

Vers de nouvelles méthodes 3D de capture du mouvement : la photogrammétrie p   et la lasergramm  trie comme techniques d'acquisition

Auteur : Willems, Henri

Promoteur(s) : Billen, Roland

Facult   : Facult   des Sciences

Dipl  me : Master en sciences g  ographiques, orientation g  omatique et g  om  trologie,    finalit   sp  cialis  e

Ann  e acad  mique : 2017-2018

URI/URL : <http://hdl.handle.net/2268.2/4878>

Avertissement    l'attention des usagers :

Tous les documents plac  s en acc  s ouvert sur le site le site MatheO sont prot  g  s par le droit d'auteur. Conform  ment aux principes   nonc  s par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, t  l  charger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte int  gral de ces documents, les diss  quer pour les indexer, s'en servir de donn  es pour un logiciel, ou s'en servir    toute autre fin l  gale (ou pr  vue par la r  glementation relative au droit d'auteur). Toute utilisation du document    des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage    respecter les droits moraux de l'auteur, principalement le droit    l'int  grit   de l'oeuvre et le droit de paternit   et ce dans toute utilisation que l'utilisateur entreprend. Ainsi,    titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son int  gralit  , l'utilisateur citera de mani  re compl  te les sources telles que mentionn  es ci-dessus. Toute utilisation non explicitement autoris  e ci-avant (telle que par exemple, la modification du document ou son r  sum  ) n  cessite l'autorisation pr  alable et expresse des auteurs ou de leurs ayants droit.

UNIVERSITÉ DE LIÈGE
Faculté des Sciences
Département de Géographie

Vers de nouvelles méthodes 3D de capture du mouvement

La photogrammétrie et la lasergrammétrie comme techniques d'acquisition

*Mémoire présenté par **Henri WILLEMS**
pour l'obtention du titre de :*

***Master en sciences géographiques,
orientation géomatique et géométrologie
à finalité spécialisée***

*Année académique : **2017 – 2018***

Date de défense : 28 juin 2018

Promoteur : Prof. R. BILLEN

Lecteurs : Prof. R. WARNANT & Prof. S. SCHMITZ

UNIVERSITÉ DE LIÈGE
Faculté des Sciences
Département de Géographie

Vers de nouvelles méthodes 3D de capture du mouvement

La photogrammétrie et la lasergrammétrie comme techniques d'acquisition

*Mémoire présenté par **Henri WILLEMS**
pour l'obtention du titre de :*

***Master en sciences géographiques,
orientation géomatique et géométrologie
à finalité spécialisée***

*Année académique : **2017 – 2018***

Date de défense : 28 juin 2018

Promoteur : Prof. R. BILLEN

Lecteurs : Prof. R. WARNANT & Prof. S. SCHMITZ

REMERCIEMENTS

En prélude de ce travail de fin d'études, j'aimerais témoigner ma reconnaissance à toutes les personnes qui ont contribué à son aboutissement.

Premièrement, mes remerciements vont à mon promoteur Roland BILLEN pour m'avoir accompagné, conseillé et aiguillé pendant ces mois de recherche.

Ensuite, ma gratitude va à l'ensemble de l'Unité de Géomatique, pour la bonne humeur et l'atmosphère de travail agréable qui y règne. Au sein de celle-ci, j'aimerais personnellement remercier les assistants Florent POUX et Benoit JONLET. Ces derniers ont toujours trouvé des réponses à mes questions et cela peu importe les problèmes auxquels j'ai fait face.

Finalement, j'adresse des remerciements chaleureux à l'ensemble de mes proches, sans qui je ne serais pas là où j'en suis aujourd'hui. Ainsi, je remercie ma famille, et en particulier mes parents, pour m'avoir appuyé dans mes choix et sans qui mes études n'auraient pas été possibles. Merci également à mes amis, en particulier mes colocataires du Curtius Gang, grâce à qui j'ai toujours eu le moral. Et enfin, merci à Oksana, parce que je n'ai pas le choix...

TABLE DES MATIERES

1	INTRODUCTION	4
1.1	<i>Définition et domaines d'application</i>	5
1.2	<i>Contexte de recherche</i>	6
2	ÉTAT DE L'ART	7
2.1	<i>Les systèmes d'acquisition non-optiques.....</i>	7
2.1.1	<i>Les systèmes mécaniques.....</i>	7
2.1.2	<i>Les systèmes acoustiques.....</i>	7
2.1.3	<i>Les systèmes magnétiques.....</i>	7
2.2	<i>Les systèmes d'acquisition optiques</i>	8
2.2.1	<i>Les systèmes basés sur des capteurs RGB-D</i>	8
2.2.2	<i>Les systèmes mono-capteur.....</i>	9
2.2.3	<i>Les systèmes à deux capteurs</i>	10
2.2.4	<i>Les systèmes multi-capteurs.....</i>	11
3	ORIENTATION ET HYPOTHESES DE RECHERCHE	14
3.1	<i>Contraintes</i>	14
3.2	<i>Ressources disponibles</i>	14
3.3	<i>Orientation de la recherche.....</i>	15
3.4	<i>Hypothèses de recherche</i>	19
4	DÉVELOPPEMENT THÉORIQUE	20
4.1	<i>Objet en mouvement</i>	20
4.2	<i>Sans marqueurs placés sur l'objet.....</i>	21
4.2.1	<i>Acquisition du mouvement</i>	21
4.2.2	<i>Traitement des données</i>	27
4.2.3	<i>Rendu 3D</i>	27
4.3	<i>Avec marqueurs placés sur l'objet.....</i>	29
4.3.1	<i>Acquisition du mouvement</i>	30
4.3.2	<i>Traitement des données</i>	34
4.3.3	<i>Rendu 3D</i>	34
4.5	<i>Un système de capture de mouvement basé sur la lasergrammétrie</i>	35
4.5.1	<i>Acquisition du mouvement</i>	35
4.5.2	<i>Extraction du mouvement</i>	37
5	APPLICATIONS.....	38
5.1	<i>Objet en mouvement</i>	38
5.2	<i>Sans marqueurs placés sur l'objet.....</i>	39
5.2.1	<i>Acquisition du mouvement</i>	39
5.2.2	<i>Traitement des données</i>	40
5.2.3	<i>Rendu 3D</i>	44
5.3	<i>Avec marqueurs placés sur l'objet.....</i>	45
5.3.1	<i>Acquisition du mouvement</i>	45
5.3.2	<i>Traitement des données</i>	47
5.3.3	<i>Rendu 3D</i>	48
5.4	<i>Un système de capture de mouvement basé sur la lasergrammétrie</i>	49
5.4.1	<i>Acquisition du mouvement</i>	49
5.4.2	<i>Extraction du mouvement</i>	49
6	DISCUSSION	56
7	CONCLUSION.....	59
8	BIBLIOGRAPHIE	60
9	ANNEXES.....	63

1 INTRODUCTION

Depuis que les technologies de *capture de mouvement* ont vu le jour dans les années 90, ces dernières n'ont eu de cesse d'évoluer. Elles se sont, au fur et à mesure des années, étendue à des applications toujours plus diverses, le tout en utilisant des techniques de plus en plus variées, menant à des résultats 3D de plus en plus précis et réalistes.

Parmi les techniques utilisées pour développer les méthodes de capture de mouvement, beaucoup ont déjà fait leurs preuves et atteint leurs limites dans leurs applications respectives. Néanmoins, la photogrammétrie et la lasergrammétrie, deux techniques de reconstitution 3D ayant déjà fait leurs preuves en matière de modélisation d'objet fixe, ne sont qu'à leur début en ce qui concerne la capture et la modélisation 3D du mouvement.

Ce travail de fin d'études a pour but d'établir les perspectives de ces deux techniques en matière de capture de mouvement. Pour atteindre son objectif, ce dernier se structure en six étapes principales.

Premièrement, une présentation de la capture du mouvement, de ses domaines d'application et du contexte de recherche sera réalisée dans cette introduction afin de préciser le sujet abordé et justifier les raisons d'un travail de fin d'étude portant sur ce sujet.

Deuxièmement, un état de l'art technique sera établi. Ce point présentera le détail des systèmes d'acquisition développés à ce jour en matière de capture de mouvement, ainsi que leurs avantages et leurs inconvénients respectifs.

Troisièmement, le détail des contraintes et ressources disponibles pour la réalisation de ce travail de fin d'études sera dressé. La confrontation de ces contraintes et ressources avec les résultats de l'état de l'art et d'une série de tests effectués par nos soins nous permettra d'orienter le travail et établir nos hypothèses de recherche.

Quatrièmement, un développement théorique de nos connaissances en matière de photogrammétrie et de lasergrammétrie sera réalisé afin de proposer une méthodologie d'acquisition, de traitement et de rendu 3D nous permettant de valider ou invalider chacune de nos hypothèses de recherche.

Cinquièmement, la méthodologie développée théoriquement sera mise en pratique et le détail des résultats sera présenté.

Dernièrement, une synthèse des résultats sera produite et confrontée à nos hypothèses de recherches. Cette confrontation nous permettra de discuter les résultats obtenus et de tirer les conclusions de ce travail de fin d'études.

1.1 Définition et domaines d'application

La *capture du mouvement* (*motion capture* ou *mocap* dans sa version abrégée) regroupe l'ensemble des techniques visant à acquérir les composantes tridimensionnelles (positions, vitesses, rotations) du mouvement d'un modèle ou d'un ensemble de points représentant un modèle. Dans un logiciel de rendu 3D, cette capture permet la conception d'animation dans le but de réaliser des traitements et des analyses ultérieures variant en fonction des applications (Knyaz, 2015).

Les domaines d'applications où les technologies de capture du mouvement sont les plus développées sont les industries du film et du jeu vidéo. Historiquement, c'est au sein de ces industries du divertissement que les besoins d'acquisition de mouvements d'acteurs se sont fait premièrement ressentir, permettant ainsi de donner une haute impression de réalité à différents effets spéciaux (Knyaz, 2015). Répondant à une concurrence de plus en plus grandissante, ces industries restent, à ce jour, les plus grosses sources d'avancées technologiques en la matière.

Néanmoins, on dénombre également de nouvelles sources de développements technologiques dans des milieux de plus en plus variés. Ainsi, la médecine développe maintenant des techniques précises d'analyse du mouvement de l'être humain permettant de déceler des anomalies et de proposer des modes de réhabilitation adéquats aux patients. Les exemples de *motion capture* appliqués à la biomécanique et à l'ergonomie sont multiples dans la littérature scientifique (Bonenfant et al., 2017 ; Kim, 2013 ; Nguyen et al., 2011). Dans le monde du sport de haut niveau, des systèmes permettent aussi d'obtenir des informations précises sur les mouvements des sportifs pendant l'effort. Ces mesures permettent l'amélioration continue des techniques et l'atteinte de meilleurs résultats. Par exemple, les champions de golf ont accès à des systèmes de capture de mouvement les aidant à corriger leurs techniques de swing (Urtasun, 2005). En outre, dans les industries technologiques de pointe, le paramétrage des mouvements de robots sollicite de plus en plus les avancées relatives à la *motion capture*, cela afin d'obtenir des mouvements les plus homogènes et les plus naturels possible (Stanton, 2012). Enfin, la capture de mouvement est également souvent utilisée en recherche scientifique. Des exemples typiques sont l'analyse de processus dynamiques comme le battement des ailes d'insectes, l'analyse du mouvement d'un véhicule ou d'un projectile et la modélisation d'objets 3D (Knyaz, 2015 ; Moeslund, 2006).

1.2 Contexte de recherche

Ce travail de fin d'études est né de la collaboration de la Faculté de Philosophie et Lettres et de l'Unité de Géomatique de l'Université de Liège.

Dans le cadre d'un projet de conservation et de documentation d'œuvres d'art contemporain, l'objectif principal du travail de fin d'études est de développer une méthode d'acquisition, de traitement et de rendu 3D d'objet en mouvement. Cet objectif est justifié par des besoins de recherche exprimés tant par la faculté de Philosophie et Lettres que par l'Unité de Géomatique.

Tout d'abord, nous exposerons les besoins de recherche de la faculté de Philosophie et Lettres, identifiés lors d'un entretien avec Manon D'haenens, chercheuse au sein de l'Unité d'Archéométrie. Ainsi, développer une méthode d'acquisition, de traitement et de rendu 3D d'objet en mouvement serait utile à la rédaction d'une partie du *mode d'emploi* de l'œuvre d'art cinétique « 24 boules sur 24 sommets inclinés ». Cette œuvre, réalisée en 1965 par l'artiste contemporain belge Pol Bury, se compose, comme son nom l'indique, de 24 boules de bois en mouvement continu sur un socle constitué de 24 sommets. Le mouvement des boules est assuré par des moteurs électriques cachés dans le socle. La Figure 1 illustre cette œuvre. On peut définir le *mode d'emploi* d'une œuvre d'art comme un dossier écrit contenant l'ensemble de ses caractéristiques. Il fait le lien entre la technique et l'effet artistique. La rédaction du *mode d'emploi* est indispensable pour permettre la conservation et la transmission à la postérité. Il permet de garder une trace des connaissances recueillies avant tout projet de restauration ou de déplacement. On retrouve dans le *mode d'emploi* des informations relatives à l'identification des matériaux, au détail du fonctionnement et à l'effet de l'œuvre.



Figure 1 : Photo de l'œuvre « 24 boules sur 24 sommets inclinés » de Pol Bury
(Source : Bury Catalogue 2017)

L'effet caractérise toutes les informations artistiques pouvant émaner du mouvement d'une œuvre d'art cinétique. Après discussion, il a été décidé qu'une animation 3D était le meilleur moyen de visualiser et d'analyser cet effet de l'œuvre. Cette animation conserverait ainsi la mécanique du mouvement de l'objet et permettrait de mesurer ses composantes tridimensionnelles (positions, vitesses, axes principaux, durée, etc.). C'est donc dans le but de réaliser une animation 3D qu'une méthode de capture du mouvement doit être développée.

Par la suite, des besoins de recherche émanent également de l'Unité de Géomatique. En plus d'étendre son domaine d'expertise à la conservation d'œuvre d'art cinétique contemporaine, ce projet permettrait de développer une méthode d'acquisition, de traitement et de rendu 3D d'objet en mouvement. Pour satisfaire nos attentes, cette méthode doit fournir des résultats à un niveau de précision millimétrique. Elle doit également être adaptable à d'autres applications et ne pas se limiter à la seule œuvre cinétique de Pol Bury présentée ci-dessus. De plus, notre objectif est aussi de tester la combinaison des différents moyens d'acquisition présents à l'Unité de Géomatique dans un cas de capture du mouvement. Ce travail est également l'occasion de déterminer le niveau de précision qu'il est possible d'atteindre. Les moyens techniques présents à l'unité sont le tachéomètre, le scanner laser terrestre et l'appareil photo reflex. La liste et le détail technique des moyens d'acquisition seront présentés au point 2. La concrétisation de ces objectifs passera donc par la mise en place, l'analyse et la critique de méthodes d'acquisition d'objet en mouvement utilisant la photogrammétrie et la lasergrammétrie comme techniques d'acquisition.

2 ÉTAT DE L'ART

Afin de capturer les composantes tridimensionnelles du mouvement d'un modèle (objet ou personne), de nombreuses méthodes différentes ont vu le jour au cours des dernières décennies. Chacune de ces méthodes est composée de deux systèmes : un système d'acquisition et un système de traitement sur ordinateur. Le système d'acquisition est composé de capteurs dont la nature dépend de la méthode utilisée. Son rôle est d'extraire les données brutes matérielles du mouvement et de les transmettre au système de traitement qui se charge d'effectuer les calculs et de reconstituer le mouvement sur ordinateur. Les méthodes actuelles s'appuient sur des systèmes d'acquisition non-optiques (mécaniques, acoustiques ou magnétiques) et optiques.

2.1 Les systèmes d'acquisition non-optiques

2.1.1 Les systèmes mécaniques

Surtout utilisés pour des applications en biomécanique, les systèmes mécaniques de capture de mouvement utilisent des exosquelettes dont les articulations sont équipées de goniomètres électroniques.

Un *goniomètre* est un capteur fournissant des mesures d'angles relatifs à l'aide d'un compas interne basé sur un potentiomètre. La résistance du potentiomètre change avec la rotation des deux segments de corps qui lui sont connectés. En mesurant les variations de potentiel au sein d'un goniomètre traversé par un courant électrique d'intensité constante, on peut en déduire la résistance conformément à la loi d'*Ohm*. Cette résistance est donc proportionnelle à l'angle des articulations (Seo et al., 2017).

Ces systèmes mécaniques ont l'avantage d'être indépendants de toute forme de champs magnétiques ou de réflexion d'onde électromagnétique et sont relativement rapides à mettre en place. Malheureusement, ils sont dépendants de nombreux câbles servant à communiquer les données des capteurs au système de traitement (Knyaz 2015), empêchant toute perspective de capture de mouvement ample ou d'objets de petite taille.

2.1.2 Les systèmes acoustiques

Les systèmes acoustiques sont équipés d'un ensemble de récepteurs enregistrant les sons d'émetteurs disposés sur le modèle en mouvement. Les récepteurs placés autour de la scène enregistrent les moments d'acquisition des sons spécifiques à chacun des émetteurs. Connaissant la vitesse du son et la différence entre les moments d'émission et de réception, les distances émetteurs-récepteurs sont calculées et la position 3D de chaque émetteur est déterminée par triangulation des distances émetteurs-récepteurs (Seo et al., 2017).

Les systèmes acoustiques sont soumis à des inconvénients comme la restriction du mouvement imposée par les câbles des émetteurs, le nombre limité d'émetteurs pouvant être utilisés en même temps et la susceptibilité aux échos et aux sources de bruits extérieurs au système. Ces inconvénients limitent leur champs d'applications (Knyaz, 2015).

2.1.3 Les systèmes magnétiques

Les systèmes magnétiques utilisent comme marqueurs un ensemble d'aimants modélisant les points-clés du mouvement. Autour de la scène, un ensemble de récepteurs dérive les positions relatives des aimants par rapport à une antenne en fonction de l'orientation de leur champ magnétique respectif. (Yabukami et al., 2000 ; Bodenheimer et al., 1997). Ces systèmes sont économiques, précis et rapides pour des captures de mouvements simples, mais ont le désavantage d'être également limités par des câbles, ainsi que d'être soumis à de possibles interférences magnétiques avec des objets métalliques environnants (Knyaz, 2015).

2.2 Les systèmes d'acquisition optiques

Les systèmes optiques utilisent les données issues de capteurs d'image 2D ou 3D afin d'extraire les caractéristiques du mouvement d'un objet dans l'espace (Poppe, 2010). L'état des recherches dans le domaine permet de mettre en évidence plusieurs systèmes optiques. Ces derniers se différencient par le nombre et le type de capteurs utilisés, ainsi que par la présence ou non de marqueurs sur l'objet.

2.2.1 Les systèmes basés sur des capteurs RGB-D

Un capteur ou une caméra RGB-D (*Red-Green-Blue-Depth*) est un système de détection qui capture des images dans la longueur d'onde du visible et qui fournit en plus une information de profondeur pour chacun des pixels. L'information de profondeur est possible grâce à un émetteur et une caméra infrarouge présents au sein du capteur. L'émetteur recouvre la scène de rayons infrarouges qui seront réfléchis par l'ensemble des surfaces touchées avec plus ou moins d'intensité en fonction de la distance à la caméra. La caméra infrarouge mesure cette intensité de réflexion et la distance à l'objet en mouvement peut ainsi être déduite (Henry et al., 2012).

Sur base d'une ou plusieurs caméras, il est alors possible de déterminer les positions d'objets dans l'espace. En biomécanique, ces systèmes RGB-D permettent d'extraire des données de mouvements basées sur un squelette de marqueurs attaché à un corps en mouvement. Plus encore, il existe à ce jour des logiciels de capture de mouvement se servant de données RGB-D et fournissant des solutions pour reconstituer des scènes de mouvement complètes en 3D sans l'utilisation de marqueur (Seo et al, 2017). Ces algorithmes sont basés sur un modèle prédéfini par l'utilisateur. Ils peuvent ainsi traquer le mouvement d'un individu en faisant correspondre la surface du modèle avec l'information de profondeur extraite de la caméra infrarouge. Les données de mouvement peuvent ensuite être exportées sous le format d'échange de fichiers désiré (BVH, ASK, C3D ou autres). Ces logiciels fournissent également des algorithmes de post-traitement capables d'extraire le bruit de mesure et de filtrer le mouvement afin de le rendre plus homogène.

