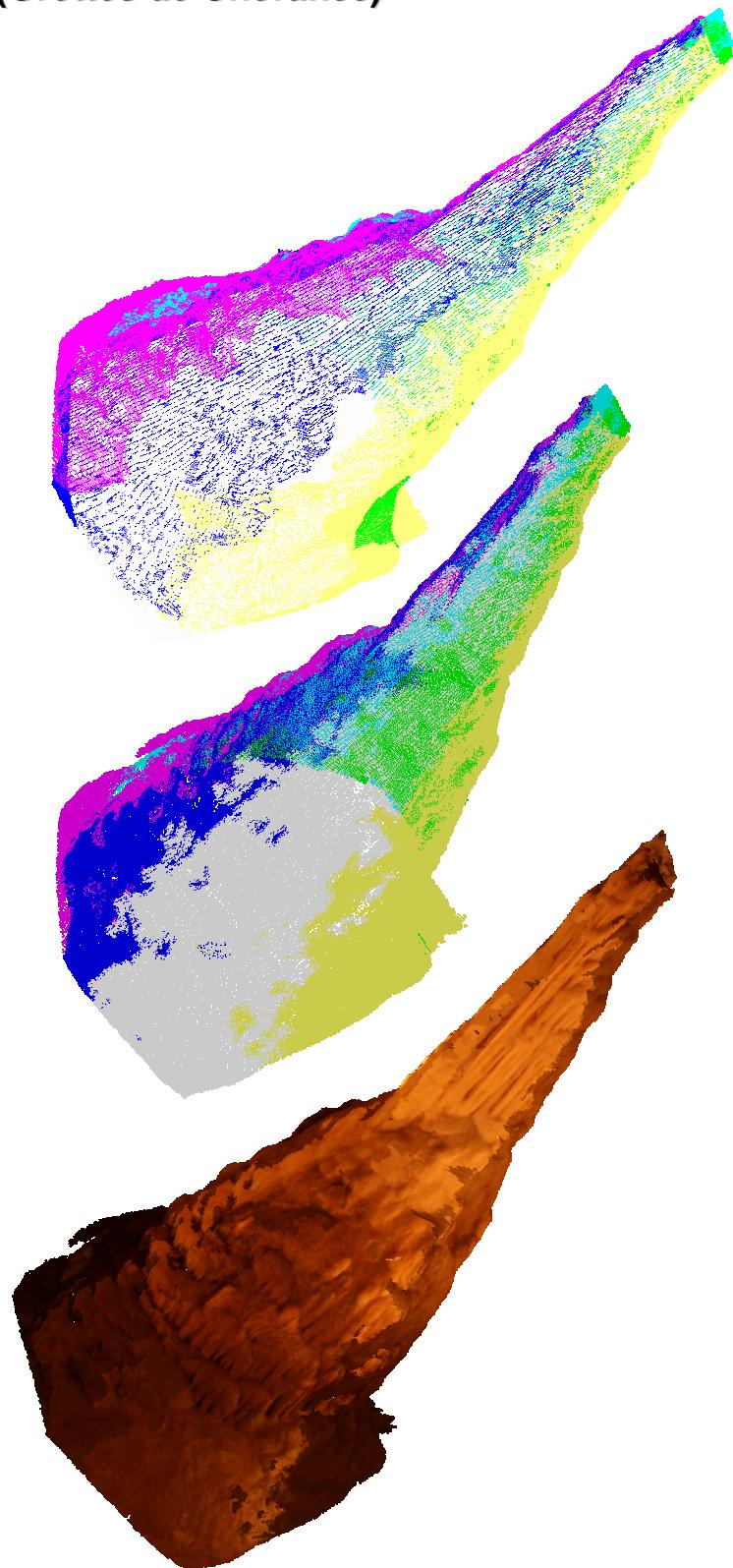


Figure 35 : Stalagmite en Alpes. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturée

Caractéristique de l'objet : bien texturé, points de vue reparties correctement, sans variations importants de la surface, par contre nuages de points bruités

Le modèle obtenu assez correct. Par contre, on peut retrouver des petits « trous » (préfiltrage).

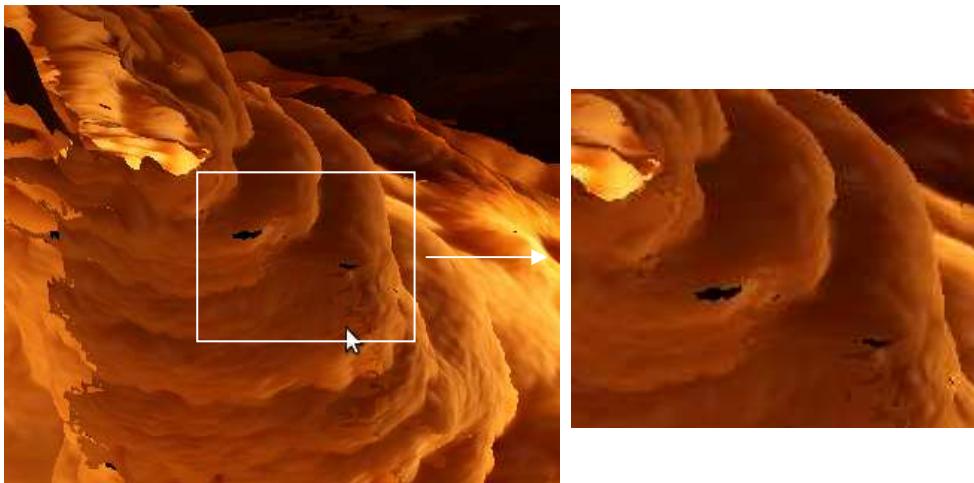


Figure 36 : Stalagmite. Défaut du résultat « trous ».

5.4. Eglise de St- Laurent (Lattes)

Caractéristique de l'objet : manque de texture (vitres, une partie du mur d'un bâtiment), répartition des points de vue n'est pas assez dense, variations importants de la surface au niveau de l'objet, nuages de points bruités (murs couvert par la végétation).

Le modèle issu de la fusion « propre », par contre on peut retrouver les parties manquantes.