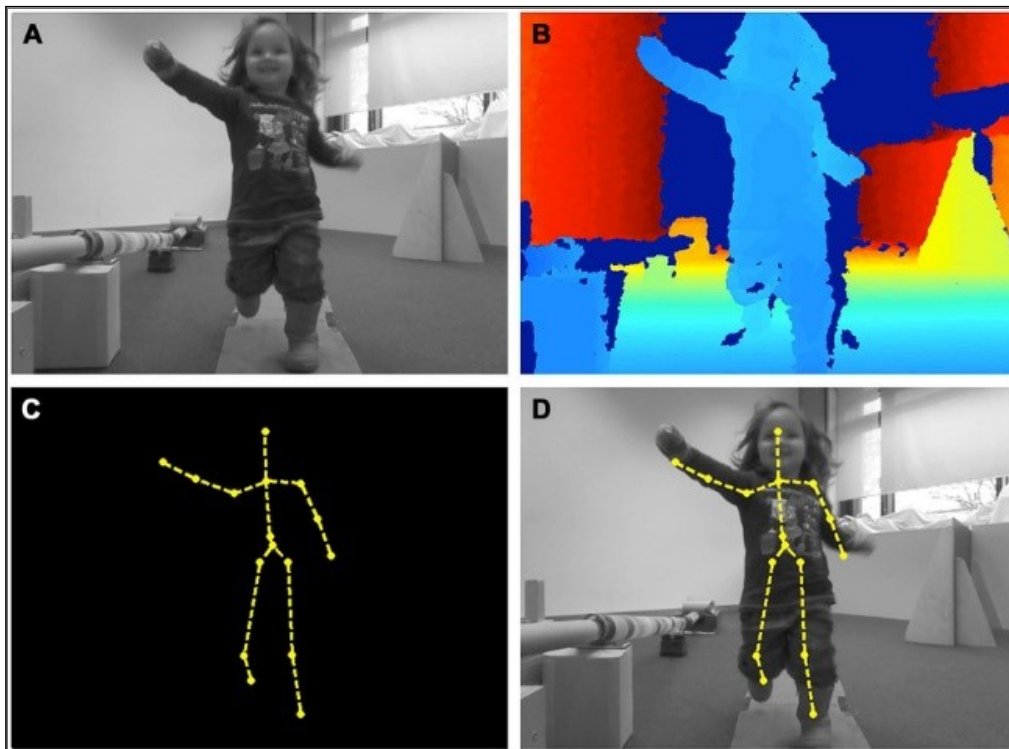


Figure 2 : Exemple d'image RGB-D liée à un modèle de corps prédéfini.
(Source : Hepach, 2015)

A ce jour, ces algorithmes sont aussi capables de gérer les occlusions se produisant lors de la capture du mouvement de corps humain par exemple (Hepach, 2015). Un exemple d'image RGB-D liée à un modèle de corps prédéfini est présenté à la Figure 2.

Les avantages des systèmes basés sur des caméras RGB-D sont nombreux. Parmi ces avantages, nous pouvons citer l'absence de marqueurs ou de capteurs attachés à l'objet en mouvement. Cette absence permet par la même occasion de s'affranchir de câbles pouvant gêner le mouvement. La technologie associée est également économique et facile à mettre en place pour la collecte de données.

Néanmoins, l'utilisation de ces systèmes présente certains inconvénients. Ainsi, la lumière infrarouge issue des capteurs limite l'usage de cette technologie à l'intérieur pour ne pas interférer avec le rayonnement solaire. De plus, la courte portée de pareil capteur (maximum 4 mètres) est une autre limitation majeure (Seo et al., 2017).

Initialement développée pour l'industrie du jeu vidéo avec la caméra KINECT de MICROSOFT, la technologie RGB-D est aujourd'hui utilisée dans d'autres applications de capture du mouvement comme le monitoring de maquette de barrage par exemple.

2.2.2 Les systèmes mono-capteur

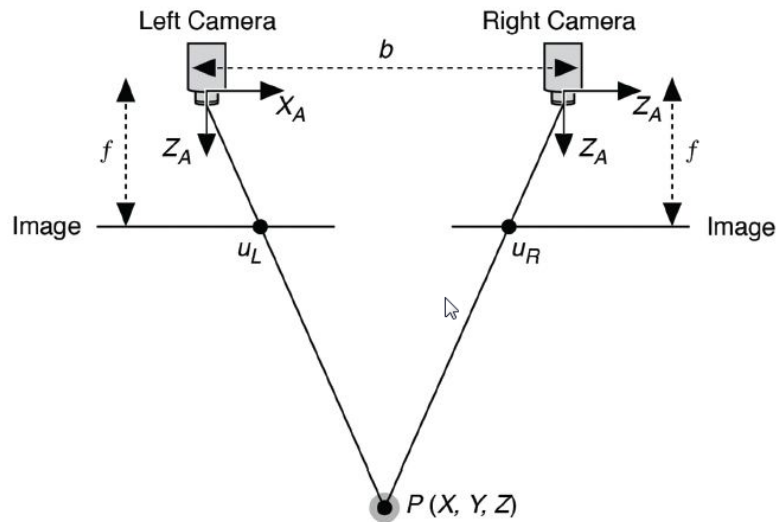
Les systèmes mono-capteur de capture du mouvement ont été développés avec l'avènement des techniques de vision par ordinateur. Ces systèmes ont pour objectif de générer des données en suivant les caractéristiques de surface d'un objet identifié dynamiquement au sein d'une vidéo. La détection des objets en mouvement et le traçage d'objets basé sur le mouvement ont des applications importantes dans des milieux comme la reconnaissance faciale, la surveillance du trafic et la sécurité automobile (Ren et al., 2018). La problématique du suivi d'objets basé sur le mouvement peut être divisée en deux parties : la détection d'objets en mouvement dans chaque image d'une vidéo et le groupement temporel des détections correspondant au même objet.

Pour ce qui est de la détection d'objets en mouvement dans chaque image d'une vidéo, trois étapes sont requises. La première est l'utilisation d'algorithme de soustraction d'arrière-plan basé sur des modèles de mélange gaussiens ou *Gaussian Mixture Model (GMM)*. Il s'agit de modèles statistiques utilisant une combinaison complexe de fonctions de densité. Ces modèles servent à estimer les paramètres de distribution (variance, moyenne, et amplitude) de plusieurs gaussiennes représentant les valeurs de pixels au sein d'une image. Ces paramètres sont choisis selon un critère de maximum de vraisemblance afin d'approximer le plus possible la distribution des pixels représentant l'objet en mouvement au sein d'une image. La deuxième étape est une suite d'opérations morphologiques sur les pixels extraits afin d'éliminer le bruit de mesure et constituer des groupes de pixels de même propriété radiométriques, appelés *blobs*. Enfin, la troisième étape consiste en la détection des *blobs* connectés et susceptibles de correspondre au même objet en mouvement sur base de leur information radiométrique. Cette détection porte le nom d'analyse blob, ou *blob analysis* (Buch, 2011).

Une fois les objets détectés au sein d'une image, l'association dans le temps de ces objets est nécessaire afin de suivre leur mouvement au sein d'une vidéo. Le mouvement de chaque objet est estimé par un *filtre de Kalman*. Ce type de filtre estime les états d'un système dynamique à partir d'une série de mesures incomplètes ou bruitées. Ainsi, il assigne une probabilité d'appartenance d'objet à un mouvement en fonction des informations extraites des images précédentes (Kotsialos, 2001). Un *filtre de Kalman* peut permettre de suivre plusieurs objets en mouvement dans la même vidéo.

2.2.3 Les systèmes à deux capteurs

Les systèmes à deux capteurs se basent sur le principe de la *stéréovision*. La stéréovision regroupe l'ensemble des techniques mises en œuvre pour reproduire une perception du relief à partir de deux images planes prises depuis deux points de vue proches. A ce titre, on constate que les techniques regroupées dans la littérature scientifique sous le nom de *stéréovision* se basent sur des principes fondamentaux de *photogrammétrie*. Bien qu'il soit rarement fait mention du terme « *photogrammétrie* » dans la littérature scientifique relative à la capture du mouvement, ce constat permet, entre autres, de justifier l'existence de ce mémoire en géomatique et géométrie.



Avec :

b : ligne de base, soit la distance entre les deux capteurs

f : distance focale d'un capteur

X : axe des X d'un capteur

Z : axe optique d'un capteur

P : point réel défini par les coordonnées X , Y et Z

u_L : projection du point réel P en une image acquise par le capteur de gauche

u_R : projection du point réel P en une image acquise par le capteur de droite

Figure 3 : Schéma illustrant le principe fondamental des systèmes à stéréovision.

(Source : <http://www.ni.com/>)

Les systèmes de stéréovision nécessitent que les capteurs soient parallèles et aient la même distance focale. Une illustration du principe est visible à la Figure 3. Puisque les deux capteurs sont séparés d'une distance b , appelée *base*, la position d'un même point réel P est donc différente sur les images 2D issues de ces capteurs. Les abscisses de ces points homologues u_L et u_R sont donc définies par :

$$u_L = f * \frac{X}{Z}$$

$$u_R = f * \frac{X - b}{Z}$$

La distance entre points homologues projetés est appelée « *disparité* ». Sur base de celle-ci, la *profondeur de champ* peut être calculée, c'est-à-dire la distance entre le système de vision stéréoscopique et le point réel.

$$disparité = u_L - u_r = f * \frac{b}{Z}$$

$$profondeur = f * \frac{b}{disparité}$$

Dans la réalité, les hypothèses idéales énoncées pour le système simplifié de vision stéréoscopique ne peuvent pas s'appliquer. En effet, une installation de mesure stéréoscopique est toujours plus complexe et le parallélisme des capteurs n'est jamais totalement respecté. Les capteurs, même les plus performants, introduisent aussi des déformations qu'il est nécessaire de mesurer et de corriger avant toute reconstitution 3D. Pour cela, l'étalonnage des capteurs à l'aide d'une grille d'étalonnage prise sous différents angles peut aider.

En stéréovision, la première étape consiste donc à identifier les pixels homologues au sein d'une paire d'images stéréoscopiques. Par triangulation, on peut ensuite établir une représentation 3D d'un objet présent sur l'image par intersection des rayons homologues issus de chaque point de vue. Plus le nombre de pixels homologues est grand, plus le nombre de points 3D qu'il est possible de reconstituer est grand sur une seule paire d'images. La sélection des pixels homologues peut se faire à la main, mais des méthodes de corrélation existent pour tenter d'obtenir des correspondances pour chaque pixel dans des images stéréoscopiques. Ces méthodes sont moins précises que la sélection manuelle, mais se révèlent très utiles lorsqu'un grand nombre d'images est à traiter, comme c'est le cas lors de la capture du mouvement d'un objet sur base de vidéos par exemple. Dans ce cas, une caméra stéréoscopique (à deux points de vue) peut être utilisée. Connaissant la longueur de la base qui sépare les deux points de vue, la caméra stéréo mesure la distance de visibilité directe sur base du principe illustré à la Figur. Il est alors possible, avec des niveaux de précisions variant en fonction du système utilisé, de représenter en 3D des objets filmés depuis chacun des points de vue de la caméra.

Une méthode de capture de mouvement par stéréovision a été proposée en 2014 (Starbuck et al. 2014). Cette méthode compare également ces résultats avec les systèmes RGB-D. En laboratoire, cette méthode s'est révélée comparable à l'approche basée sur les capteurs RGB-D. Des avantages supplémentaires inhérents à l'approche par stéréovision ont aussi été mis en avant. Ainsi, cette dernière ne souffre pas des conditions environnementales et peut être appliquée tant à l'intérieur qu'à l'extérieur. De plus, la plage de fonctionnement de la caméra stéréoscopique n'est pas limitée à la précision d'un capteur de profondeur, mais dépend du champ de vision de l'objectif, de la séparation des lentilles, de la taille de l'image et de la résolution des pixels. Cependant, le calcul de l'information de profondeur à partir de deux images nécessite des connaissances mathématiques en géométrie spatiale plus poussées que pour les systèmes RGB-D (Seo et al. 2017).

2.2.4 Les systèmes multi-capteurs

Le but des systèmes de capture de mouvement multi-capteurs (plus de deux capteurs) est d'estimer des coordonnées 3D sur base également de principes fondamentaux de photogrammétrie. Ces systèmes ont l'avantage de capturer le mouvement d'un objet sous des points de vue plus nombreux qu'en stéréovision classique.

Plusieurs méthodes de capture du mouvement à multiples points de vue sont présentes dans la littérature scientifique (Han and Lee, 2013 ; Corazza et al., 2006 ; Gruen, 1997). Ces méthodes se basent sur des prises de vues synchronisées réparties autour d'un modèle en mouvement. Partant du principe que les prises de vues, souvent des vidéos, sont synchronisées, il est alors possible de découper ces prises de vues en images. Cette opération dépend de la fréquence des vidéos, exprimée en nombre d'images par seconde. Ainsi, une vidéo filmée à une résolution de 1080 pixels de largeur et à une fréquence de 30 images par seconde, permettra d'extraire 30 images toutes les secondes à cette résolution. Si le recouvrement entre les images issues des différents points de vue est suffisant (plus de deux tiers de l'image) et que l'angle entre les prises de vue n'excède pas 15 degrés, il est alors possible de représenter l'état du modèle en 3D à une fréquence de 30 modèles par seconde. Si un réseau de points de contrôle a été mis en place pour que ces différents modèles soient reconstruits dans le même système de coordonnées, il est possible d'orienter ces différents modèles dans leur position relative. Il est alors également possible d'afficher chacun des modèles les uns à la suite des autres dans une animation 3D.

Plus le nombre de dispositifs de prise de vue est élevé et plus la répartition de ces derniers autour du modèle est uniforme, plus il est possible de reconstituer précisément chacun des instants du mouvement.

Il existe deux types de systèmes de capture de mouvement à multiples capteurs. Les premiers avec marqueurs, et les seconds sans marqueurs placés sur le modèle.

- **Avec marqueurs placés sur le modèle :**

Les avantages de marqueurs placés sur l'objet en mouvement sont les suivants.

Premièrement, ces marqueurs diminuent le nombre de points dont il faut déterminer les positions en 3D. En effet, si le modèle peut être rattaché à ces marqueurs, il n'est alors pas nécessaire de calculer l'ensemble des points du modèle. On peut se limiter au calcul des positions des marqueurs pour déterminer les composantes du mouvement et, lors de la construction de l'animation, rattacher une représentation 3D de notre modèle à ces marqueurs. Cette méthode est applicable si le modèle en mouvement est rigide. En biomécanique par exemple, cette méthode ne permettrait pas de mesurer l'ensemble des mouvements d'un corps étant donné la variation des positions relatives entre les membres. Il faudrait pour cela diviser le corps humain en une multitude d'entités considérées comme rigides (tête, avant-bras, jambes, etc.) et appliquer cette méthode pour chacune des entités.

Deuxièmement, la présence de marqueurs influence la précision de la reconstitution des différents modèles 3D et donc la précision de la capture du mouvement. Ces marqueurs permettent de contraindre l'orientation relative des images permettant la reconstitution. La détermination des positions 3D de ces points sera donc plus précise. L'aspect négatif de cette méthode est que la précision de la reconstitution du mouvement du modèle ne dépend que d'un nombre restreint de points. Une erreur dans la détermination de ces points engendrerait des erreurs directes sur l'ensemble de la reconstitution du mouvement.

- **Sans marqueurs placés sur le modèle :**

L'avantage de l'absence de marqueurs placés sur le modèle est qu'il serait possible de capturer le mouvement d'objet non rigide (ex. : corps humain). De plus, développer une méthode sans marqueurs permet des applications réalistes de capture de mouvement. Dans l'industrie du jeu vidéo, la capture des mouvements faciaux permet de créer des personnages aux expressions fidèles à la réalité. Une absence de marqueurs permet de reproduire en 3D la totalité du visage, sans passer par un rattachement à un modèle quelconque. Ces systèmes sans marqueurs donnent des résultats proches de la réalité, mais nécessitent des prises de vues rapprochées et sont souvent associés avec des technologies plus poussées analysant l'homographie des prises de vue. Sans capteurs, on peut donc difficilement se limiter à une reconstruction photogrammétrique simple du modèle. Des méthodes d'analyse d'images peuvent aider.

Les systèmes de capture de mouvement multi-capteurs dépendent donc des caractéristiques de ces derniers (longueur focale, résolution des pixels, recouvrement et angle entre les prises de vues) et de leur calibration. Les conditions environnementales limitent peu cette technique et les données peuvent toujours être modifiées en post-traitement si les conditions d'éclairage ou les couleurs sont limitantes (Seo et al. 2017).

On peut distinguer six étapes composant les méthodes multi-capteurs de capture et de représentation du mouvement en 3D.

1. La mise en place du dispositif de capteurs, de points de contrôle et de marqueurs

La mise en place du dispositif de capteurs dépend du nombre et de la résolution des capteurs utilisés, des limites de recouvrement et d'angles imposés entre les prises de vue et de la précision désirée du modèle final. Ces paramètres vont influencer directement la répartition géométrique des capteurs autour du modèle. La mise en place des points de contrôle est indispensable pour permettre d'assembler l'ensemble des modèles dans le but d'en faire une animation 3D. À cette fin, la disposition des points de contrôle doit être pensée pour favoriser un maximum de précision sur chacune des reconstitutions 3D qui seront générées. La possible occlusion de ces points par le modèle en mouvement doit également être prise en compte. Dans le cas de marqueurs placés sur l'objet en mouvement, il est également nécessaire de prévoir des marqueurs qui peuvent être détectés depuis chacun des points de vue, faute de quoi la position de ces derniers ne pourrait pas être calculée à chaque instant de la vidéo et l'animation finale ne pourrait pas être compilée.

2. La synchronisation des prises de vue

Afin d'être capable de reconstituer un modèle 3D de l'objet en mouvement à chaque instant depuis des points de vue séparés, il est nécessaire de synchroniser les prises de vues. Depuis des vidéos synchronisées, il est alors possible d'extraire des ensembles d'images correspondant au même instant de prise de vue. Plusieurs méthodes sont possibles pour synchroniser des prises de vue. Ainsi, on peut y arriver à l'aide de la bande-son des vidéos, de leurs images, ou bien en démarrant les vidéos au même instant de manière électronique. Ces trois pistes seront explorées pour cette recherche.

3. Le découpage en images en fonction de la fréquence des prises de vue

Dû à l'impossibilité de reconstruire des modèles 3D directement depuis des vidéos issues de points de vue différents, il est nécessaire de découper ces vidéos en images. Les vidéos étant synchronisées, elles peuvent donc être découpées à l'aide de logiciels adéquats.

4. La construction des modèles 3D pour chacun des instants du mouvement

À chaque instant découpé, il est possible de reconstituer en 3D l'objet en mouvement ou de calculer la position des marqueurs placés sur l'objet. Cette reconstitution est possible à l'aide de plusieurs logiciels.

5. L'interpolation temporelle des différents modèles 3D

Un des points clés de la capture du mouvement est l'interpolation temporelle des différents modèles 3D créés. En effet, une fois les modèles construits, nous ne pouvons représenter qu'une discrétisation du mouvement. Ceci est dû au fait qu'un découpage en images des vidéos a été nécessaire. Des méthodes d'interpolation 3D appliquées à l'ensemble du modèle ou simplement aux marqueurs peuvent permettre d'avoir une idée précise du mouvement continu de l'objet dans l'espace.

6. La représentation du mouvement dans une animation 3D

L'animation 3D est le résultat final de la capture du mouvement. Elle peut être construite à partir de plusieurs logiciels différents, propriétaires ou *open source*.

Nous détaillerons ces six étapes au point 4 « Développement théorique » de ce mémoire.

3 ORIENTATION ET HYPOTHESES DE RECHERCHE

Après un état de l'art des techniques existantes en matière de capture de mouvement, il est maintenant nécessaire d'identifier les différentes contraintes liées à la conception d'une méthode d'acquisition, de traitement et de rendu 3D d'objet en mouvement. La confrontation de ces contraintes avec l'état de l'art et les ressources disponibles au sein de l'Unité de Géomatique, ainsi qu'aux résultats d'une série de tests effectués par nos soins, nous permettra d'orienter notre recherche et d'établir nos hypothèses de travail.

3.1 Contraintes

Les différentes contraintes sont liées soit à la mise en place du dispositif d'acquisition, soit à l'œuvre d'art ou l'objet dont il est nécessaire de capturer le mouvement. Nous pouvons lister les contraintes suivantes :

- L'animation 3D finale doit être géoréférencée dans un système de coordonnées local afin de décomposer le mouvement et en mesurer les composantes tridimensionnelles (positions, vitesses, orientations, etc.).
- La méthode à développer requiert du matériel d'acquisition déjà présent dans l'Unité de Géomatique.
- L'œuvre ne peut pas être déplacée pour l'acquisition des données. Il est donc nécessaire de mettre au point une méthode présentant un dispositif d'acquisition adaptable et déplaçable. Cette contrainte rencontre la volonté de l'Unité de Géomatique de développer une méthode d'acquisition modulable.
- Afin de ne pas altérer l'œuvre d'art, cette dernière ne pourra pas servir de support à des marqueurs définitifs pouvant l'endommager ou la modifier définitivement.
- L'acquisition des données devant se faire in situ, la méthode d'acquisition doit être pensée pour être effectuée en une fois et dans une limite de temps inférieure à 4 heures.
- Afin de rendre compte de l'*effet de l'œuvre*, le détail des différents matériaux composant cette dernière doit être visible dans le rendu 3D final. Il est donc nécessaire de prévoir un mode d'acquisition permettant cette distinction.

3.2 Ressources disponibles

La méthode d'acquisition devant se limiter à l'utilisation des ressources présentes au sein de l'Unité de Géomatique, il convient de faire l'état des ressources matérielles et informatiques disponibles.

• Ressources matérielles :

Les ressources matérielles disponibles pour la mise en place d'une méthode d'acquisition sont les suivantes :

- Une station totale LEICA TPS 1205
- Un scanner laser LEICA P30
- Un boîtier reflex CANON EOS 5D Mark III
- Un boîtier reflex CANON EOS 70D
- Un objectif reflex SIGMA DG EX 15 mm 1 : 2,8
- Un objectif reflex SIGMA DG EX 8 mm 1 : 3,5
- Un objectif reflex CANON EF 24-105 mm 1 : 4
- Une télécommande sans fil CANON RC-6

Les caractéristiques de précision de l'ensemble du matériel sont disponibles dans l'Annexe 1.

• Ressources informatiques

Les ressources informatiques disponibles pour le traitement des données issues de l'acquisition sont les suivantes :

- Le logiciel CONTEXT CAPTURE : logiciel créé par la société BENTLEY qui permet de produire des modèles 3D à haute résolution sur base d'un minimum de trois photographies d'un sujet.
- Le logiciel CYCLONE : logiciel créé par la société LEICA permettant de traiter, assembler et nettoyer les nuages de points 3D issus d'un scanneur laser terrestre.
- Le logiciel CLOUDCOMPARE : logiciel *open source* d'édition, de traitement et d'affichage de *meshs* (maillage surfacique triangulaire) et de nuages de points 3D. Il permet notamment de calculer les distances entre deux nuages de points 3D ou un nuage de points et un *mesh*.
- Le logiciel AUTOCAD MAP 3D : logiciel de dessin assisté par ordinateur créé par la société AUTODESK.
- Le logiciel COVADIS : logiciel de topographie, de terrassement, d'infrastructure et de génie civil créé par la société GEOMEDIA et pouvant servir comme extension dans le logiciel AUTOCAD MAP 3D.
- 3DS MAX : logiciel de modélisation et d'animation 3D, développé par la société AUTODESK. C'est un logiciel de référence dans le domaine de l'infographie 3D et de la *motion capture*.

3.3 Orientation de la recherche

Avant d'aller plus loin dans la démarche de recherche, il est nécessaire de confronter l'état de l'art effectué au chapitre n° 3 avec l'ensemble des objectifs, contraintes et ressources ainsi qu'avec un ensemble de tests effectués préalablement.

Premièrement, lorsqu'on compare les ressources matérielles et informatiques de l'Unité de Géomatique avec les ressources nécessaires pour les systèmes de capture du mouvement présentés dans l'état de l'art, nous nous rendons compte que nous ne disposons d'aucune ressource matérielle permettant de mettre en place un système d'acquisition mécanique, acoustique ou magnétique. Il est donc nécessaire de se limiter à un système d'acquisition optique. Un système d'acquisition optique permettrait la distinction des différents matériaux de l'objet en mouvement, répondant ainsi à la dernière de nos contraintes. Parmi les systèmes optiques, les systèmes utilisant des capteurs RGB-D peuvent être écartés pour les mêmes raisons. Bien que les capteurs RGB-D sont des appareils relativement peu chers en comparaison avec le matériel présent à l'Unité de Géomatique, le but n'est pas d'acquérir du nouveau matériel, mais d'explorer les possibilités de capture du mouvement du matériel existant. Nous opterons donc pour une méthode d'acquisition optique non RGB-D.

Deuxièmement, la réponse à la question du nombre de capteurs optiques à utiliser pour développer un système d'acquisition optique (mono capteur, à deux capteurs ou à capteurs multiples) va également dépendre des objectifs de recherche. Comme mentionné au début de ce point 3.3, des tests ont été effectués par nos soins pour chacun des systèmes optiques afin de déterminer lequel d'entre eux correspondait le mieux aux objectifs de recherche, compte tenu des contraintes et ressources établies.

- **Un système optique mono-capteur**

Pour les systèmes de capture du mouvement optique mono capteur, nous avons effectué des tests sur les systèmes de détection et le suivi du mouvement d'un modèle. Des exemples de ces tests sont visibles à la Figure 4 pour la détection et le suivi de visage en mouvement, et à la Figure 5 pour la détection et le suivi de mouvement d'objet.



Figure 4 : Tests de détection et de suivi de visage en mouvement à l'aide d'une analyse blob et d'un filtre de Kalman.

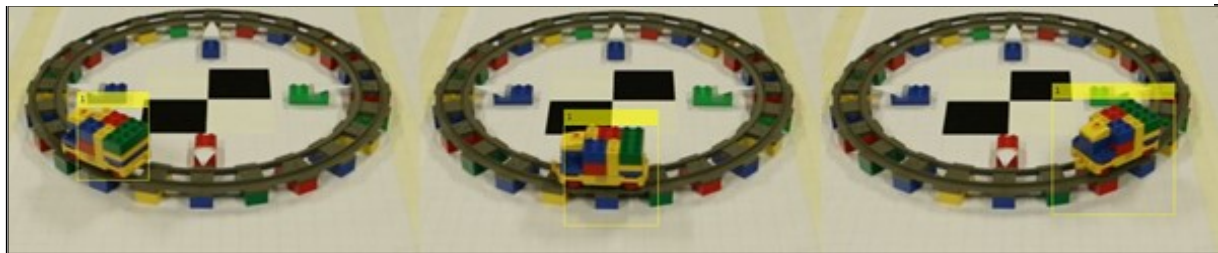


Figure 5 : Tests de détection et de suivi d'objet en mouvement à l'aide d'une analyse blob et d'un filtre de Kalman.

Les tests effectués préalablement à la recherche permettent de tirer les conclusions suivantes.

Premièrement, les algorithmes testés pour la capture du mouvement sur base d'un seul capteur ne permettent pas de déceler le mouvement à une précision suffisante que pour rencontrer la précision millimétrique des objectifs de recherche. De plus, perfectionner ces algorithmes aurait nécessité des connaissances supplémentaires en *deep* et *machine learning*. Ce perfectionnement aurait pu, à lui seul, faire l'objet d'un travail de fin d'études, mais n'aurait pas, dans ce cas, rencontré les objectifs de ce mémoire.

Deuxièmement, les possibilités qu'offre la capture du mouvement optique mono-capteur ne correspondent pas aux objectifs de recherche. En effet, ces systèmes permettent, à l'aide d'algorithmes variés, de détecter, de suivre et de classifier des objets dans une image ou une vidéo 2D. Ces perspectives sont des thématiques de prédilection pour des applications en vision assistée par ordinateur, mais ne constituent aucun intérêt pour notre recherche. Notre objectif n'étant pas de détecter le mouvement ou de le traquer, mais bien de capturer ces composantes tridimensionnelles. Pour cela, un minimum de deux points de vue est nécessaire. Il est en effet difficile d'extraire une information tridimensionnelle d'un seul capteur optique RGB. Dans la littérature scientifique, certains y sont arrivés (Zhu et al., 2016), mais, dans ce cas, les niveaux de précision atteints ne rencontrent pas nos objectifs de recherche.

Pour toutes ces raisons, la piste d'un système de capture du mouvement optique mono capteurs a été écartée.

- **Un système optique à deux capteurs**

Pour les méthodes de capture de mouvement optique à deux capteurs (stéréovision), des tests de reconstitution 3D ont été entrepris afin de comprendre les limitations techniques et d'évaluer les niveaux de précision auxquels il est possible de s'attendre. Un exemple de test de reconstitution 3D est visible à la Figure 6 et un exemple de calibration pour l'orientation relative des caméras est visible à la Figure 7.

Les tests effectués permettent de tirer les conclusions suivantes.

Premièrement, bien que les systèmes à stéréovision permettent de reconstituer des modèles 3D d'objets à des niveaux de précision acceptables, ces reconstitutions ne pourront jamais représenter l'ensemble de l'objet tout simplement parce qu'elles sont limitées par les deux points de vue du système. Il serait donc difficile de reconstituer un modèle 3D complet d'un objet en mouvement basé uniquement sur un système de capture par stéréovision.



Figure 6 : Test de reconstitution 3D par stéréovision.



Figure 7 : Test de calibration pour l'orientation relative du système à stéréovision

Deuxièmement, à la suite des tests que nous avons effectués, il a été mis en évidence que pour reconstituer un modèle 3D par stéréovision, la longueur focale des lentilles et la résolution des deux prises de vues devaient être identiques afin de calibrer le système de stéréovision et calculer l'orientation relative des deux capteurs. Idéalement, pour obtenir un système d'acquisition opérationnel, il est aussi nécessaire de posséder deux capteurs aux caractéristiques techniques identiques. Les deux appareils photo reflex présents à l'Unité de Géomatique permettent tous deux de filmer à une résolution de 1920 pixels par 1080 pixels, mais ils ne possèdent pas le même format de capteur. En effet, le CANON EOS 5D Mark III est un appareil plein-format, *full frame* en anglais, et le CANON EOS 70D est un appareil APS-C, *Advanced Photo System type-C*. Le deuxième possède un capteur 1,6 fois plus petit que le premier. Cette différence de capteurs engendre des complications pour le calcul de l'orientation relative des caméras, et donc la reconstitution d'un modèle 3D. De plus, cette complication est accentuée par le fait que nous n'avons pas en notre possession deux objectifs de longueurs focales identiques.

Pour ces deux raisons, la piste de capture de mouvement par stéréovision a été abandonnée.

- **Un système optique multi-capteurs**

Pour les systèmes optiques multi-capteurs, des tests de reconstitution 3D d'objet en mouvement sous différents angles ont été effectués. Un exemple de résultat 3D est visible à la Figure 8.



Figure 8 : Test de reconstitution 3D multi-capteurs d'un objet en mouvement.

Les tests effectués permettent de tirer les conclusions suivantes.

Premièrement, en plaçant des capteurs optiques tout autour d'un objet, il est possible de capturer l'intégralité de l'objet en 3D. Cette disposition permettrait donc de représenter la totalité du mouvement de l'objet dans une animation 3D.

Deuxièmement, le nombre de ressources matérielles disponibles est un facteur limitant. En utilisant une technique nécessitant un nombre élevé de capteurs optiques, il est impossible d'effectuer la capture du mouvement en une seule prise, tout simplement parce que nous n'avons pas à notre disposition un tel nombre de capteurs. Ce système optique multi-capteurs pourrait néanmoins s'appliquer à un mouvement facilement reproductible. En effet, si le mouvement peut être reproduit facilement et avec une précision suffisante, il est possible d'effectuer une succession de prises de vues sous un angle différent, et ce jusqu'à l'acquisition totale de notre objet en mouvement sous tous les angles nécessaires pour le reconstituer en 3D à la précision voulue.

Pour ces raisons, la piste de capture de mouvement optique multi-capteurs est une piste qui sera explorée. La question reste à savoir si cette méthode doit être appliquée avec ou sans marqueurs posés sur l'objet.

L'option avec capteurs posés sur l'objet permettrait, comme exposé dans l'état de l'art, une plus grande précision sur le modèle et une plus grande facilité de reconstitution du mouvement. De plus, elle permettrait de se limiter uniquement au calcul de la position des marqueurs afin de venir y rattacher une représentation 3D de l'objet. Contraint par le fait que l'objet en mouvement est une œuvre d'art, placer ces capteurs sur l'objet ne serait possible que dans l'éventualité où ces derniers soient facilement retirables et ne détériorent pas l'œuvre.

L'option sans marqueurs posés sur l'objet présente, quant à elle, le désavantage d'être moins précise, mais a l'avantage de permettre de reconstituer l'objet en 3D sans l'endommager, ou bien d'être appliquée à des objets au mouvement plus complexe sur lesquels la pose de marqueurs est impossible.

Pour ces raisons, les cas avec et sans marqueurs placés sur l'objet seront explorés.

- **Le scanneur laser terrestre**

Bien qu'aucune mention de la lasergrammétrie n'ait été aperçue dans la littérature scientifique relative à la capture du mouvement, nous aimerions connaître les possibilités du scanneur laser LEICA P30 en la matière. Pour cela, une méthode d'acquisition utilisant le scanneur laser terrestre sera également pensée et testée.

3.4 Hypothèses de recherche

Après avoir réalisé un état de l'art de la recherche en matière de capture de mouvement, établi les objectifs de recherches, les contraintes et les ressources disponibles, ainsi qu'orienter notre recherche à l'aide de tests effectués par nos soins, nous sommes maintenant à même d'exposer les hypothèses qui vont guider ce travail de fin d'études.

Les hypothèses de recherche sont les suivantes :

- **H1** : « *Un système optique et multi-capteurs de capture du mouvement, basé sur la photogrammétrie, permet de développer une méthode d'acquisition, de traitement et de rendu 3D d'objets en mouvement, à une précision millimétrique.* »
 - **H1-A** : « *Ce système de capture du mouvement optique multi-capteurs n'inclut pas des marqueurs posés sur l'objet en mouvement.* »
 - **H1-B** : « *Ce système de capture du mouvement optique multi-capteurs inclut des marqueurs posés sur l'objet en mouvement.* »
- **H2** : « *Un système de capture du mouvement basé sur la lasergrammétrie permettrait de développer une méthode d'acquisition, de traitement et de rendu 3D d'objets en mouvement, à une précision millimétrique.* »

4 DÉVELOPPEMENT THÉORIQUE

Afin de répondre à chacune des hypothèses de recherche, trois méthodes d'acquisition seront développées sur base théorique dans ce chapitre.

Tout d'abord, une présentation de l'objet en mouvement utilisé pour le test de ces méthodes sera faite.

Ensuite, afin de répondre à notre première hypothèse (H1), nous développerons, sur base de nos connaissances théoriques, deux méthodes optiques et multi-capteurs permettant de capturer le mouvement, l'une avec marqueurs placés sur l'objet et l'autre sans marqueurs placés sur l'objet. L'application de ces deux méthodes permettra de valider ou d'invalidier les sous-hypothèses H1-A et H1-B.

Finalement, pour répondre à notre deuxième hypothèse de recherche (H2), nous développerons également une méthode de capture du mouvement basée sur la lasergrammétrie.

4.1 Objet en mouvement

N'ayant pas accès de façon permanente à l'œuvre d'art cinétique « 24 boules sur 24 sommets inclinés » et souhaitant développer une méthode de capture adaptable à d'autres objets en mouvement, les tests seront effectués sur un objet de forme plus simple qu'il sera possible de manipuler facilement en laboratoire.

Pour cela, le mouvement de cet objet doit satisfaire plusieurs conditions. Premièrement, comme pour l'œuvre d'art cinétique, ce mouvement doit être facilement reproductible. Deuxièmement, les tests se feront sur un objet se déplaçant à vitesse constante. Bien que cela restreigne l'ensemble des objets auxquels la méthode développée pourrait s'appliquer, cela permettra plus de facilité lors du traitement des données et du rendu 3D du mouvement.

Ainsi, l'objet en mouvement utilisé pour les tests sera une locomotive miniature de la marque LEGO. Ce jouet présente une forme parallélépipédique et est motorisé. Il se déplace à vitesse constante sur un support fixe, de forme circulaire ou courbe, fait de rails en plastique (voir Figure 9). Le support de rails est stable et ne se déforme pas dans le temps. L'avantage de ce dispositif est que les rails imposent un mouvement régulier et reproductible à l'objet. La reproductibilité du mouvement est indispensable étant donné les moyens techniques limités que nous disposons pour développer nos acquisitions. En effet, ne possédant pas les ressources matérielles nécessaires à l'Unité de Géomatique pour implémenter ces méthodes de manière idéale, un mouvement parfaitement reproductible permet d'effectuer l'acquisition du mouvement depuis des points de vue différents à des instants différents. Il sera alors possible de capturer le mouvement sous différents angles avec le même capteur en faisant comme si nous disposions de plusieurs capteurs identiques.

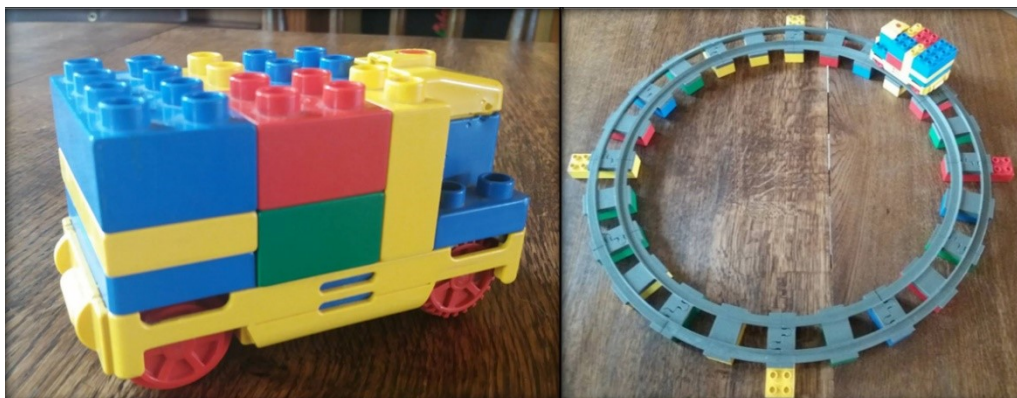
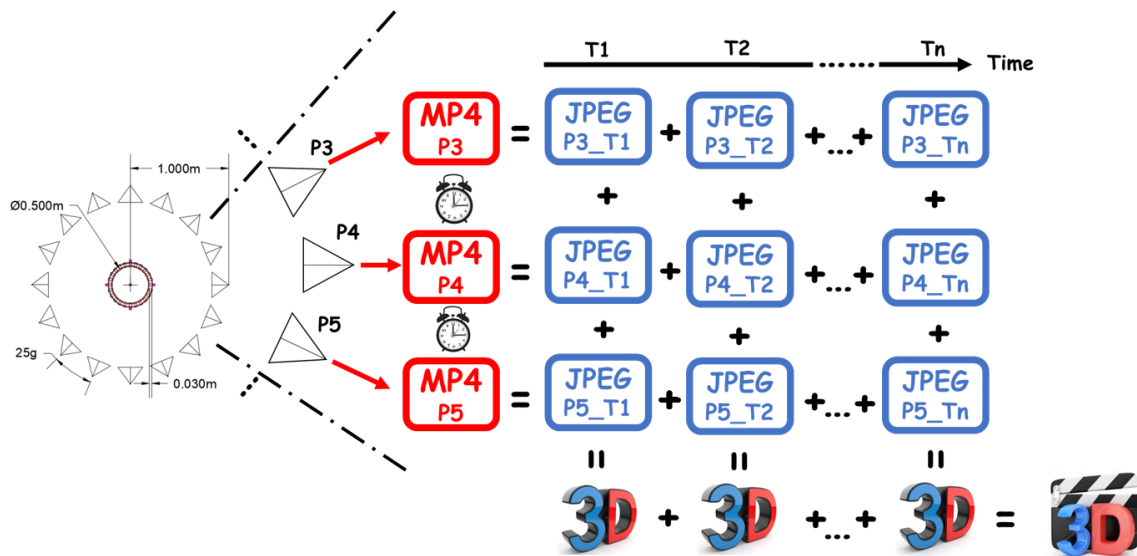


Figure 9 : Objet utilisé pour les expériences et support du mouvement circulaire

4.2 Sans marqueurs placés sur l'objet

Sur base de l'état de l'art, il a été décidé que le système optique multi-capteurs sans marqueurs placés sur l'objet sera composé d'un ensemble de points de vue desquels le mouvement de l'objet pourra être entièrement observé. Depuis chacun de ces points de vue, l'objet sera filmé à la plus haute fréquence possible, ainsi qu'à la plus grande résolution du capteur. Par la suite, il sera alors possible, moyennant synchronisation des vidéos, d'extraire une image pour chaque instant du mouvement et depuis chacun des points de vue. Sur base d'un ensemble d'images correspondant au même instant du mouvement, une reconstitution photogrammétrique en 3D est alors théoriquement possible. La succession de ces reconstitutions dans un logiciel spécialisé permettra de réaliser une animation de l'objet en 3D à chaque instant du mouvement (voir Figure 10).