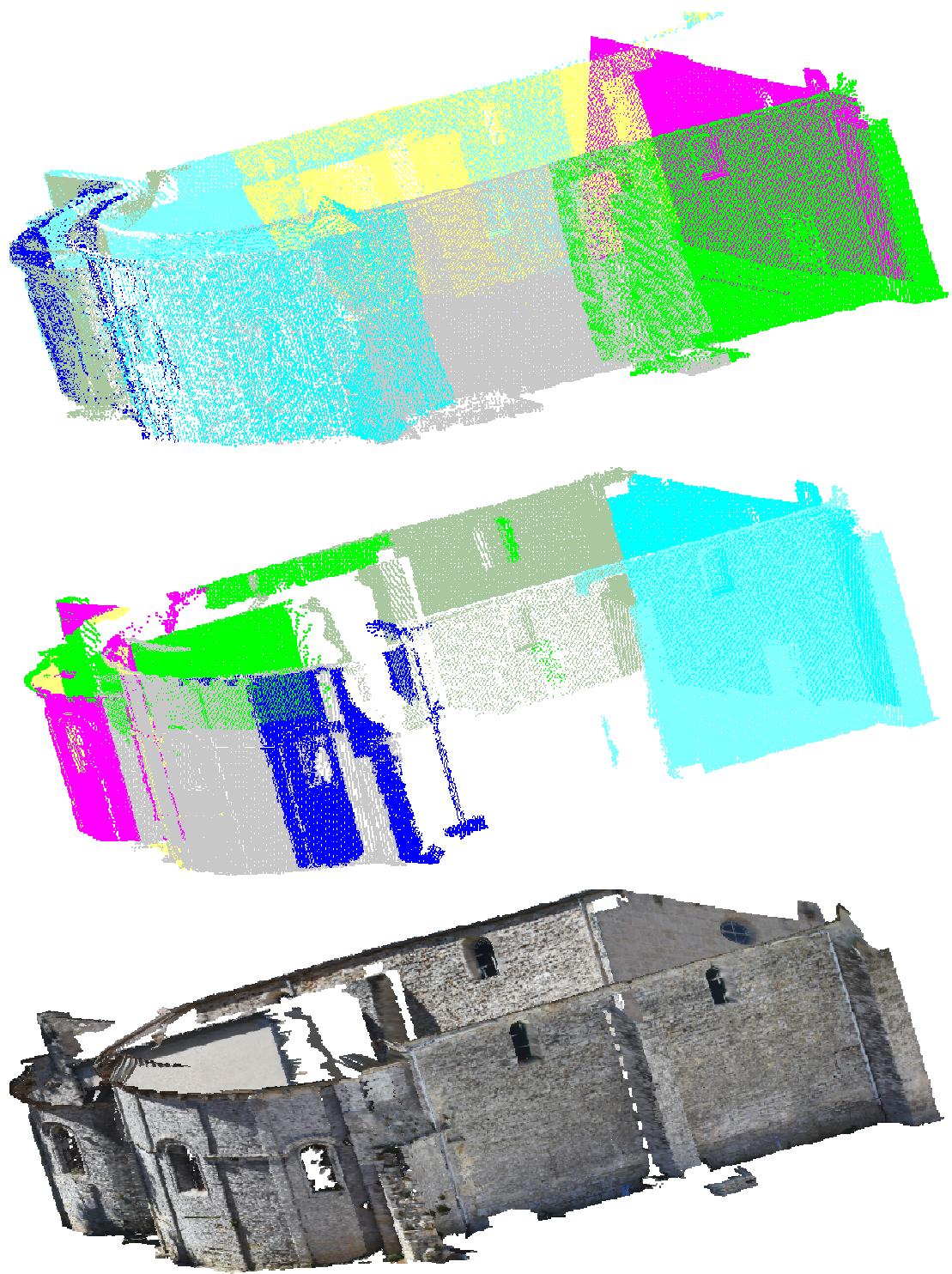


Figure 37 : Eglise de St-Laurent de Lattes. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturé

5.5. Aiguille du Midi (Massif du Mont-Blanc)

Caractéristique de l'objet : manque de texture (neige), répartition des points de vue n'est pas assez dense (l'existence des parties de grande angle d'incidence), variations importants de la surface dans les plusieurs plans

Le modèle obtenu plus « propre ». Par contre, beaucoup de parties manquantes.