Avec :

P_n : Position du point de vue n

T_n : Instant n du mouvement

Figure 10 : Schéma de la méthode optique et multi-capteurs pour capturer le mouvement sans marqueurs placés sur l'objet.

Sur base de ces suppositions théoriques, on peut maintenant établir précisément notre méthode optique et multi-capteurs de capture du mouvement sans marqueurs placés sur l'objet. Le développement d'une telle méthode peut être divisé en trois étapes principales. Chacune de ces étapes doit être pensée de manière à aboutir à une méthode fonctionnelle. Ces trois étapes sont l'acquisition du mouvement, le traitement des données et le rendu 3D.

4.2.1 Acquisition du mouvement

Pour ce qui est de l'acquisition du mouvement, nous devons répondre à trois questions avant de développer une méthode d'acquisition fonctionnelle. Comment disposer les capteurs ? Comment référencer les modèles dans un système local ? Et comment synchroniser les prises de vues ? Conscient qu'il n'est pas possible de tester toutes les possibilités de réponse, une seule piste de solution sera abordée et testée pour chacune de ces questions.

• Comment disposer les capteurs ?

Grâce à notre état de l'art et à notre expérience pratique de la photogrammétrie, nous pouvons affirmer qu'une disposition, à intervalle régulier, de capteurs autour de l'objet est une disposition optimale pour reconstituer un modèle photogrammétrique complet en 3D. La distance entre les points de vue et l'objet en mouvement ainsi que le nombre de points de vue sont, quant à elles, des variables qui vont dépendre des caractéristiques techniques du capteur et des conditions de reconstitution 3D du logiciel que nous utiliserons, à savoir CONTEXT CAPTURE.

Ainsi, la taille de pixels projetée sur notre modèle sera notre facteur limitant. De ce facteur découlent des variables comme la résolution et la précision du modèle généré. Il est donc important de fixer ce facteur avant toute acquisition. Comme mentionné dans les contraintes initiales, la taille du pixel projeté doit être au maximum de 0,001 m. Nous allons maintenant déterminer la distance focale et la distance capteur-objet nécessaires pour atteindre cette valeur.

Dans notre cas, le capteur utilisé sera un CANON 5D MARK III. Ce dernier possède une taille de capteur de 36 x 24 mm. L'objectif dont il est équipé est un CANON EF 1 : 2,8 d'une distance focale de 24 mm. La résolution d'image est de 1920 x 1080 pixels pour une fréquence de prise de vue vidéo de 30 *frames per second* ou *fps*

Connaissant maintenant les caractéristiques du capteur et de son objectif, il est possible de déterminer la distance capteur-objet minimum nécessaire pour obtenir une taille de pixel projeté de maximum 0,001 m sur le modèle 3D. On peut déterminer ces valeurs avec la formule ci-dessous.

Soit :

$$\begin{array}{ll}
 Ri : \text{Résolution image} & Lc : \text{Largeur du capteur} \\
 Pp : \text{Taille du pixel projeté} & Dc : \text{Distance capteur} \\
 Df : \text{Distance focale} &
 \end{array}$$

On a donc :

$$Dc = \frac{Df}{Lc} * Ri * Pp = \frac{0,024}{0,024} * 1080 * 0,001 = 1,080 \text{ m}$$

Après calcul, on constate que la distance capteur-objet minimum permettant la taille de pixels projetés désirée est 1,080 m. Afin de se donner une marge de sécurité, il a été décidé de placer le capteur à 1,000 m du centre de dispositif.

Après avoir déterminé la distance capteur-objet nécessaire, il est utile de trouver le nombre de points de vue suffisant pour réaliser nos reconstitutions. Les recommandations des développeurs du logiciel CONTEXT CAPTURE imposent un minimum de 3 points de vue présentant un recouvrement minimum de deux tiers de l'image et un écart entre les prises de vues de minimum 16,667 gons (15 degrés). En suivant ces recommandations, il serait nécessaire d'avoir un minimum de 24 points de vue, soit un capteur placé tous les 16,667 gons autour de l'objet. Néanmoins, la multiplication des points de vue engendre des complications et du temps supplémentaire pour la synchronisation en post-traitement. Des tests que nous avons effectués ont mis en évidence que le logiciel pouvait reconstituer complètement notre objet en 3D avec seulement 16 points de vue différents.

Ainsi, 16 capteurs optiques seront disposés autour de l'objet en mouvement avec un écart angulaire de 25,000 gons et à une distance de 1,000 m du centre du dispositif. Sur base de ces informations, le schéma présenté à la Figure 11 a été dressé à l'aide du logiciel AUTOCAD MAP 3D.

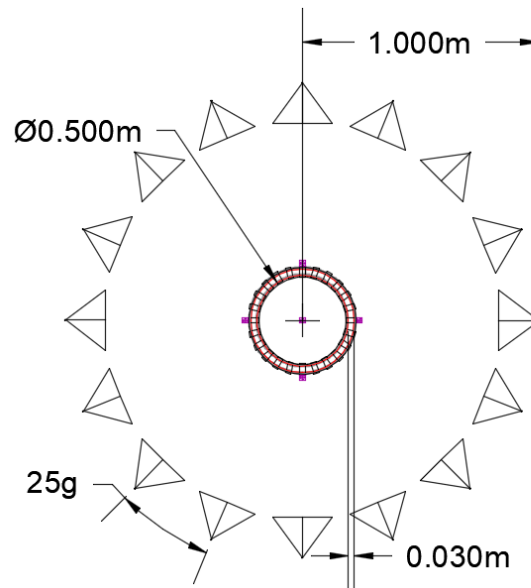


Figure 11 : Disposition des capteurs pour la méthode optique et multi-capteurs sans marqueurs placés sur l'objet.

• Comment référencer les modèles 3D dans un système de coordonnées local ?

Afin de calculer les composantes du mouvement dans le système international des unités de mesure, d'implanter correctement les positions des capteurs calculées au point précédent et d'utiliser l'espace disponible dans le laboratoire de la manière la plus optimale possible, il est nécessaire d'effectuer nos mesures dans un système de coordonnées local.

Pour cela, la station totale LEICA TPS1205 et le scanneur laser LEICA P30 seront utilisés afin de modéliser la pièce dans laquelle nous effectuerons nos mesures. Des cibles de scanneur seront placées dans la pièce et un repère sera placé à l'autre bout de cette dernière afin de procéder à une orientation du tachéomètre. Après avoir mis ce dernier en station à l'emplacement désiré, l'emplacement de la station totale sera marqué au sol à l'aide du plomb laser et les centres des différentes cibles seront levés. Par la suite, 4 scans (un dans chaque coin de la pièce) à une résolution de 3,1 mm à 10 m seront effectués. Sur chacun de ces scans, l'emplacement marqué au sol de la station totale sera levé avec une canne sur laquelle une cible sera placée. Ces 4 scans seront assemblés sur ordinateur à l'aide du logiciel CYCLONE, en utilisant comme nuage de points de référence celui contenant les coordonnées des cibles dans le système local.

Afin d'obtenir uniquement des coordonnées positives, une translation du nuage de points final sera effectuée de manière à placer l'origine du système de coordonnées à l'extérieur de la pièce. Ce changement de coordonnées s'effectuera à l'aide du logiciel AUTOCAD MAP 3D.

Une fois dans le nouveau système de coordonnées, on superposera, en vue de dessus, le schéma du dispositif de prise de vue présenté au point précédent de la meilleure façon possible en fonction de l'espace disponible dans la pièce. Une fois le schéma superposé, les futures positions des capteurs, de la station totale et du repère dans le nouveau système de coordonnées seront exportées à l'aide du logiciel COVADIS. Ces coordonnées permettront de réorienter la station dans la pièce et implanter au sol les positions des capteurs. Elle permettra également d'implanter correctement le centre du support de l'objet en mouvement.

La Figure 12 montre une vue supérieure du nuage de points assemblé sur lequel les positions des capteurs ont été superposées. Cette procédure sera la même pour l'implantation des positions du scanneur pour le système de capture de mouvement basé sur la lasergrammétrie.

Lorsque l'implantation des emplacements des capteurs et du centre du dispositif sera effectuée, le support et l'objet en mouvement pourront être installés. 5 cibles photogrammétriques (1 au centre et 4 autour du dispositif) seront alors mises en place afin de pouvoir réaliser les modèles 3D dans le système local et estimer les précisions sur nos reconstitutions. Ces 5 cibles seront visibles depuis chacun des points de vue. Leurs centres seront levés à la station totale et leurs coordonnées calculées dans le logiciel COVADIS (voir Figure 13).

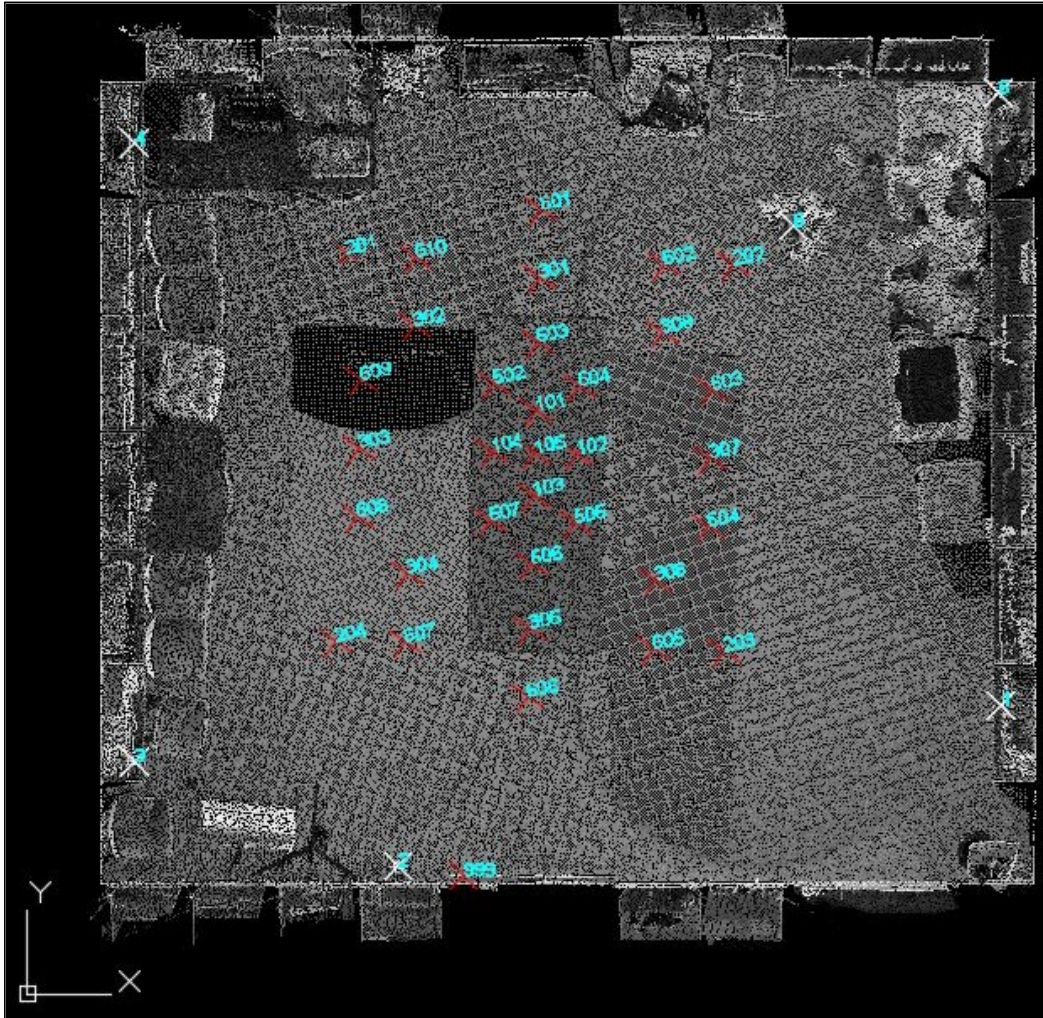


Figure 12 : Vue de dessus du nuage de points du laboratoire avec superposition des futurs emplacements des capteurs.

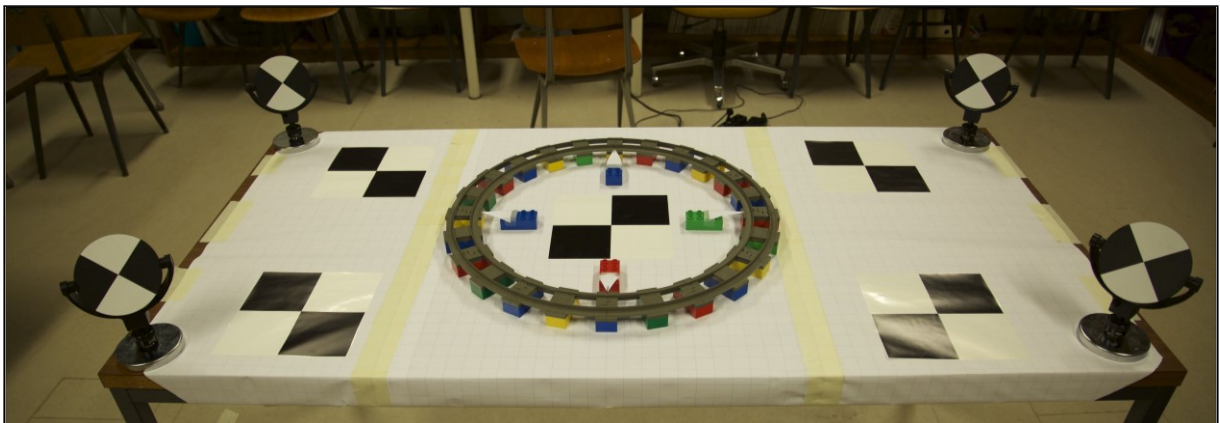


Figure 13 : Dispositions des 5 cibles photogrammétriques à proximité du support (cibles carrées).

- **Comment synchroniser les prises de vues ?**

La synchronisation des prises de vues est indispensable pour assurer la correspondance des instants filmés depuis chacun des points de vue. En effet, reproduire le modèle 3D d'un objet à un instant précis n'est possible que si chacune des images en entrée correspond au même instant précis.

Pour assurer la synchronisation des prises de vue, plusieurs techniques sont possibles en fonction du nombre de capteurs présents. Dans une situation idéale, nous aurions à notre disposition 16 capteurs que nous placerions autour de l'objet en mouvement. Dans ce cas, trois méthodes de synchronisation sont possibles.

Premièrement, des logiciels comme MATLAB ou encore DIGICAMCONTROL permettent, une fois l'ensemble des capteurs connectés à un ordinateur, de synchroniser les prises de vue en les lançant toutes depuis le même logiciel. Pour connaître les capacités de ces deux logiciels, nous avons effectué des tests sur un cas simplifié à deux capteurs. Ainsi, les capteurs CANON 5D MARK III et CANON 70D ont été connectés à un ordinateur à l'aide de deux câbles USB 2.0. Ces tests ont montré que, malgré la synchronisation à l'aide du logiciel, une latence et un décalage des prises de vue, dépendant du type de capteurs et du type de câbles, étaient présents. De plus, aucun de ces deux logiciels n'accepte jusqu'à 16 capteurs en même temps. Pour choisir cette option de synchronisation, nous aurions donc dû disposer de 16 capteurs et 16 câbles USB identiques. Un logiciel spécifique permettant la synchronisation de 16 appareils, ou bien des connaissances en programmation suffisantes que pour programmer le démarrage simultané de 16 capteurs sans latence, auraient également été nécessaires.

Deuxièmement, les capteurs en notre possession permettent le démarrage à l'aide de la télécommande infrarouge CANON RC-6. Les deux capteurs peuvent donc être démarrés simultanément, et ce avec des résultats satisfaisants pour notre application. Cette option aurait pu être choisie si nous possédions suffisamment de capteurs identiques pouvant être démarrés par la même télécommande infrarouge.

Troisièmement, les vidéos issues des capteurs peuvent être synchronisées en post-traitement à l'aide du son enregistré lors de la vidéo. Ainsi, les vidéos peuvent démarrer à des moments distincts et il suffit, par exemple, de claquer des mains pour obtenir un repère auditif qui permettra de synchroniser les vidéos en post-traitement. Cette méthode, utilisée dans de nombreuses applications, est assez précise et à l'avantage de ne pas nécessiter de logiciel coûteux, ni de câbles pour la synchronisation. Elle a, en revanche, le désavantage de ne pas pouvoir être automatisée, car une étape de post-traitement manuelle sur logiciel est nécessaire pour synchroniser les prises de vues.

La synchronisation à l'aide d'un logiciel, d'une télécommande ou du son n'est donc pas possible étant donné les ressources matérielles dont nous disposons.

Dans notre situation, nous ne pourrions donc disposer que d'un seul capteur, car les différences techniques entre les deux capteurs en notre possession ne nous permettent pas de les utiliser simultanément pour notre application. Les prises de vues doivent donc être effectuées sous des angles différents et à des moments différents. Cela est possible uniquement grâce à la répétabilité du mouvement de l'objet. Ainsi, il est théoriquement possible de synchroniser les prises de vues uniquement sur base du son ou des images issues des capteurs. Deux solutions sont alors possibles : placer un repère auditif, c'est-à-dire un son émis à un moment clé du mouvement permettant d'identifier un instant précis de ce dernier ; ou bien placer un repère visuel sur l'objet et le support, ce qui permettrait également d'identifier un instant-clé. Ces deux solutions imposent une étape de post-traitement manuel supplémentaire. Un choix est donc à faire entre un repère auditif ou visuel.

La première option est plus précise que la deuxième, car elle ne dépend pas de la fréquence des prises de vue (30 fps). Le son est enregistré en continu. Un instant du moment peut donc être

identifié de manière plus précise à l'aide d'un repère auditif qu'un repère visuel qui ne pourrait être visible que 30 fois par seconde. Néanmoins, si l'objet n'émet pas de son particulier lié à son mouvement, cette option implique la mise en place d'un dispositif d'émission de bruit lors d'un passage de l'objet en mouvement. Ce dispositif pourrait facilement être mis en place dans le cas de l'objet choisi, mais serait difficilement adaptable à l'œuvre d'art cinétique ou à d'autres objets en mouvement.

La seconde option, bien que moins précise, car dépendante de la fréquence de prise de vue, permettrait de s'affranchir du son de la vidéo et d'adapter la méthode à d'autres objets en mouvement. Des tests effectués par nos soins ont permis de déterminer si une fréquence de 30fps était suffisante pour identifier correctement un instant dans un mouvement à l'aide d'un repère visuel. Ces tests se sont avérés positifs.

Pour toutes ces raisons, nous avons fait le choix d'un repère visuel. Il a été décidé d'utiliser des témoins de papier de forme triangulaire. Ces témoins sont composés de quatre repères fixes attachés au support et d'un repère mobile attaché à l'objet en mouvement (voir Figure 14). Pointés vers l'intérieur du dispositif et à plusieurs endroits sur le support pour éviter les occlusions, ces témoins vont nous permettre d'identifier des instants de passages depuis plusieurs points de vue. Ainsi, une fois l'ensemble du mouvement de l'objet capturé en vidéo, ces dernières seront synchronisées en identifiant les moments de passage du témoins triangulaire mobile devant les témoins fixes. De cette manière, il sera possible d'extraire une image depuis chacun des points de vue, et ce à tout instant du mouvement.