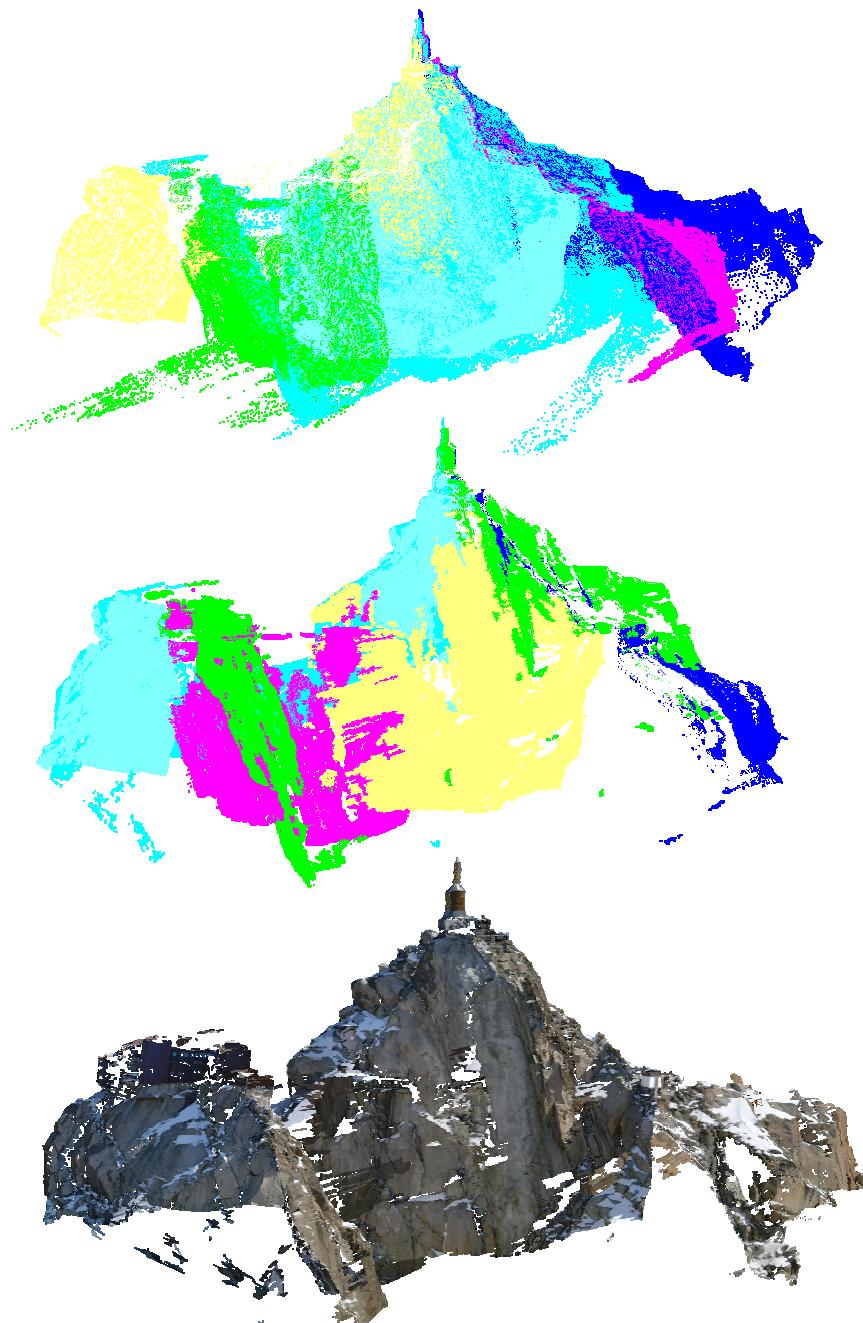


Figure 38 : Aiguille de Midi en Alpes. De haut en bas : modèle avant la fusion, modèle après la fusion, modèle après la fusion texturé

5.6. Bilan

Pour tester la complexité et l'efficacité de l'algorithme plusieurs objets avec des caractéristiques différentes ont été traités. Ces objets diffèrent selon le gabarit, la texture, la complexité de surface et la configuration de point de vue. En analysant les résultats obtenus, on peut noter plusieurs avantages de l'algorithme tels que: l'efficacité dans la majorité de cas, la suppression des erreurs d'où l'obtention d'un modèle « propre ». De plus, il ne nécessite pas un long temps de calcul. Par contre, l'inconvénient est l'exigence d'une « bonne » répartition des points de vue qui assure l'incidence fiable pour chaque point de l'objet.

Le tableau ci-dessous présente le résultat du calcul de la redondance avant et après la fusion.

	Pilier	Nubienne	Stalagmite	Eglise de St-Laurent	Aiguille du Midi
Redondance de données avant la fusion [%]	42%	33%	38%	36%	46%
Redondance de données après la fusion [%]	0,75%	0,18%	0,89%	0,19%	0,35%

Figure 39 : Tableau contenant les valeurs de redondance avant et après la fusion.

La redondance qui n'a pas été supprimé se centralise vers les raccords entre les cartes fusionnées (Figure 40). Cela peut être issu des passages entre les différentes géométries de chaque carte.

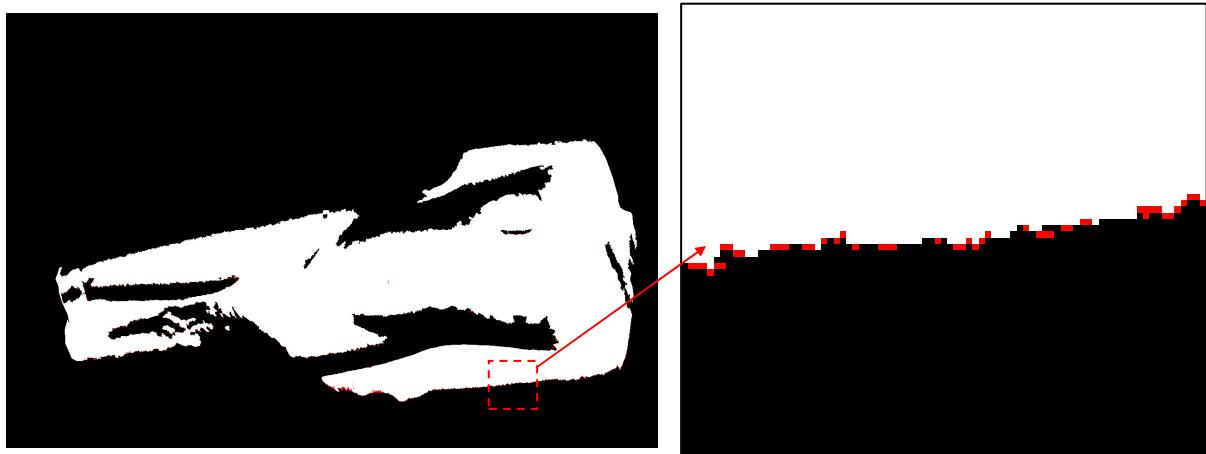


Figure 40 : Redondance aux raccords entre les cartes fusionnées (pixels rouge).

CONCLUSION

Le travail présent dans ce document propose une méthode de fusion de cartes de profondeur en photogrammétrie architecturale. Nous suivons la stratégie qui consiste à conserver l'information la plus fiable de chaque carte. Pour cela, nous appliquons plusieurs critères de fiabilité : angle d'incidence, corrélation, critère d'étirement de triangle. Nous proposons deux approches de fusion. La différence entre ces approches est la manière avec laquelle la visibilité a été formalisée. La première approche a été basée sur le calcul de la distance euclidienne entre deux points. Par ailleurs, le problème de visibilité posé de cette façon n'a pas été efficace pour tous les types d'objet. En conséquence, nous étions amenés à mettre en place un processus qui permet de manière rigoureuse la gestion du problème d'occultation. Pour cela, nous avons élaboré d'abord, une étape d'unification (à un moment donné) de la géométrie et ensuite nous introduisons le Z-buffer. Cette solution a apporté une amélioration significative quant au résultat de la fusion. Pour obtenir un « masque de définition » assez régulier, deux filtres morphologiques ont été appliqués. D'abord, la dilatation géodésique qui permet de boucher des petits « trous », puis l'érosion géodésique pour supprimer des petits artefacts. Le calcul de la fusion est précédé par le préfiltrage qui a pour effet la suppression de l'information la plus faible.

Les modèles 3D issus de la fusion sont satisfaisants. Le traitement n'est pas lourd en termes de calculs. Par contre, la méthode exige des prises de vue bien réparties est assez denses qui permettent d'assurer les incidences faibles pour chaque point de l'objet. Si tel n'est pas le cas, les valeurs d'incidences peuvent être élevées. Par conséquence, nous pouvons rencontrer des « trous » dans le modèle. Un autre problème peut provenir du filtrage morphologique. Dans cette étape, les pixels de plus faible information peuvent être groupés en deux ou trois donc par la suite rétablis dans le processus de filtrage.

Néanmoins, il reste encore quelques problèmes à résoudre. Les nuages concernés par la fusion n'ont pas le même échantillonnage donc nous pouvons retrouver des discontinuités entre les cartes fusionnés. Une approche de type « blending » permettant de « gommer » les discontinuités aux raccords entre les cartes peut être envisagé comme la prochaine étape de la fusion.

BIBLIOGRAPHIE

- [1] CURLESS B., LEVOY M., *A volumetric method for building complex models from range images*, SIGGRAPH '96: Proceeding of the 23rd annual conference on Computer Graphics and interactive techniques, New York, NY, USA, 1996, p.303-312
- [2] HILTON A., STODDART A., ILLINGWORTH J., WINDEATT T., *Reliable Surface Reconstruction from Multiple Range Images*, ECCV '96: Proceeding of the 4th European conference on Computer Vision-Volume, London, UK, 1996, p.117-126
- [3] SOUCY M., LAURENDEAU D., *A general Surface Approach to the Integration of a Set of Range Views*, IEEE, 1995, p.344-358
- [4] TURK G., LEVOY M., *Zippered polygon meshes from range images*, SIGGRAPH '94 : Proceeding of the 21st annual conference on Computer Graphics and interactive techniques, New York, NY, USA, 1994, p.311-318
- [5] PITO R., *Mesh integration based on co-measurements*, IEEE, 1996

[6] GARGALLO P., STURM P., *Bayesian 3D Modeling from Images using Multiple Depth Maps*, IEEE, San Diego, Californie, 2005, p.885-891

ITK - Segmentation & Registration Toolkit. <http://www.itk.org/>

Qt 4.5: Qt Reference Documentation. <http://doc.trolltech.com/4.5/index.html>

OpenGL SDK . <http://www.opengl.org/sdk/docs/man/>

Annexe 1 « Formule d'image »

Formule d'image

« Formule d'image » est une relation entre un point terrain et sa projection dans l'image.

Soit un point en coordonnées terrain $(x \ y \ z)^T$.