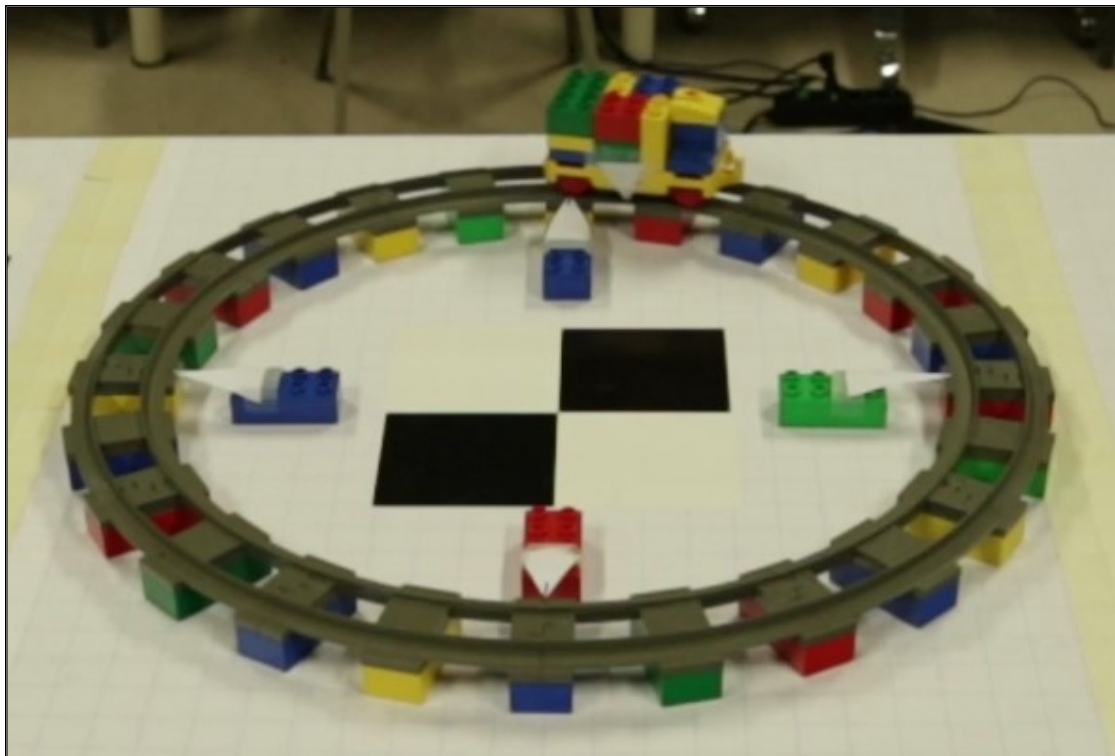


Figure 14 : Disposition des témoins de papier sur le support et sur l'objet en mouvement.

4.2.2 Traitement des données

Le traitement des données peut être divisé en deux étapes : la première est le découpage temporel des vidéos et la seconde est la reconstitution 3D de chaque instant du mouvement.

- **Découpage temporel**

Une fois les vidéos acquises depuis chacun des points de vue, il est nécessaire de les découper en images avant toute reconstitution 3D. Afin de discrétiser au minimum notre mouvement, il convient de découper nos vidéos à la fréquence maximale possible, soit 30 images par seconde. Beaucoup de logiciels permettent ce type de découpage. Nous utiliserons pour cela le logiciel CONTEXT CAPTURE. Le résultat de cette étape est un ensemble de photos correspondant à une discrétisation temporelle du mouvement depuis chacun des points de vue.

C'est à cette étape que les témoins de papier permettront la synchronisation des vidéos. Le but va donc être de retrouver, dans chacun des points de vue, les images correspondant à un même passage de témoin. Quand les images correspondantes au même instant du mouvement seront retrouvées, on pourra alors sélectionner le nombre d'images désirées temporellement en amont et en aval de ce témoin en fonction de l'intervalle de temps que nous désirons représenter. Ainsi si nous voulons représenter 1 seconde du mouvement, nous sélectionnerons, dans chaque point de vue, 15 images avant et après le passage du témoin de papier, soit 30 images correspondant à 30 instants successifs.

- **Reconstitution 3D**

Il est maintenant possible de reconstituer en 3D chacun des instants du mouvement. Pour cela, nous utiliserons à nouveau le logiciel CONTEXT CAPTURE. Pour chaque instant, 16 images issues de chacun des points de vue seront importées dans le logiciel. Pour des raisons de temps de calcul, une aérotriangulation de contrôle sera effectuée. Les coordonnées des points de cinq cibles photogrammétriques calculées à l'aide du logiciel COVADIS permettront de contraindre notre reconstitution et d'évaluer si les précisions de cette dernière rencontrent nos prérequis. Ensuite, une limitation de l'emprise de la reconstruction sera appliquée pour correspondre uniquement à l'emprise du mouvement et la reconstitution d'un *mesh* 3D sera effectuée.

A la fin de cette étape, nous disposerons donc de 1 *mesh* 3D référencé dans le système de coordonnées local et correspondant à un instant précis du mouvement de l'objet. Si l'étape est répétée 30 fois avec des jeux d'images correspondant à des instants successifs, nous disposerons de 30 *meshs* 3D référencés dans le même système de coordonnées local et correspondant à 30 instants successifs du mouvement de l'objet.

4.2.3 Rendu 3D

Pour rappel, l'objectif final de la méthode est de réaliser une animation 3D d'un objet en mouvement. Il existe de nombreux logiciels permettant de faire de l'animation sur base de *mesh* 3D. Les plus répandus sont développés par la société AUTODESK chez laquelle on retrouve des logiciels comme MAYA, MOTION BUILDER et 3DS MAX.

Sans expérience dans l'utilisation de logiciel d'animation 3D, une revue de la littérature disponible nous a permis d'identifier le logiciel 3DS MAX comme étant le plus simple d'utilisation. Ce logiciel nous permet également d'aboutir à l'objectif de recherche que nous avons fixé, c'est-à-dire la réalisation d'une animation 3D précise, référencée dans un système de coordonnées local.

Le logiciel 3DS MAX propose plusieurs modes de réalisation afin d'obtenir une animation 3D. Les deux modes les plus populaires parmi les développeurs d'animations sont le mode *pose to pose* et le mode *trajectory*. Ces deux modes seront utilisés dans ce travail de fin d'études.

Dans le cas de la méthode optique et multi-capteurs sans marqueurs placés sur l'objet, nous utiliserons la méthode *pose to pose*.

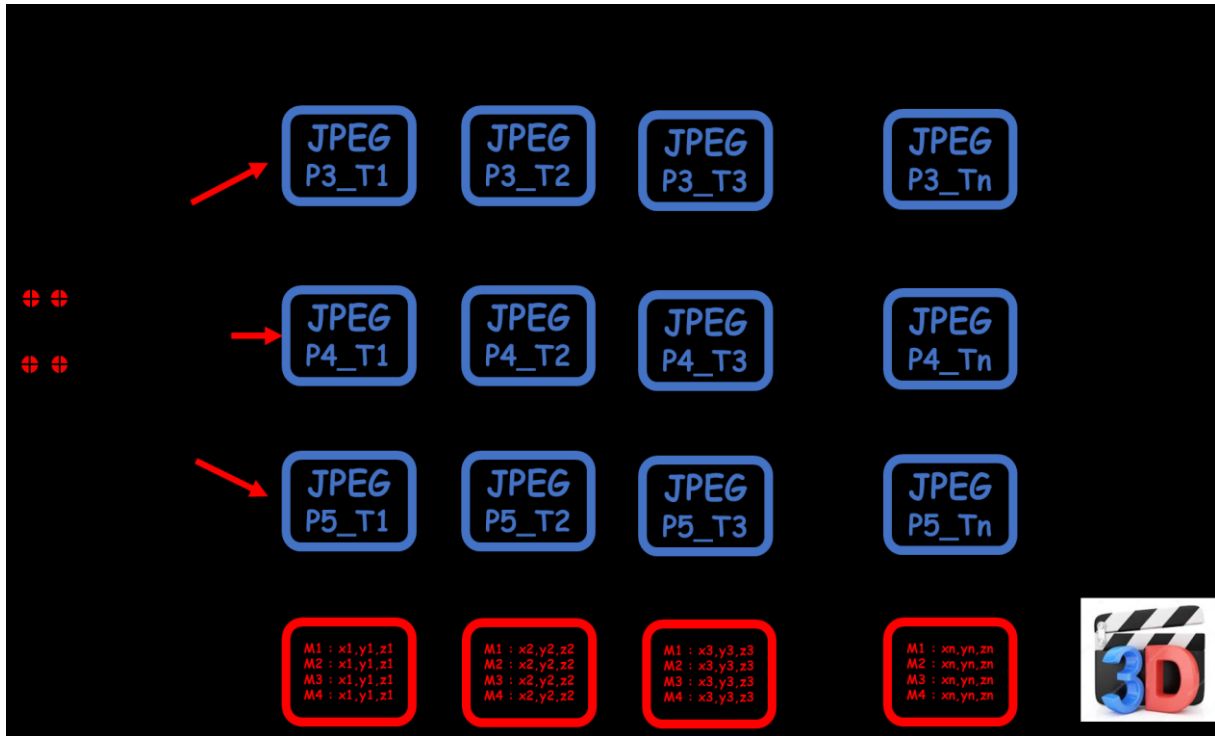
Comme son nom l'indique, le mode *pose to pose* est un mode utilisé en animation permettant de définir les poses clés (aussi appelées *keyframes*) du mouvement d'un objet ou d'un personnage. Ce mode interpole ensuite ces poses dans des cadres intermédiaires afin d'animer le sujet d'une pose à l'autre. On retrouve ce mode d'animation tant dans l'animation classique que dans l'animation 3D sur ordinateur.

Pour cette méthode, l'animation se fera en trois étapes. La première consiste en l'import des différents *meshs* 3D, chacun des *meshs* correspondant à un instant du mouvement que nous voulons représenter. La deuxième étape consistera à renseigner au logiciel chacune des poses qu'il doit considérer pour calculer l'animation. Chaque *mesh* correspondra donc à une *keyframe*. Et la troisième étape consistera à interpoler le mouvement entre les différentes poses et créer une animation 3D fluide de notre objet en mouvement.

Dans le cas de la méthode optique et multi-capteurs de capture de mouvement avec marqueurs placés sur l'objet, nous utiliserons le mode *trajectory* (voir point 4.3.3).

4.3 Avec marqueurs placés sur l'objet

Cette méthode, développée sur base de nos connaissances théoriques, comporte deux grandes étapes : la première étant la modélisation de l'objet et de ses marqueurs au repos, et la deuxième étant le suivi des marqueurs présents sur cet objet pendant le mouvement. Ainsi, s'il est possible, à l'aide des techniques photogrammétriques, de modéliser l'objet au repos et ensuite de connaître les positions des marqueurs placés sur cet objet au cours du temps, il est alors possible de modéliser la trajectoire des marqueurs et donc de l'objet en mouvement (voir Figure 15). Notons que cette méthode, contrairement à la méthode sans marqueurs développée précédemment, a pour hypothèse de départ que l'objet à modéliser ne se déforme pas dans le temps.



Avec :

M_n : position du marqueurs n

P_n : Position du point de vue n

T_n : Instant n du mouvement

Figure 15 : Schéma de la méthode optique et multi-capteurs pour capturer le mouvement avec marqueurs placés sur l'objet.

Sur base de ces suppositions théoriques, on peut établir précisément notre méthode optique et multi-capteurs de capture du mouvement avec marqueurs placés sur l'objet. Le développement de cette méthode peut également être divisé en trois étapes principales. Ces trois étapes sont l'acquisition du mouvement, le traitement des données et le rendu en 3D.

4.3.1 Acquisition du mouvement

Hormis les questions habituelles de référencement dans un système local de coordonnées, de disposition des capteurs et de synchronisation des prises de vues qu'il est nécessaire de se poser avant toute acquisition, il est également judicieux de trouver les réponses aux questions suivantes. Comment modéliser l'objet au repos ? Et comment connaître les positions des marqueurs pendant le mouvement ?

- **Comment modéliser l'objet au repos ?**

Pour obtenir un modèle 3D de l'objet au repos par photogrammétrie, nous appliquerons les recommandations du logiciel CONTEXT CAPTURE afin de modéliser un objet en 3D. Ces recommandations sont les suivantes.

- Les images initiales permettant de reconstituer le modèle doivent faire en sorte que chaque partie de l'objet soit au moins photographier depuis 3 points de vue différents.
- Le recouvrement minimum entre deux prises de vues doit être de deux tiers d'image et l'angle entre les prises de vues doit être de maximum 15 degrés.
- La longueur focale utilisée pour le capteur lors des prises de vue doit toujours être identique. Notons que le logiciel peut automatiquement estimer les distorsions liées à la lentille de l'objectif.
- Pour obtenir le modèle le plus précis possible, les photos seront prises à la plus grande résolution.
- L'exposition des prises de vue doit permettre d'éviter le flou de bougée, ainsi que le bruit de mesure et la sur/sous-exposition du sujet qui peuvent altérer sérieusement la construction du modèle 3D.
- Bien que n'affectant pas la reconstruction 3D du modèle, une lumière ambiante et constante sera, autant que possible, utilisée pour la prise de vue ; les variations de lumières et les lumières directes ayant des effets d'ombres non-esthétiques sur la texture du modèle.

Pour des objets simples, respecter ces recommandations représente un ensemble de 30 à 50 images à une résolution de 6016 x 4000 pixels, effectuées avec le CANON 5D MARK III sur lequel l'objectif CANON 24 mm 1 : 2,8 sera monté. Le produit de la reconstitution sera un *mesh* 3D sur lequel les marqueurs seront visibles et facilement identifiables.

- **Comment référencer notre modèle dans un système local ?**

Pour les mêmes raisons que pour la méthode précédente, il est essentiel de référencer notre modèle 3D dans un système local de coordonnées. La méthode proposée au point 4.2.1 est une méthode de référencement connue des topographes familiers de la photogrammétrie. A ce point, l'emploi de cette méthode nous permet de limiter nos erreurs et de nous concentrer sur le problème de synchronisation de prises de vue dont nous essayons d'évaluer la précision. Comme nous l'avons déjà mentionné, il s'agit d'une méthode non automatisable, car elle nécessite de pointer les points de contrôle manuellement avant d'effectuer l'aérotiangulation.

Dans le but de minimiser les interventions en post traitement, nous allons donc tester pour notre méthode de capture multi-capteurs avec marqueurs posés sur l'objet, une méthode de référencement automatique dans un système de coordonnées local. Conscient que le test d'une nouvelle méthode peut introduire des erreurs ou des niveaux de précisions trop faibles dans notre reconstitution 3D, nous essayerons d'évaluer les précisions qu'il est possible d'atteindre. Pour cela, il faut disposer de marqueurs que le logiciel de reconstitution peut automatiquement détecter et géoréférencer. Dans cette optique, nous allons utiliser une fonction du logiciel CONTEXT CAPTURE qui permet d'éditer des QR Codes. Le logiciel est capable d'éditer à la fois des QR Codes géoréférencés et des QR Codes simples.

QR Codes géoréférencés :

Pour référencer notre modèle dans un système local, nous utiliserons des QR Codes géoréférencés. Un QR Code géoréférencé est un QR Code contenant les coordonnées X, Y et Z de son centre dans le système de coordonnées désiré. Si le QR Code a été renseigné au logiciel comme étant un point de contrôle, il est alors possible d'implanter ce dernier au bon endroit sur le site avant d'effectuer les prises de vue. Lors de l'aérotriangulation des images, le logiciel pourra alors automatiquement repérer le QR Code, détecter son centre et lui assigner le statut de point de contrôle avec les coordonnées désirées. Cette méthode permet de s'affranchir du post-traitement chronophage visant à cliquer à la main, avant reconstitution, l'ensemble des points de contrôle implantés. Néanmoins, la précision de cette méthode n'est pas connue et dépend de l'algorithme utilisé pour lire et détecter le centre de QR Codes. Nous tenterons donc de déterminer la précision qu'il est possible d'atteindre avec le logiciel CONTEXT CAPTURE en utilisant des QR Codes géoréférencés. Un exemple de QR Codes géoréférencé est présent à la Figure 16.

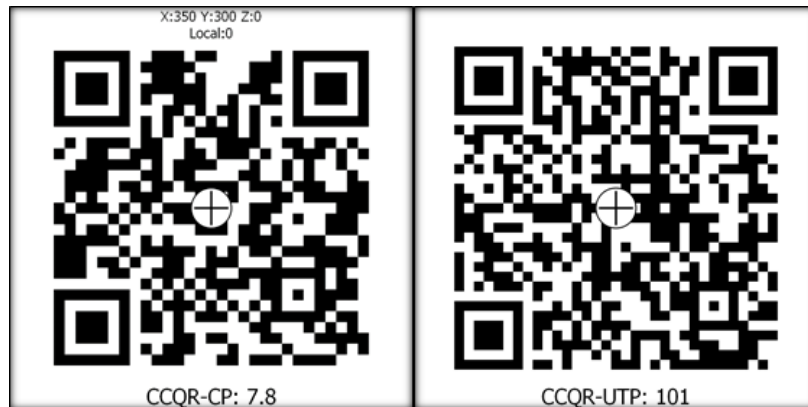


Figure 16 : Exemple de QR Code géoréférencé (A gauche) et de QR Code simple (à droite). Le QR Code géoréférencé contient comme information les coordonnées de son centre (X : 350 Y:300 Z:0), le système de coordonnées dans lequel il s'inscrit (0 pour local), son type (CCQR-CP pour Control point) et son nom (7.8). Le QR Code simple contient comme information son type (CCDR-UTP pour User Tie Point) et son nom (101).

Pour nous géoréférencer notre modèle dans un système de coordonnée local, un nombre minimum de QR Codes détectés est requis. Ainsi, il est possible de reconstituer un modèle photogrammétrique géoréférencé à partir d'un minimum de 3 points de contrôle. Néanmoins, pour être plus précis, il convient de pouvoir estimer les erreurs et effectuer des corrections en compensant les résidus par moindres carrés. Pour cela, un minimum de 4 points de contrôles détectés est nécessaire, la précision augmentant avec le nombre de points de contrôle détectés.

Après tests, nous avons identifié que l'ensemble des QR Codes présents sur la scène n'était pas toujours détecté. Ainsi, le nombre de QR Codes détectés varie en fonction des paramètres de prises de vue (exposition, ouverture, vitesse d'obturation, etc.), de la résolution de l'image, de l'inclinaison de la prise de vue et de l'algorithme de détection. De plus, l'objet en mouvement peut obstruer un ou plusieurs points de contrôle lors de son passage. Pour être certains de détecter un minimum de 4 points de contrôle à chaque instant du mouvement et éviter des occlusions par l'objet, on maximisera le nombre de points de contrôle présents sur la scène lors des prises de vue.

La disposition idéale est donc à un lit de QR Codes proportionnels à la taille de l'objet et de son mouvement. Nous choisissons un lit de 11 x 11 QR Codes carrés de 0,050 m de côté, ce qui correspond à un plan géoréférencé localement de 0,500 m x 0,500 m. Pour s'assurer que les centres des QR Codes correspondent aux coordonnées tridimensionnelles qu'ils renseignent. Le centre de ces derniers a été implanté précisément sur un plan au format A1 et à l'échelle 1:1 à l'aide du logiciel AUTOCAD MAP 3D (voir Figure 17). C'est donc sur ce lit que prendra place l'objet en mouvement.

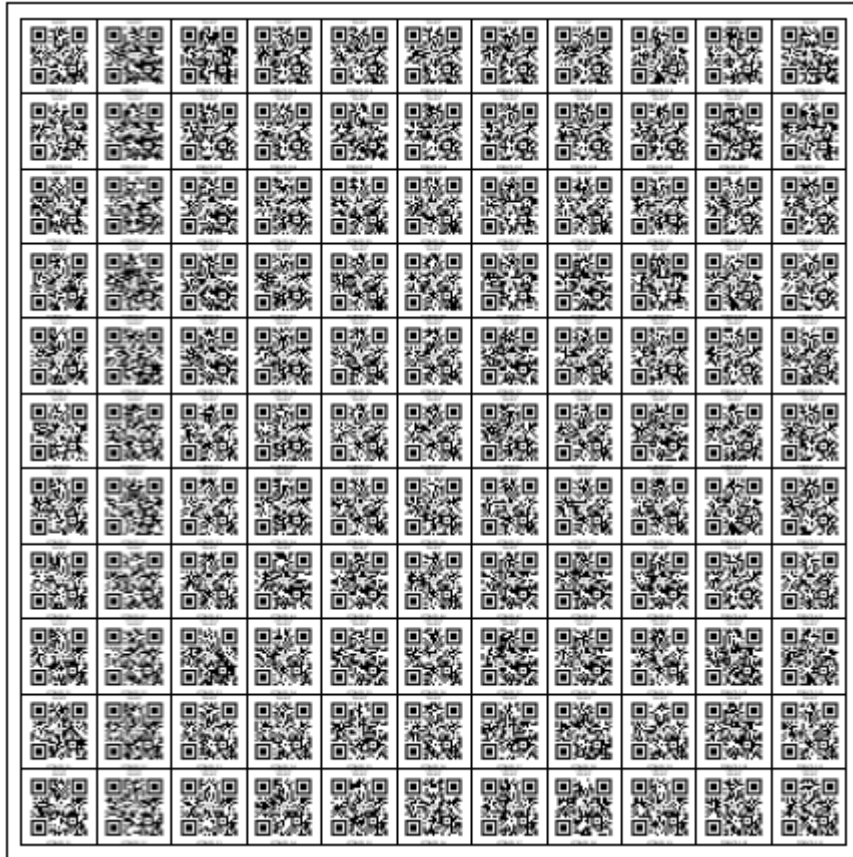


Figure 17 : Plan géoréférencé au format A1 et à l'échelle 1:1 de 11x11 QR Codes de 5 cm de cotés.