Soit la $k^{ième}$ pose de caméra définie par les paramètres O_k et R_k .

Les coordonnées d'un point lié à la caméra :

$$\begin{pmatrix} x_k' \\ y_k' \\ z_k' \end{pmatrix} = R_k * \begin{pmatrix} x \\ y \\ z \end{pmatrix} - O_k$$

La projection « canonique » reliant un point et sa direction de rayon :

$$\pi \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{x'}{z'} \\ \frac{y'}{z'} \\ 1 \end{pmatrix} = \begin{pmatrix} x' & y' \\ z' & z' \end{pmatrix} = (u \ v)$$

La relation entre une direction dans le repère camera et la position d'image :

$$(i \ j) = \mathfrak{I}(u \ v) = D(F * ((u \ v) + PP))$$

Donc la « formule d'image » :

$$(i \ j) = \mathfrak{I} \left(\pi \left(R_k * \begin{pmatrix} x \\ y \\ z \end{pmatrix} - O_k \right) \right)$$

Annexe 2 Relèvement dans l'espace

Trois points sont connus simultanément en coordonnées image I_1, I_2, I_3 et en coordonnées terrain T_1, T_2, T_3 . Trouver R (matrice de rotation) et O (position du sommet de la caméra) tels que $I_l = \mathfrak{I}(\pi(R^*(T_l - O)))$ où \mathfrak{I} (paramètres intrinsèques) de la caméra sont connus.

Il suffit de connaître les coordonnées caméras C_1, C_2, C_3 des trois points terrain.

On note : $C_{1,2} = \overrightarrow{C_1 C_2}$

$$\text{On a : } R(T_l - O) = C_l, \text{ donc } R \begin{pmatrix} T_{1,2} \\ T_{1,3} \\ T_{1,2} \wedge T_{1,3} \end{pmatrix} = \begin{pmatrix} C_{1,2} \\ C_{1,3} \\ C_{1,2} \wedge C_{1,3} \end{pmatrix}$$

On note \vec{u}_l la direction de rayon en coordonnées caméras donc :

$$\vec{u}_l = \mathfrak{I}^{-1}(I_l) = \begin{pmatrix} \frac{D_x^{-1}(I_l) - PP_x}{F} \\ \frac{D_y^{-1}(I_l) - PP_y}{F} \\ 1 \end{pmatrix}$$

Il suffit donc connaître la distance des points au centre de projection puisque $\overrightarrow{C_l} = d_l \vec{u}_l$ avec $d_l > 0$ (point de bon côté de la caméra).

Translation et rotation conservent les distances donc : $\|d_2 \vec{u}_2 - d_1 \vec{u}_1\|^2 = \|\overrightarrow{C_1 C_2}\|^2 = \|\overrightarrow{T_1 T_2}\|^2 = D_{1,2}$,

$$\|d_3 \vec{u}_3 - d_1 \vec{u}_1\|^2 = D_{1,3} \text{ et } \|d_3 \vec{u}_3 - d_2 \vec{u}_2\|^2 = D_{3,2}$$

$$\text{D'où : } \|d_2 \vec{u}_2 - d_1 \vec{u}_1\|^2 = \frac{D_{1,2}}{D_{1,3}} \|d_3 \vec{u}_3 - d_1 \vec{u}_1\|^2 = \rho_1 \|d_3 \vec{u}_3 - d_1 \vec{u}_1\|^2 \quad \text{donc :} \quad \left\| \frac{d_2 \vec{u}_2 - \vec{u}_1}{d_1} \right\|^2 = \rho_1 \left\| \frac{d_3 \vec{u}_3 - \vec{u}_1}{d_1} \right\|^2$$

$$\text{et } \left\| d_2 \vec{u}_2 - d_1 \vec{u}_1 \right\|^2 = \frac{D_{1,2}}{D_{3,2}} \left\| d_3 \vec{u}_3 - d_2 \vec{u}_2 \right\|^2 = \rho_2 \left\| d_3 \vec{u}_3 - d_2 \vec{u}_2 \right\|^2 \quad \text{donc : } \left\| \frac{d_2}{d_1} \vec{u}_2 - \vec{u}_1 \right\|^2 = \rho_2 \left\| \frac{d_3}{d_1} \vec{u}_3 - \frac{d_2}{d_1} \vec{u}_2 \right\|^2$$