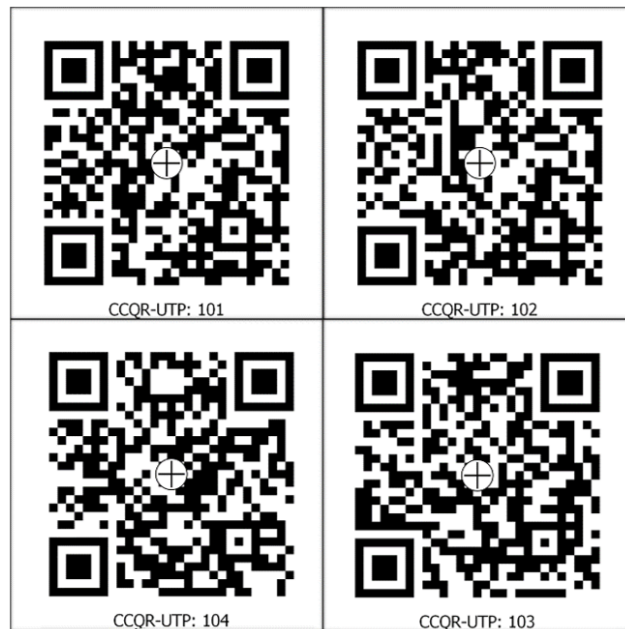


Figure 18 : Disposition des 4 QR Codes simples servant de marqueurs sur l'objet en mouvement

QR codes simples :

L'autre type de QR Codes éditables par le logiciel CONTEXT CAPTURE sont des QR Codes simples (non-géoréférencés) qui peuvent contenir une information sur leur type et sur leur code. Ces types de QR Codes sont considérés par le logiciel comme des *Users Tie Point*. Un exemple de QR Codes simple est présent à la Figure 16. Si le QR Code est placé sur l'objet lors des prises de vues, il peut être détecté par le logiciel et être identifié comme *Tie Point*. Lors de l'aérotriangulation, si suffisamment de QR Codes géoréférencés ont été détectés, les coordonnées des centres des QR Codes simples pourront alors être calculées par le logiciel avec une précision que nous tenterons d'évaluer.

Devant développer une méthode qui n'altère pas l'objet de façon définitive, les QR Codes simples sont une alternative acceptable car ils peuvent servir de marqueurs non-définitifs. Nous placerons donc sur l'objet des QR Codes simples qu'il sera possible de retirer par la suite. Pour déterminer avec précision la position de l'objet en mouvement, il faut également un minimum de 4 marqueurs détectés pour déterminer la position de l'objet par moindres carrés. Sur l'objet, on placera donc 4 QR Codes carrés de 0,030 m * 0,030 m de côté. Chacun de ces QR Codes contiendra l'information de son code permettant de l'identifier en post traitement. La détection des QR Codes simples dépend des mêmes facteurs que pour les QR Codes géoréférencés. La répartition des QR Codes simples servant de marqueurs sur l'objet est visible à la Figure 18.

• Comment disposer les capteurs ?

Utilisant le même capteur et les mêmes objectifs que pour la méthode d'acquisition précédente, nous garderons la même disposition et le même nombre de capteurs, c'est-à-dire 16 capteurs répartis tous les 25 gons autour de l'objet. Néanmoins, afin de favoriser la détection des QR Codes qui seront placés à l'horizontale, nous effectuerons les prises de vue à une distance plus petite du centre (0,050 m) et avec une hauteur par rapport à l'objet plus importante. Sachant que la détection des QR Codes dépend de l'inclinaison des prises de vues (α et β), nous pouvons voir l'intérêt d'une modification de la hauteur du capteur sur la Figure 19.

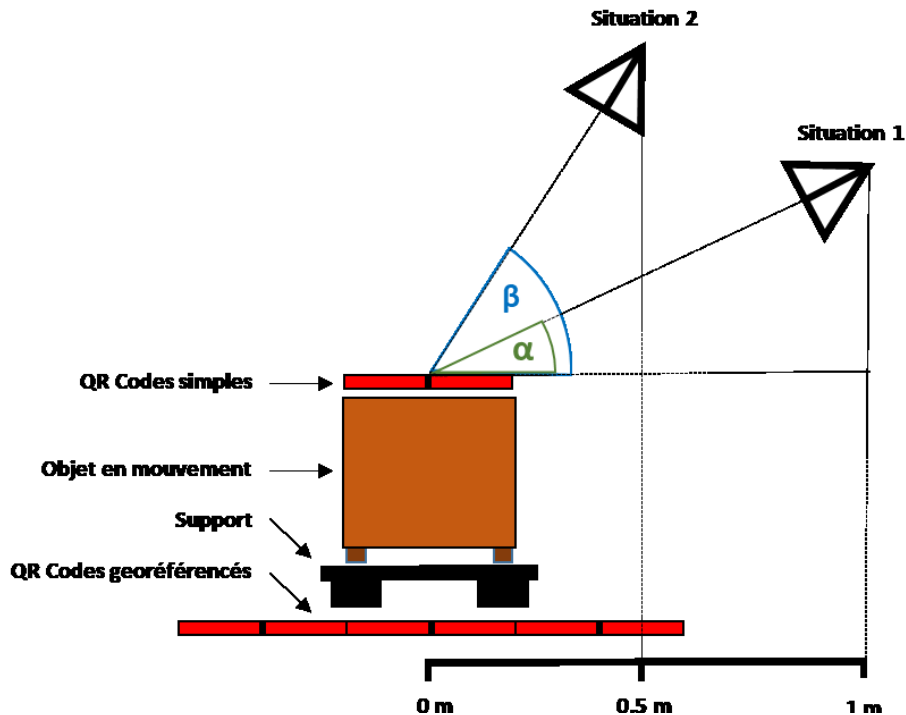


Figure 19 : Illustration de l'influence de la hauteur du capteur pour la détection des QR Codes.

- **Comment synchroniser les prises de vue ?**

L'objectif de cette méthode est d'évaluer la précision de la détection automatique de points de contrôle à l'aide de QR Codes dans une perspective de capture du mouvement.

Pour cela, les prises de vue seront systématiquement effectuées sur l'objet à l'arrêt. De cette manière, nous n'introduirons pas d'erreurs de synchronisation des prises de vue ou d'erreurs dues au flou de bougée. L'objet sera donc arrêté dans son mouvement à intervalles réguliers pour effectuer les prises de vue depuis chacune des 16 positions de capteur. Afin de ne pas perdre de temps dans l'acquisition des données, nous répéterons cette opération 7 fois de manière à pouvoir calculer les positions des marqueurs à 7 instants différents du mouvement. Une limitation à 7 poses est suffisante pour déterminer, par la suite, si la trajectoire de l'objet peut être calculée sur base de position déterminée par des QR Codes.

4.3.2 Traitement des données

Comme mentionné dans le point 1.1.1, les positions des marqueurs seront calculées automatiquement par le logiciel CONTEXT CAPTURE lors de l'aérottriangulation. Connaissant les positions relatives des QR Codes placés sur l'objet, nous pourrons, si suffisamment de QR Codes sont détectés, ajuster les coordonnées par moindres carrés et déterminer la précision sur ces dernières.

Une fois les coordonnées calculées et ajustées, nous interpolerons ces derniers afin d'établir la trajectoire de chacun des marqueurs détectés. Si la trajectoire des marqueurs est connue, il sera alors possible d'animer l'objet en 3D. La méthode utilisée pour le calcul et l'interpolation de la trajectoire, ainsi que les résultats de cette méthode seront présentés au point 5.3.

4.3.3 Rendu 3D

Toujours avec le logiciel 3DS MAX, nous utiliserons cette fois-ci le mode *trajectory*. Comme son nom l'indique, le mode *trajectory* est un mode utilisé en animation permettant d'animer un objet sur base d'une trajectoire ou d'une suite de coordonnées à suivre par certains points clés présents sur l'objet à animer.

Pour cette méthode, l'animation se fera en trois étapes. La première consiste en l'import des coordonnées de la trajectoire des points clés de l'objet, chacune des trajectoires étant propre à son marqueur. La deuxième étape consistera à importer le *mesh* 3D de l'objet fixe. Et la troisième étape consistera à rattacher les marqueurs à leur trajectoire respective afin de créer une animation 3D fluide de notre objet en mouvement.

4.5 Un système de capture de mouvement basé sur la lasergrammétrie

Voulant également connaître les perspectives de la lasergrammétrie en matière de capture du mouvement et ainsi répondre à notre dernière hypothèse de recherche, il a été décidé de mettre en place une méthode de capture du mouvement à l'aide du scanner laser. L'objectif dans ce cas-ci ne sera pas d'aller jusqu'à l'obtention d'un rendu 3D de l'objet en mouvement, mais de tirer des conclusions sur les possibilités offertes par le scanner laser en la matière. Nous chercherons donc à extraire d'une scène uniquement le mouvement de l'objet. Pour cela, nous avons développé une méthode en deux parties composées tout d'abord de l'acquisition de l'objet en mouvement, et ensuite de l'extraction du mouvement en post-traitement.

4.5.1 Acquisition du mouvement

Afin d'extraire les points appartenant au mouvement de l'objet dans un nuage de points, il a été décidé de scanner notre scène à deux reprises depuis 4 points de vue différents. A la première reprise, uniquement le support du mouvement sera scanné et, à la deuxième reprise, l'objet sera mis en mouvement sur le support et le tout sera scanné. Sur base de ces 2 ensembles de 4 scans, nous assemblerons deux nuages de points différents dans le même système de coordonnées. Cela permettra, en superposant les nuages de points en post-traitement, d'identifier les points propres au mouvement de l'objet sur base de la distance entre ces derniers.

- **Disposition des capteurs**

Pour les mêmes raisons qu'au point 4.2.1, il est nécessaire d'optimiser l'espace de prise de vue tout en étant certain de garder une résolution inférieure à 0,001 m au sein du nuage de points final.

Pour cela, il faut déterminer 3 paramètres : la résolution optimale de notre scanner, la distance entre le centre du dispositif et l'appareil de levé, ainsi que l'RMS (*Root Mean Square*) erreur engendrée sur base des caractéristiques techniques (erreur sur les distances et erreurs angulaires) du scanner LEICA P30.

Les caractéristiques techniques sont les suivantes :

Erreur sur distance (eD): **0,012 + 10 ppm mètres**

Erreur angulaire horizontale (eHZ) : **0,0025 gons**

Erreur angulaire verticale (eV): **0,0025 gons**

Le scanner est capable de lever à différentes résolutions de point. Ces résolutions sont présentées à l'utilisateur comme des résolutions à 10 m. Pour chaque résolution à 10 m, on peut calculer la distance scanner-objet nécessaire pour une résolution de 0,001 m dans notre nuage de points final.

Soit :

Résolution à 10 m (m): **Res_10**

Resolution à la distance D (m): **Res(D)**

Distance entre le scanner et l'objet (m): **D**

On peut écrire :

$$Res(D) = \left(\frac{Res_{10}}{10} \right) * D$$

On peut également déterminer l'erreur sur les distances due à *eHZ*, *eV* et *eD*, ainsi que déduire les RMS erreurs résultantes.

Soit :

erreur sur distance lié à eHZ (m) : $eHZ(D)$

erreur sur distance lié à eV (m) : $eV(D)$

erreur de linéarité (m) : $elin(D)$

On peut écrire :

$$eHZ(D) = \text{atan}(\text{rad}(eHZ)) * D$$

$$eV(D) = \text{atan}(\text{rad}(eV)) * D$$

$$elin(D) = 0,012 + 0,0010 * \frac{D}{1000}$$

Avant de calculer la RMS erreur correspondante, il faut également pouvoir quantifier le bruit de mesure. La fonction suivante permet de représenter le bruit de mesure en fonction de la distance scanneur-objet. Cette fonction est établie sur base d'un échantillon de mesures distance-bruit (voir Annexe 2), la moyenne des mesures distances-bruits sur trois types de surfaces (noire, grise et blanche) a été effectuée et une régression polynomiale de degrés 2 a été établie. On obtient donc la fonction de régression suivante comme fonction du bruit de mesure :

Bruit de mesure moyen en fonction de la distance (m) : $B(D)$

$$B(D) = 0,0007046 * D^2 - 0,002234 * D + 0,5737$$

Enfin :

Root Mean Square error en fonction de la distance : $RMS(D)$

$$RMS(D) = \sqrt{eHZ(D)^2 + eV(D)^2 + elin(D)^2 + B(D)^2}$$

Ainsi, via un script MATLAB que nous avons développé (l'Annexe 3), nous avons établi le Tableau 1 présentant, pour une résolution sur objet de 0,001 m, la résolution à 10 m qu'il est nécessaire d'introduire dans le scanneur, ainsi que la distance scanneur-objet correspondante et la RMS résultante.

Tableau 1: Résolution à 10 m et distance scanneur-objet pour un résolution 0,001m, ainsi que la RMS erreur correspondante.

Résolution à 10 m	Distance scanneur-objet	RMS erreur
50,0 mm	0,1950 m	0,0013 m
25,0 mm	0,3950 m	0,0013 m
12,5 mm	0,7950 m	0,0014 m
6,3 mm	1,5850 m	0,0014 m
3,1mm	3,2250 m	0,0014 m
1,6 mm	6,2450 m	0,0014 m
0,8 mm	12,4950 m	0,0016 m

Il a donc été choisi d'effectuer une succession de 4 scans entourant l'objet avec un écart angulaire de 100 gons. Les scans seront éloignés de l'objet d'une distance de 2 m et la résolution choisie sera de 3,1 mm à 10m. Le choix de ces paramètres nous assure une résolution inférieure à 0,001 m dans notre nuage de points final. Un schéma du dispositif a été tracé (voir Figure 20) et suivant la même méthode que décrite au point 4.2.1, les positions du scanneur ont pu être implantées au sol.

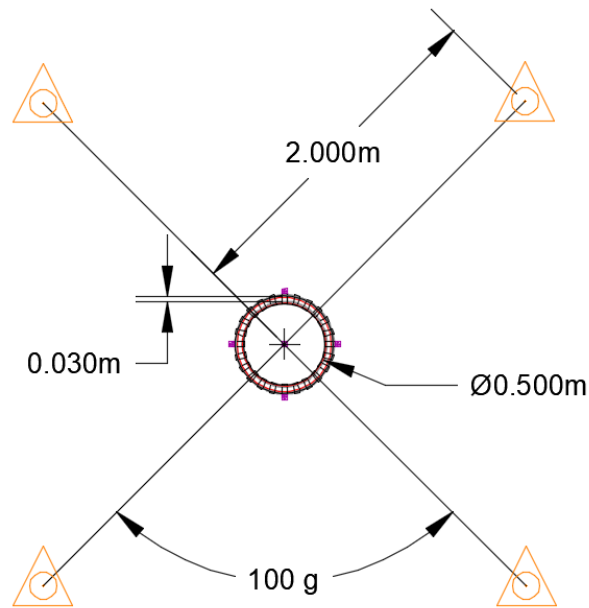


Figure 20 : Schéma d'implantation des positions du scanneur.

- **Référencement dans un système local et assemblage des scans**

Pour assembler les scans, les centres d'un ensemble de cibles disposées dans toute la pièce seront levés à la station totale dans le système de coordonnées local défini au point 4.2.1.

Les 4 scans du support et les 4 scans de l'objet en mouvement seront assemblés séparément de manière à obtenir deux nuages de points distincts, superposables dans le même système de coordonnées.

4.5.2 Extraction du mouvement

Afin d'extraire l'objet en mouvement, les deux nuages de point seront comparés dans le logiciel CLOUDCOMPARE. Une analyse des distances entre les nuages de points sera effectuée afin de séparer les points du support, les points propres au mouvement et les points correspondant au bruit de mesure. Les détails et les résultats de cette analyse seront présentés au point 5.4.

5 APPLICATIONS

Après avoir développé théoriquement trois méthodes d'acquisition, nous les mettrons en pratique dans cette partie. Nous mettrons en évidence leurs résultats, ainsi que leurs avantages et inconvénients. Pour chacune d'entre elles, les résultats seront décrits avec la même hiérarchie qu'au point 4. Nous commencerons donc par l'acquisition du mouvement, ensuite le traitement des données et nous finirons avec le rendu 3D.

5.1 Objet en mouvement

Avant d'entrer dans les détails des résultats des tests, il est nécessaire de critiquer l'objet choisi pour l'application de nos méthodes. En effet, l'objet sélectionné pour les tests présente certains points négatifs qui ont été mis en lumière lors de l'acquisition et du traitement des données.

Premièrement, l'objet utilisé possède une surface plastique engendrant de nombreux reflets avec la lumière ambiante. Cette surface légèrement réfléchissante perturbe la qualité des prises de vue et induit du bruit de mesure, et donc des erreurs non quantifiables lors de la reconstruction 3D. Le problème du bruit se présente tant pour la reconstruction photogrammétrique que pour la reconstitution à l'aide du scanneur laser, avec tout de même plus d'importance pour la reconstruction photogrammétrique. La présence d'une surface dentelée, propre à la marque LEGO, au-dessus de l'objet accroît également le bruit de mesure. Le volume de la locomotive n'étant pas un parallélépipède parfait, il est difficile d'estimer le bruit de mesure en comparant le modèle avec sa forme connue. Obtenir une reproduction précise, tant au repos qu'en mouvement, est donc difficile avec cet objet. La Figure 21 illustre le problème du bruit de mesure sur le résultat de la reconstitution 3D de l'objet en mouvement.

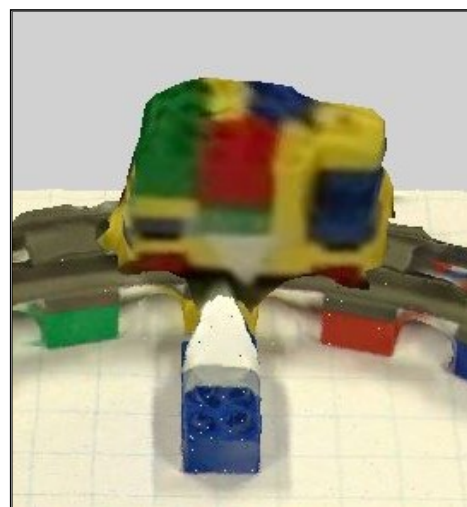


Figure 21 : Bruit de mesure sur reconstitution 3D

Deuxièmement, l'hypothèse d'un support de rails en plastique fixe ne peut pas être prise en compte. En effet, lors de l'acquisition de l'objet en mouvement, il a visuellement été constaté que, à la suite de l'effort continu exercé par l'objet en mouvement sur les rails, ces derniers avaient tendance à dévier légèrement de leur position initiale bien qu'elles aient été fixées au sol à plusieurs endroits. Au maximum, une déviation de 3 mm entre la position initiale, au début de l'acquisition, et la position finale, à la fin de l'acquisition, a été mesurée. Cette déviation, difficilement mesurable sur l'intégralité du support à cause du bruit de mesure, devra être prise en compte lors de la discussion des résultats des différentes méthodes.

Ainsi, des avantages cités au point 4.1, il ne reste que le caractère économique du dispositif qui peut justifier son emploi pour nos tests. Afin d'accroître la précision des tests et de déterminer les possibilités offertes par les méthodes développées, l'objet utilisé aurait dû être de couleur mate et de forme beaucoup plus simple. Afin d'assurer la reproductibilité du mouvement, le support du mouvement utilisé aurait également dû être de couleur mate et complètement fixe au sol. Un exemple de dispositif idéal est un chariot sur rails semblable à ceux utilisés en laboratoire de physique mécanique pour évaluer les forces.

Ne possédant pas de dispositifs respectant l'ensemble de ces critères, nous avons néanmoins choisis de continuer notre expérimentation avec notre dispositif initial. Nous tiendrons bien sûr compte des éléments que nous venons d'énoncer dans le point 6 « Discussion ».