$$\text{On pose: } \lambda_2 = \frac{d_2}{d_1}, \quad \lambda_3 = \frac{d_3}{d_1} \quad \text{donc : } \left\| \lambda_2 \vec{u}_2 - \vec{u}_1 \right\|^2 = \rho_1 \left\| \lambda_3 \vec{u}_3 - \vec{u}_1 \right\|^2 \quad \left\| \lambda_2 \vec{u}_2 - \vec{u}_1 \right\|^2 = \rho_2 \left\| \lambda_3 \vec{u}_3 - \lambda_2 \vec{u}_2 \right\|^2$$

$$\text{Si on connaît } \lambda_2 \text{ on pourra calculer } d_2 \text{ en utilisant } \left\| d_2 \vec{u}_2 - d_1 \vec{u}_1 \right\|^2 = D_{1,2} \text{ et} \\ d_2^2 \left\| \lambda_2 \vec{u}_2 - \vec{u}_1 \right\|^2 = D_{1,2}.$$

$$\text{On a : } \left\| \lambda_2 \vec{u}_2 - \vec{u}_1 \right\|^2 = \rho_1 \left\| \lambda_3 \vec{u}_3 - \vec{u}_1 \right\|^2 \quad \text{et } (A\lambda_2^2 + B\lambda_2 + C) = (A'\lambda_3^2 + B'\lambda_3 + C')$$

$$\text{Où : } A\lambda_2^2 + B\lambda_2 + P(\lambda_3) = 0 \quad \text{donc : } \lambda_2 = \frac{-B \pm \sqrt{B^2 - 4AP(\lambda_3)}}{2A} = D + \sqrt{Q(\lambda_3)}$$

$$\text{Ensuite : } \left\| \lambda_2 \vec{u}_2 - \vec{u}_1 \right\|^2 = \rho_2 \left\| \lambda_3 \vec{u}_3 - \lambda_2 \vec{u}_2 \right\|^2 \quad \text{et } A Q(\lambda_3) + B \sqrt{Q(\lambda_3)} + C = A' Q(\lambda_3) + \sqrt{Q(\lambda_3)} R_1(\lambda_3) + R_2(\lambda_3)$$

$$\text{Donc : } S_1 \sqrt{Q(\lambda_3)} = S_2(\lambda_3) \quad S_1^2 Q(\lambda_3) = S_2^2(\lambda_3)$$

C'est une équation de 4^{ième} degré en λ_3

Nombre de solutions entre 2 et 8 mais moins de la moitié ont un sens physique.

En général, une information supplémentaire (4^{ième} point) en nécessaire pour enlever l'ambiguïté.

Annexe 3 Orientation relative entre deux images

On a deux poses 1 et 2. On note T un point terrain et P_1^C, P_2^C ses coordonnées dans les repères liés aux caméras.

L'orientation relative de 2 par rapport à 1, est l'orientation de 2 dans le repère lié à la pose 1.

Soit $O_{2,1}, R_{2,1}$, le couple définissant cette orientation, par définition: $P_2^C = R_{2,1} * (P_1^C - O_{2,1})$.

Par définition on a aussi $P_2^C = R_2 * (T - O_2)$ et $P_1^C = R_1 * (T - O_1)$.

$$\text{Donc : } P_2^C = R_{2,1} * ((R_1 * (T - O_1)) - O_{2,1}) = R_{2,1}R_1(T - O_{2,1} - R_1^{-1}O_1)$$

$$\text{Où : } R_2 = R_{2,1}R_1, \text{ et } O_2 = O_{2,1} + R_1^{-1}O_1$$

Avec au minimum, cinq points de liaisons on peut calculer l'orientation relative d'une image par rapport à l'autre à un facteur d'échelle près.

Annexe 4 Géométrie épipolaire

Définition 1 : Une géométrie épipolaire d'un couple d'images est un paire de rotations R_1^E, R_2^E qui appliquées aux caméras permettent de rendre parallèle les plans d'image et de mettre la base dans le plan d'image.