5.2 Sans marqueurs placés sur l'objet

5.2.1 Acquisition du mouvement

Pour ce qui est de l'acquisition du mouvement pour la méthode de capture optique et multi-capteurs du mouvement sans marqueurs placés sur l'objet, nous décrivons les résultats de la procédure de référencement dans un système local et de détermination des coordonnées des cibles photogrammétriques.

- **Référencement et détermination des coordonnées des cibles.**

La procédure décrite au point 4.2.1 a permis de déterminer les coordonnées de la station totale et du repère dans le système de coordonnées local à l'aide d'un nuage de points assemblé depuis 4 scans distincts. L'assemblage des scans à l'aide des cibles réparties dans l'ensemble de la pièce a permis d'obtenir un nuage de points d'une densité inférieure à 1 point par millimètre au centre du dispositif.

Connaissant les précisions (voir point 4.2.1) et la résolution utilisée du scanner laser (3.1mm à 10m), nous pouvons, étant donné la disposition et le recouvrement entre les positions du scanner, estimer une RMS erreur sur nos coordonnées inférieures à 0,0014 m. Ces coordonnées ont été utilisées pour mettre le tachéomètre en station, ainsi que pour implanter les positions des capteurs et le centre du support de l'objet en mouvement.

Ainsi, le support et l'objet en mouvement ont pu être montés au centre du dispositif et 5 cibles photogrammétriques ont été placées autour du dispositif (voir Figure 22). Les centres de ces cibles, respectivement 11, 12, 13, 14 et 15 ont été levés à la station totale et leurs coordonnées calculées dans COVADIS. Les coordonnées des cibles dans le système de coordonnées local sont reprises dans le Tableau 2.

Tableau 2: Coordonnées des cibles photogrammétriques dans le système de coordonnées local.

N° Cibles	X	Y	Z
11	14.508	8.6008	1.1334
12	14.0612	8.6114	1.1363
13	14.0622	9.5866	1.1375
14	14.5322	9.5780	1.1370
15	14.2716	9.0821	1.1353

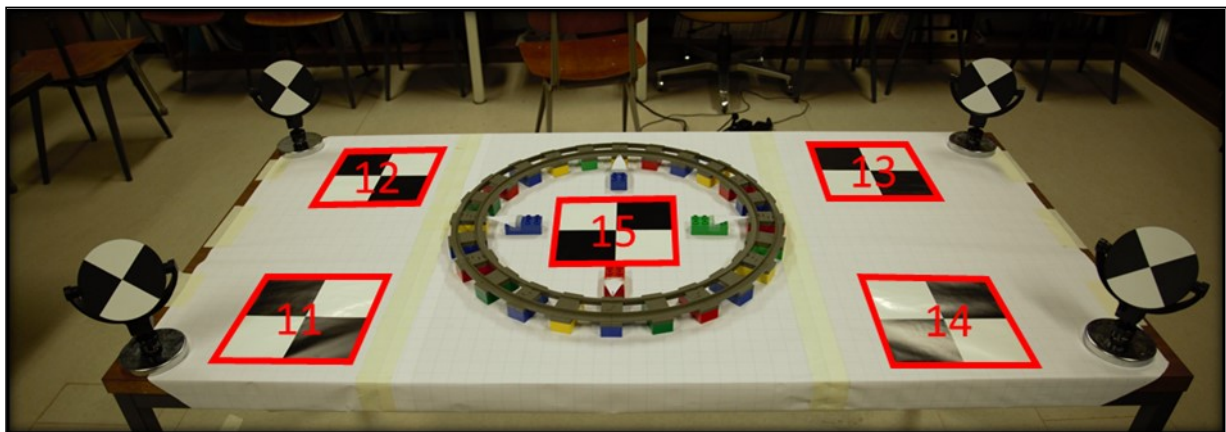


Figure 22 : Répartition des 5 cibles photogrammétriques autour du dispositif.

5.2.2 Traitement des données

Pour ce qui est du traitement des données, nous détaillerons les résultats de la synchronisation des prises de vue, de la préparation des images à la reconstitution 3D, ainsi que la construction des modèles 3D elle-même.

- **Synchronisation des prises de vue**

Le découpage des vidéos de 1980 x 1080 pixels s'est fait à l'aide du logiciel CONTEXT CAPTURE à une fréquence de 30 images par seconde.

Il est important de remarquer que la fréquence de 30 images par seconde des vidéos discrétise le mouvement étudié. En effet, entre deux images successives, l'objet se déplace et cette partie du mouvement ne peut pas être acquise par le capteur. Il est possible d'estimer visuellement la distance parcourue par l'objet en mouvement entre deux images successives de la vidéo. Dans notre cas, cette distance est de plus ou moins 0,005 m. Cette distance doit être prise en compte pour estimer les erreurs possibles de synchronisation.

En effet, comme mentionné au point 4.2.2, la synchronisation des prises de vues s'est faite à la main en parcourant les images afin d'identifier un même passage du témoin papier depuis chacun des 16 points de vue. Comme expliqué ci-dessus, la synchronisation ne permet pas d'identifier précisément le même moment vu que l'objet se déplace de 0,005 m entre chaque prise de vue. Cette distance parcourue entre les prises de vue induit une erreur de synchronisation qui, dans le cas le plus défavorable (voir Figure 23), peut induire une erreur de plus ou moins 0,0025 m lors de l'identification du passage de témoin. Cette erreur est une erreur fixe qui se propagera à l'ensemble des prises de vues servant à reconstituer nos modèles 3D. Les conséquences de ces erreurs de synchronisation devront être prises en considération avant de tirer des conclusions sur notre méthode optique et multi-capteurs de capture du mouvement.

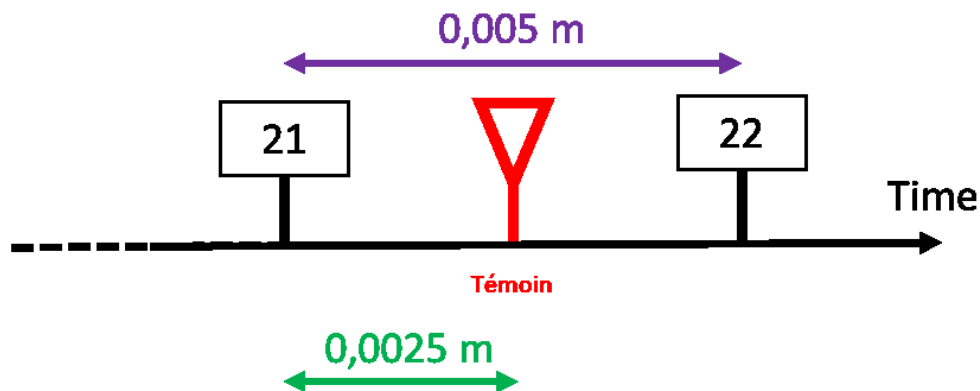


Figure 23 : Illustration du cas le plus défavorable lors de la synchronisation des prises de vue

- **Préparation des images à la reconstitution 3D**

Une fois le passage du témoin identifié pour chacun des points de vue, une sélection manuelle des 15 images suivantes et précédentes le passage de témoin a été réalisée de manière à isoler la même seconde du mouvement depuis chacun des 16 points de vue. Nous partons du principe que les résultats obtenus pour une seconde du mouvement suffisent à critiquer et tirer des conclusions sur la méthode utilisée.

Par la suite, il a fallu classer et regrouper les prises de vue pour les importer dans CONTEXT CAPTURE. Jusqu'alors groupées par point de vue, les images ont été classées par instant de manière à obtenir 30 dossiers de 16 images correspondant au même instant. Afin de gagner du temps dans le traitement des 480 images que ce traitement représente, ce classement a été fait de manière automatique à l'aide du script MATLAB présent à l'Annexe 4 (voir Figure 24).

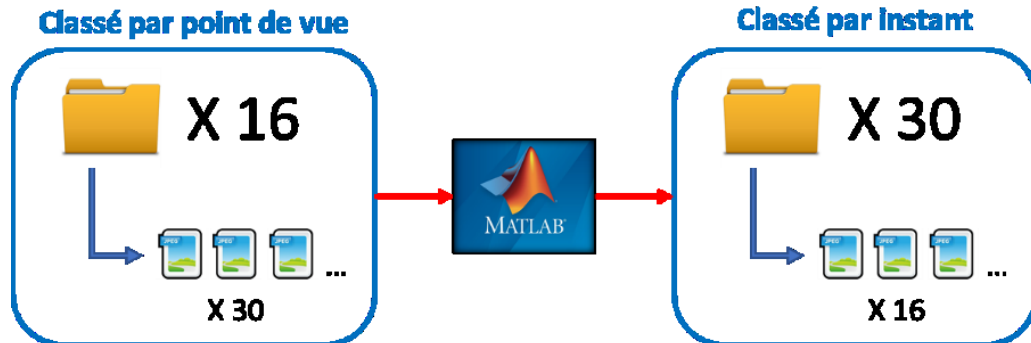


Figure 24 : Illustration des fichiers inputs et outputs du script MATLAB de classification.

Ce script permet ainsi d'automatiser une partie du traitement des données et permettrait de gagner du temps si nous désirions généraliser les traitements à l'ensemble du mouvement.

• Construction des modèles 3D

Disposant maintenant d'images de l'objet, prises sous 16 angles différents et ce pour 30 instants d'une seconde du mouvement, il est possible de reconstituer l'état du mouvement en 3D à chaque instant. Cette partie du traitement des données s'est faite en 5 étapes, et ce entièrement dans le logiciel CONTEXT CAPTURE.

- 1. Import des données inputs :** La première étape consiste à importer les 16 images correspondantes à l'instant que l'on veut modéliser. Les images sont importées dans un même *photogroup* au format JPEG. La taille du capteur (24 mm) et la focale utilisée (24 mm) sont renseignées au logiciel (voir Figure 25).

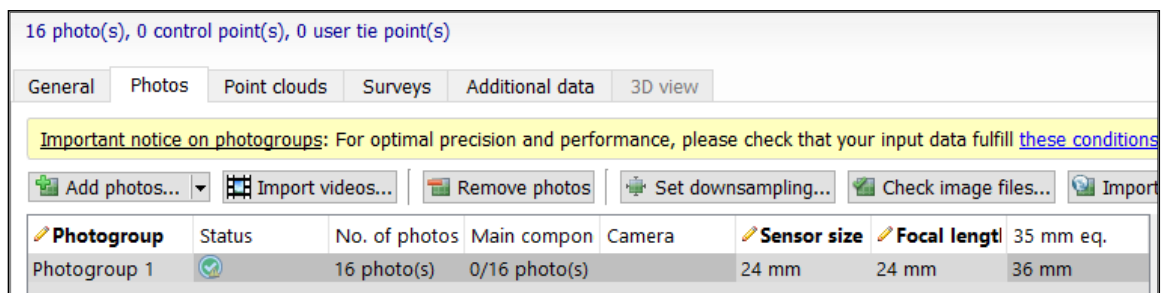


Figure 25 : Import des images inputs dans le logiciel CONTEXT CAPTURE

- 2. Aérotriangulation pour l'orientation relative :** La deuxième étape consiste à effectuer une première aérotriangulation afin de réaliser l'orientation relative des prises de vue. L'objectif d'une aérotriangulation est de déterminer, pour chaque photo, la position exacte du capteur au moment de la prise de vue. À l'aide de cette position, le logiciel pourra alors déterminer l'orientation des images, créer des modèles stéréoscopiques et lire correctement les coordonnées indiquées pour la reconstitution du modèle.
Dans les paramètres de l'aérotriangulation spécifiés par l'utilisateur, la sélection des images paires doit être renseignée comme devant se faire de manière séquentielle, car les images sont déjà ordonnées par point de vue.

Effectuer cette étape préparatoire permet de déterminer si les données en inputs sont suffisantes pour reconstituer le modèle et si l'erreur RMS de reprojection est suffisamment faible. Cette affirmation est contrôlée grâce au rapport d'aérotriangulation établie par le logiciel. Un exemple de rapport d'aérotriangulation pour une orientation relative est présent à l'Annexe 5.

Notons que l'ensemble des 30 aérotriangulations effectuées pour l'orientation relative des caméras ont pu être conduites avec une erreur de reprojection finale inférieure au pixel.

3. **Renseignement des points de contrôle :** Si l'aérotriangulation pour l'orientation relative s'effectue avec une erreur de reprojection correspondant à nos attentes. On peut alors renseigner l'emplacement et les coordonnées des cibles photogrammétriques sur une partie des données initiales. Ainsi, pour chaque lot de données correspondant à un instant du mouvement, le centre des 5 cibles et leurs coordonnées ont été renseignés au logiciel sur 4 images prises depuis des points de vue opposés (voir Figure 26). Notons que le choix de points de vue opposés permet de diminuer l'erreur RMS de distance au rayon, soit la distance moyenne entre la position du centre déterminé par moindres carrés et la position renseignée en cliquant sur les images.

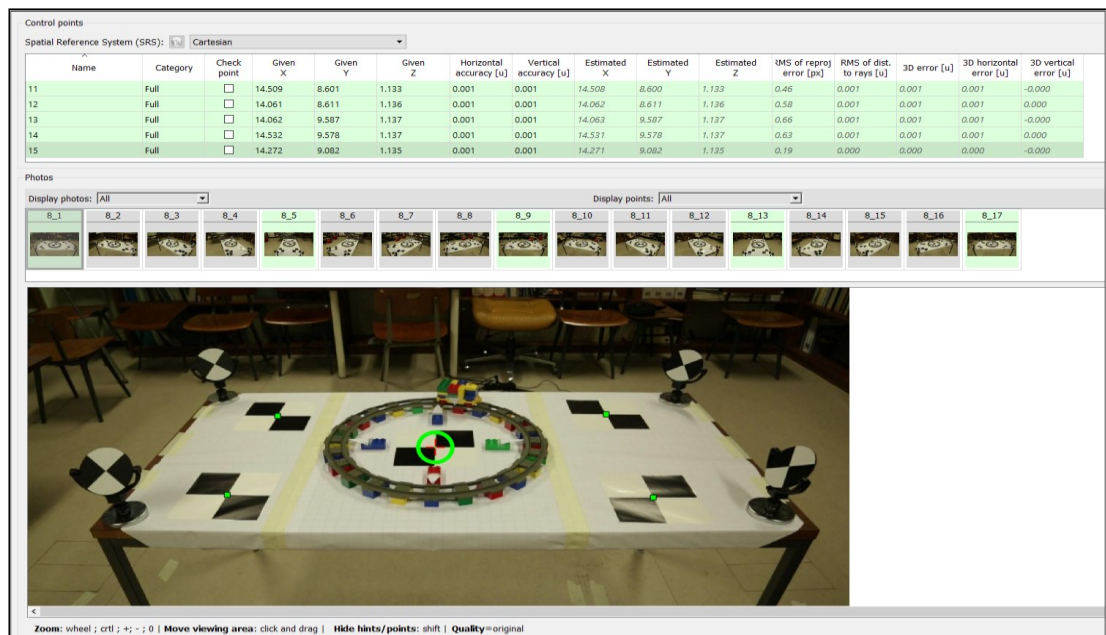


Figure 26 : Renseignement du centre des cibles et leurs coordonnées respectives au logiciel de reconstitution 3D.

4. **Aérotriangulation pour l'orientation absolue :** Une fois les cibles photogrammétriques renseignées, la quatrième étape consiste à effectuer une nouvelle aérotriangulation afin de réaliser l'orientation absolue des prises de vues dans le système de coordonnées local. Les précisions de cette aérotriangulation sont estimées à l'aide de l'erreur RMS de reprojection, exprimée en pixel, et de l'erreur RMS de distance au rayon, exprimée dans l'unité du système de coordonnées (le mètre). On peut consulter ces erreurs pour chaque point de contrôle dans le rapport d'aérotriangulation. Un exemple de rapport d'aérotriangulation pour l'orientation absolue est présent à l'Annexe 6. Les résultats de l'aérotriangulation pour chacun des 30 modèles présentent, pour chaque point homologue, des erreurs RMS de reprojection inférieure au pixel et au maximum une erreur RMS de distance au rayon de 0.002 m. Ainsi, les données en input ont été suffisantes pour l'orientation absolue de chacun des 30 modèles.

- 5. Production :** Une fois l'aérotriangulation pour l'orientation absolue effectuée et les précisions atteintes acceptables, la production du modèle photogrammétrique a pu être lancée. Avant cela, une limitation de la reconstitution du modèle a été entreprise à l'aide d'une « *bounding box* » ou « *cadre limite* ». Cette *bounding box* a permis de limiter l'ensemble de nos productions dans l'intervalle de coordonnées locales suivant (voir Figure 27) :

$$13,500\text{ m} < X < 15,000\text{ m} ; 8,000\text{ m} < Y < 10,000\text{ m} ; 0.750\text{ m} < Z < 1,500\text{ m}$$

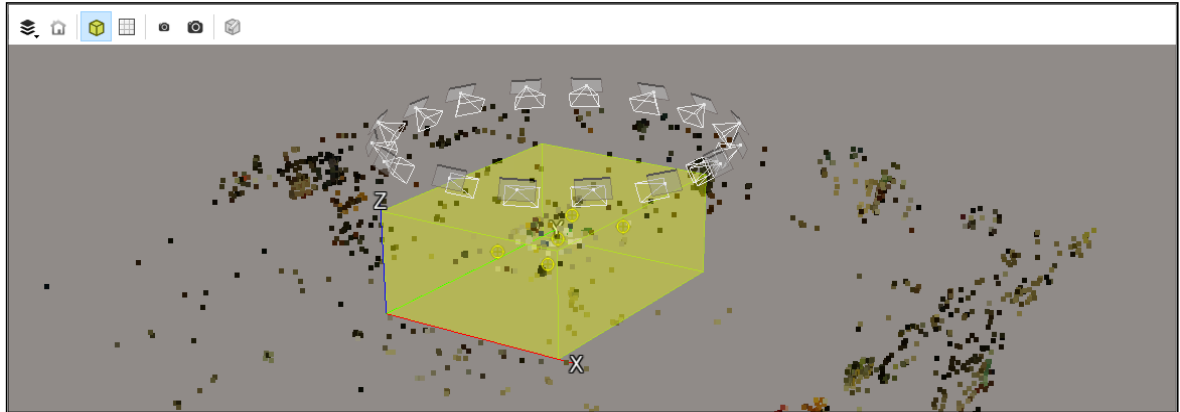


Figure 27 : Limitation de la reconstitution 3D à l'aide d'une bounding box

De cette manière, nous avons pu obtenir des productions 3D de même extension spatiale, limitées dans le même intervalle de mouvement et donc directement superposables.

Par la suite, deux productions 3D ont été réalisées pour chacun des instants, les deux sous forme de *mesh* 3D. Le premier a été fait au format de fichier OBJ, format d'échange permettant une visualisation du résultat dans CONTEXT CAPTURE, ainsi que l'import du résultat dans le logiciel de traitement de données 3D CLOUDCOMPARE (voir Figure 28). Le deuxième export a été fait au format FBX, format d'échange de la société AUTODESK qui nous a permis d'importer les résultats 3D dans le logiciel d'animation 3DS MAX.

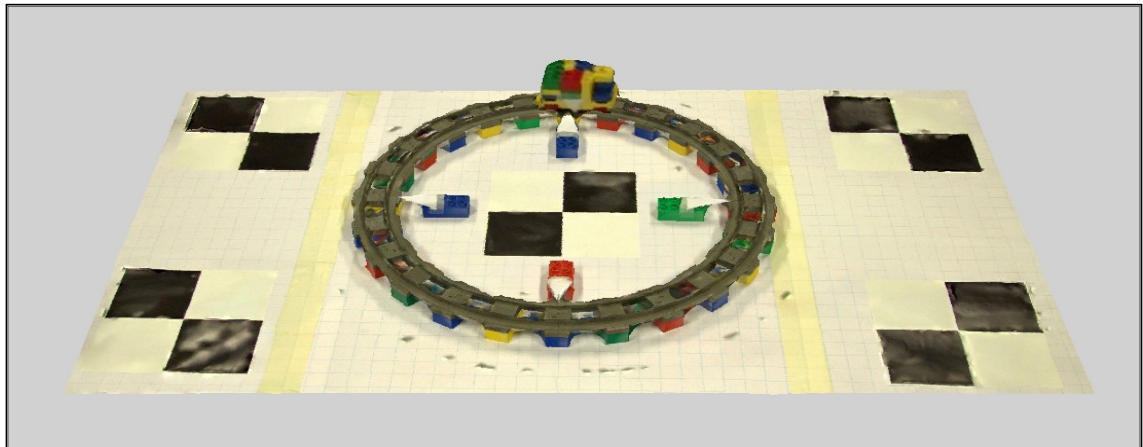


Figure 28 : Exemple de mesh 3D au format OBJ visualisé dans le logiciel CLOUDCOMPARE

A la fin de ces 5 étapes, nous avons donc, comme résultat, 30 *mesh* 3D au format OBJ et 30 *mesh* 3D au format FBX qui serviront de base au développement d'une animation 3D.