On note P_1^E, P_2^E les coordonnées des points dans les orientations épipolaires.

$$P_1^E = P_2^E + \begin{pmatrix} B \\ 0 \\ 0 \end{pmatrix} = P_2^E + \vec{X}$$

On a les relations : $P_1^C = R_1^E(P_1^E)$, et $P_2^C = R_2^E(P_2^E)$

On a donc : $P_2^C = R_2^E(P_2^E) = R_2^E(P_1^E + \vec{X}) = R_2^E({}^t R_1^E P_1^E + \vec{X})$

Soit $(x_1 \ y_1 \ z_1)_E^T$ et $(x_2 \ y_2 \ z_2)_E^T$ dans le repère épipolaire. Les faisceaux perspectifs

$(\frac{x_1}{z_1} \ \frac{y_1}{z_1} \ 1)_E^T$ et $(\frac{x_2}{z_2} \ \frac{y_2}{z_2} \ 1)_E^T$ sont homologues si les ordonnées sont égales :

$$\frac{y_1}{z_1} = \frac{y_2}{z_2} \Leftrightarrow y_1 z_2 = y_2 z_1 \Leftrightarrow 0 = y_1 z_2 - y_2 z_1 = (x_1 \ y_1 \ z_1) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

$$\text{On pose : } \mathcal{E} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \text{ on a } P_1^E \approx P_2^E \Leftrightarrow {}^t P_1^E \mathcal{E} P_2^E = 0 \text{ or } P_1^E = {}^t R_1^E P_1^C$$

On peut donc écrire : ${}^t P_1^C R_1^E \mathcal{E} {}^t R_2^E P_2^C = 0$

Définition 2 : Il existe une matrice E , dite « **matrice essentielle** » tels que $P_1^C \approx P_2^C \Leftrightarrow {}^t P_1^C E P_2^C = 0$. Elle est liée aux repères épipolaires par $E = R_1^E \mathcal{E} {}^t R_2^E$.

Pour calculer une matrice essentielle E , il faut au minimum 8 points homologues. Chaque couple de points homologues donne une relation linéaire :

$${}^t P_1^C \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} P_2^C = 0, \quad \text{où } E = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

On ajoute une contrainte pour lever

l'ambiguïté d'échelle, par exemple $i = 1$.

$$\text{On note : } \varepsilon = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \theta_x \Pi.$$

L'écriture $E = R_1^E \theta_x \Pi {}^t R_2^E = (R_1^E \theta_x) \Pi ({}^t R_2^E)$ correspond à une SVD (décomposition en valeur singulière). Donc si on a calculé E , en effectuant sa SVD on peut estimer les rotations épipolaires $E = UDV$.

$$\text{On a : } R_1^E \theta_x = \begin{pmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & S_z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & C & -S \\ 0 & S & C \end{pmatrix} U, \quad \text{où } S_x = \pm 1, S_y \dots$$

On oublie la rotation qui est sans importance (rotation autour de la base). On veut que les matrices soient directes: $S_x S_y S_z = 1$

Il y a 4 solutions algébriquement valides. Seule une a un sens physique (celle où les deux caméras voient vers l'avant).

Annexe 5 Erosion et dilatation géodésique

Distance géodésique, conditionnelle à X : d_X

- si X est fermé, il existe un arc géodésique pour toute paire de points de X
- unicité si X est simplement connexe
- X convexe $\Leftrightarrow d_X = d$

Boule géodésique : $B_{X,r}(x) = \{y \in X / d_X(x, y) \leq r\}$

Remarque : $B_X(x, r) \subseteq B(x, r)$

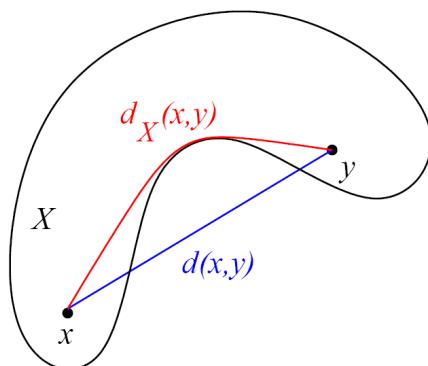


Figure 41 : Distance géodésique.

Dilatation géodésique :

$$D_X(Y, B_r) = \{x \in \Re^n / B_X(x, r) \cap Y \neq \emptyset\} = \{x \in \Re^n / d_X(x, Y) \leq r\}$$

Erosion géodésique :

$$E_X(Y, B_r) = \{x \in \Re^n / B_X(x, r) \subseteq Y\} = X \setminus D_X(X \setminus Y, B_r)$$