5.2.3 Rendu 3D

L'objectif initial était d'importer les 30 *mesh* 3D au format FBX dans le logiciel 3DS MAX afin de réaliser l'animation à l'aide du mode *pose to pose*. Cet objectif n'a pu être que partiellement atteint et cela pour deux raisons.

Premièrement, les 30 modèles correspondants aux 30 instants de la seconde que nous désirions modéliser ont pu être importés et superposés correctement dans le logiciel. En superposant ces modèles, on peut effectivement faire ressortir une première impression du mouvement de l'objet en 3D (voir Figure 29). Mais, comme expliqué au point 5.1, ces modèles présentent énormément de bruit dû à la nature des matériaux utilisés. Néanmoins, ces derniers restent parfaitement superposables et le détail d'une seconde du mouvement peut être observé en 3D.

Deuxièmement, la réalisation d'une animation 3D à l'aide du mode *pose to pose* a été impossible. En effet, le bruit de mesure et les erreurs de synchronisation font que les modèles 3D reconstitués ne sont pas comparables. On distingue bien visuellement l'objet, mais sa forme tridimensionnelle varie pour chacun d'entre eux et le logiciel est donc incapable de déduire que ces modèles représentent le même objet en mouvement.

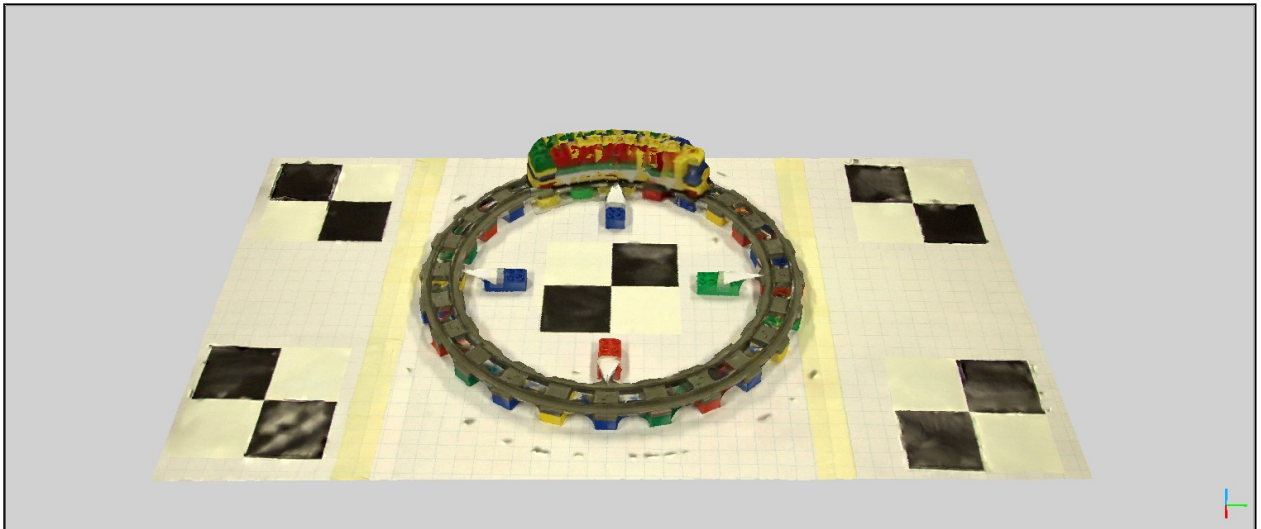


Figure 29 : Impression 3D d'une seconde du mouvement de l'objet par superposition des 30 modèles 3D correspondant à 30 instants successifs du mouvement.

5.3 Avec marqueurs placés sur l'objet

5.3.1 Acquisition du mouvement

Pour ce qui est de l'acquisition par la méthode de capture optique et multi-capteurs avec marqueurs placés sur l'objet, nous décrirons les résultats de la procédure de modélisation de l'objet au repos ainsi que celle du support. Nous terminerons par la description des résultats de la détermination des coordonnées des QR Codes simples placés sur l'objet en mouvement.

- **Modélisation de l'objet au repos**

Pour modéliser notre objet au repos, un ensemble de 67 photographies respectant les conditions de prises de vues énoncées au point 4.3.1 a été nécessaire. L'objet a été placé sur fond blanc, sans QR Codes référencés apparents. De ces images, une aérotriangulation relative a été réalisée dans le logiciel CONTEXT CAPTURE sans détection automatique des QR Codes.

Le rapport d'aérotriangulation indique une erreur RMS de reprojection de 0,84 pixel. Cette erreur étant inférieure au pixel, une reconstitution d'un *mesh 3D* a été effectuée. L'objet au repos a donc été modélisé.

- **Modélisation du support de l'objet**

Pour modéliser le support de l'objet, un ensemble de 87 photographies respectant les conditions de prises de vues énoncées au point 4.3.1 a été nécessaire. L'objet a été placé sur le support de 11x11 QR Codes mentionnés au point 4.3.1. De ces images, une aérotriangulation absolue a été réalisée dans le logiciel CONTEXT CAPTURE en activant la détection automatique des QR Codes.

Le rapport d'aérotriangulation indique la détection de 34 points de contrôle avec une erreur RMS de reprojection moyenne de 0,35 pixels et une erreur RMS de distance au rayon moyenne de 0,034 mm. Notons qu'un nettoyage manuel a été nécessaire pour supprimer les points de contrôle détectés avec une erreur RMS de reprojection supérieure au pixel, ainsi que ceux détectés sur uniquement 3 photographies ou moins. Ces erreurs étant acceptables étant donné la précision à laquelle nous voulons travailler, une reconstitution d'un *mesh 3D* a été effectuée. L'objet au repos a donc été modélisé et ce modèle a été géoréférencé dans le système de coordonnées local défini par les QR Codes (voir Figure 30).



Figure 30 : Modèle 3D géoréférencé à l'aide de QR Codes

• Détermination des coordonnées des QR Codes simples

Comme mentionné au point 4.3.1, le mouvement de l'objet sur son support a été découpé en 7 poses et chacune de ces poses a été prise en photo depuis 16 points de vue répartis autour du support. Pour chaque pose, les 16 images ont été importées dans CONTEXT CAPTURE et une aérotriangulation avec détection automatique des QR Codes a été réalisée. La détection automatique des QR Codes permet d'effectuer l'orientation absolue de nos images dans le système de coordonnées local défini par le plan de QR Codes. Elle permet également d'extraire les QR Codes simples non géoréférencés et de déterminer les coordonnées de leur centre sur base des paramètres de l'orientation absolue.

Notons que, ici aussi, un nettoyage manuel a été nécessaire pour supprimer les points de contrôle détectés avec une erreur RMS de reprojection supérieure au pixel, ainsi que ceux détectés sur uniquement 3 photographies ou moins. Le processus n'est donc pas entièrement automatisé car il nécessite de contrôler la précision des détections de QR Codes.

Pour chacun des 4 QR Codes simples placés sur l'objet, leurs coordonnées X, Y et Z dans le système de coordonnées local ont été exporté au format XML. Un tableau reprenant ces coordonnées aux 7 instants successifs a été réalisé (voir Annexe 7). De plus, un inventaire des erreurs de reprojection et de distance au rayon a été dressé (voir Tableau 3 et Tableau 4). On mesure ainsi une erreur moyenne de reprojection de 0,649 pixel et une erreur moyenne de distance au rayon de 0,106 millimètre.

Tableau 3: Erreur de reprojection (en pixel) par marqueur et par position successive du mouvement de l'objet.

<i>Instants</i>	N° des QR Codes simple				Moy.
	101	102	103	104	
1	0,440	0,590	0,550	0,380	0,490
2	0,830	0,420	0,760	0,800	0,703
3	0,540	0,450	0,680	0,800	0,618
4	0,600	0,620	0,560	1,110	0,723
5	0,560	0,530	0,690	1,130	0,728
6	0,570	0,520	0,860	0,550	0,625
7	0,670	0,360	0,710	0,880	0,655
Moy.	0,601	0,499	0,687	0,807	0,649

Tableau 4: Erreur de distance au rayon (en mm) par marqueur et par position successive du mouvement de l'objet.

<i>Instants</i>	N° des QR Codes simple				Moy.
	101	102	103	104	
1	0,069	0,092	0,087	0,061	0,077
2	0,135	0,067	0,120	0,135	0,114
3	0,085	0,073	0,110	0,137	0,101
4	0,109	0,112	0,099	0,185	0,126
5	0,101	0,087	0,117	0,181	0,122
6	0,088	0,081	0,147	0,085	0,100
7	0,104	0,055	0,109	0,138	0,102
Moy.	0,099	0,081	0,113	0,132	0,106

Considérant ces niveaux de précisions comme acceptables, nous pouvons passer au traitement des données.

5.3.2 Traitement des données

Le traitement des données consiste à calculer la trajectoire des marqueurs en interpolant les 7 positions déterminées à l'aide de la photogrammétrie. Pour cela, nous avons utilisé le logiciel MATLAB.

La méthode choisie pour l'interpolation de la trajectoire des marqueurs est une interpolation par *Spline Cubic 3D*. Cette méthode est connue pour donner une meilleure approximation de la trajectoire qu'une polynomiale de degrés $n+1$ sur l'ensemble de l'intervalle pour un échantillon de faible grandeur (7 positions dans notre cas). En effet, une polynomiale basée sur des échantillons équidistants présente de grandes oscillations près des extrémités de l'intervalle d'interpolation.

Le script que nous avons développé se trouve à l'Annexe 8.

Nous avons ainsi calculé la trajectoire de chacun des marqueurs sur base des 7 positions connues aux précisions mentionnées au point précédent (voir Figure 31). Ces trajectoires pourront être importées dans le logiciel 3DS MAX pour animer notre objet dans l'espace.

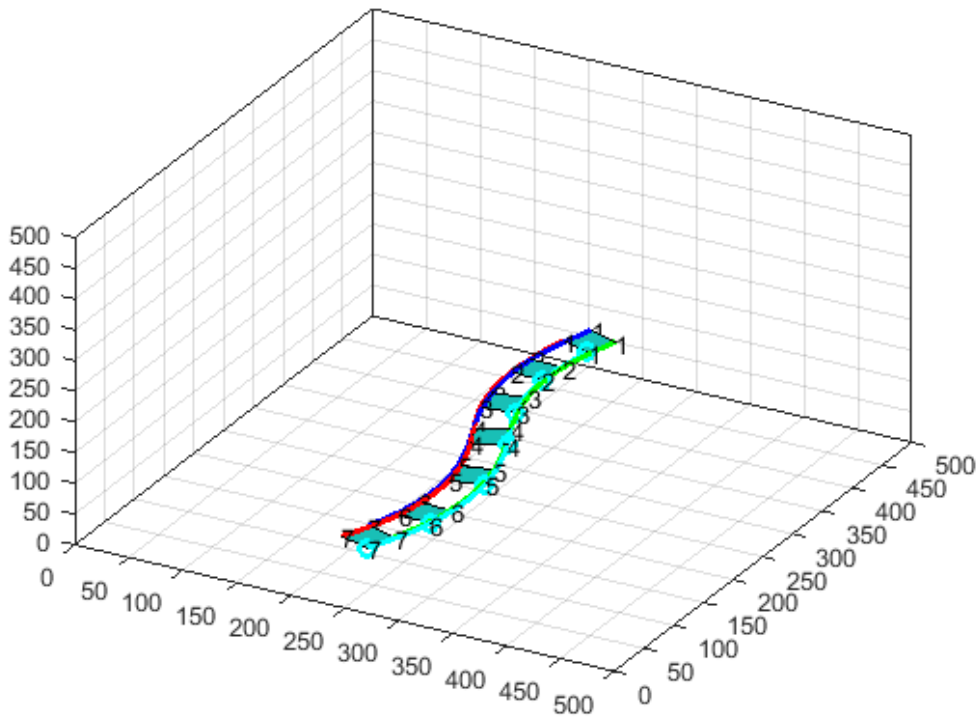


Figure 31 : Modélisation du calcul de trajectoire de marqueurs dans MATLAB

5.3.3 Rendu 3D

Possédant maintenant le support, le modèle de notre objet au repos et les trajectoires de marqueurs dans notre système de coordonnées local, nous pouvons importer le tout dans 3DS MAX afin de réaliser l'animation 3D de notre objet via le mode « *trajectory* ». Nous effectuerons deux rendus. Une première animation représentant le mouvement des marqueurs par rapport au support sera faite, suivit d'une deuxième animation représentant l'objet en mouvement sur le support.

Premièrement, nous avons importé le support et les trajectoires des différents marqueurs. Dans le logiciel, nous avons créé 4 sphères nommées respectivement de 101 à 104 représentant les différents marqueurs. Au centre de chacune de ces sphères, nous avons rattaché la trajectoire correspondante (voir Figure 32). Après cela, nous avons pu réaliser et exporter une animation, au format AVI, du mouvement des marqueurs par rapport au support. Cette animation s'est faite à une fréquence de 30 *fps* et à une résolution de 1920 x 1080 pixels. L'export de l'animation a été fait en vue orthographique, et ce sous différents points de vue afin de réaliser un montage vidéo.

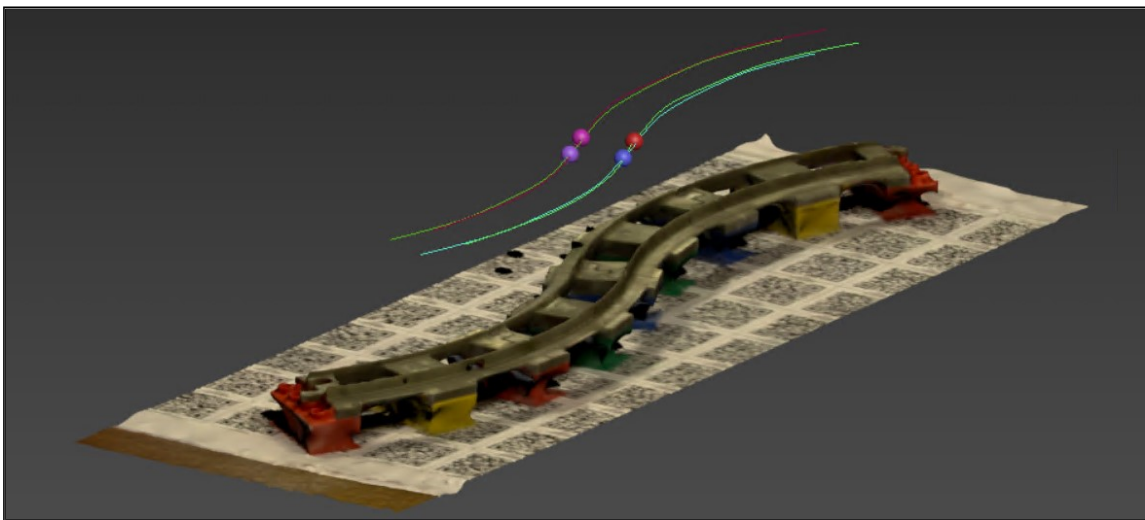


Figure 32 : Animation 3D des positions des marqueurs via les trajectoires importées.

Deuxièmement, nous avons importé le *mesh* 3D de l'objet fixe dans le logiciel afin de le mettre en mouvement. Le centre des QR Codes simples nous a permis d'effectuer la mise à l'échelle et de rattacher le centre des marqueurs au centre des sphères (voir Figure 33). Par la suite, avec la même méthodologie que pour l'animation précédente, nous avons animé l'objet et exporté une vidéo de même résolution et fréquence, et ce sous les mêmes points de vue.

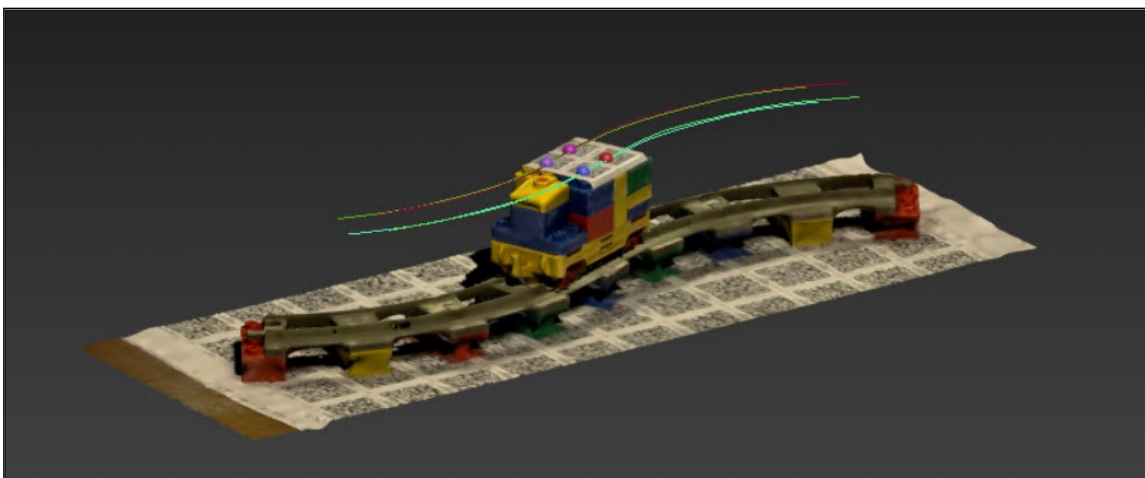


Figure 33 : Animation 3D de l'objet en mouvement par rattachement aux positions des marqueurs.

5.4 Un système de capture de mouvement basé sur la lasergrammétrie

Dans ce point, nous détaillerons les résultats de la capture du mouvement à l'aide du scanner laser et présenterons la méthode d'extraction du mouvement que nous avons développée.

5.4.1 Acquisition du mouvement

L'assemblage des scans s'est fait de la même manière qu'au point 4.2.1. Les scans ayant été effectués avec la même disposition spatiale et la même résolution (3,1mm à 10m), nous estimons une RMS erreur sur nos coordonnées inférieures à 0,0014 m. Les résultats de l'acquisition sont deux nuages de points, un issu de 4 scans acquis lorsque l'objet était en mouvement sur le support et l'autre issu de 4 scans acquis lorsque l'objet n'était pas sur le support. Nous avons détaillé la méthode d'acquisition au point 4.5.1.

5.4.2 Extraction du mouvement

Dans ce point, nous présenterons le détail, les résultats et les perspectives d'une méthode d'extraction du mouvement que nous avons développée. Cette méthode se base sur une analyse de la distance entre le nuage de points représentant le support du mouvement (rails), renseigné ci-dessous comme le « *nuage support* » (voir Figure 34) et le nuage de points représentant le support et l'objet en mouvement (rails + train en mouvement), renseigné ci-dessous comme le « *nuage support-mouvement* » (voir Figure 35).

- **Méthode**

Afin d'extraire l'objet en mouvement sur base de la distance entre le « *nuage support* » et le « *nuage support-mouvement* ». Les deux nuages, référencés dans le même système de coordonnées, ont été importés dans le logiciel CLOUDCOMPARE. Un nettoyage manuel du « *nuage support-mouvement* » a été effectué de manière à isoler l'objet et son support du reste de la scène.

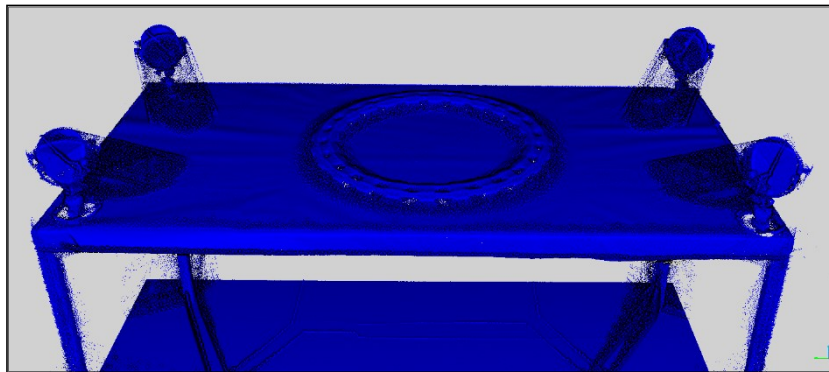


Figure 34 : Nuage de points « Support »

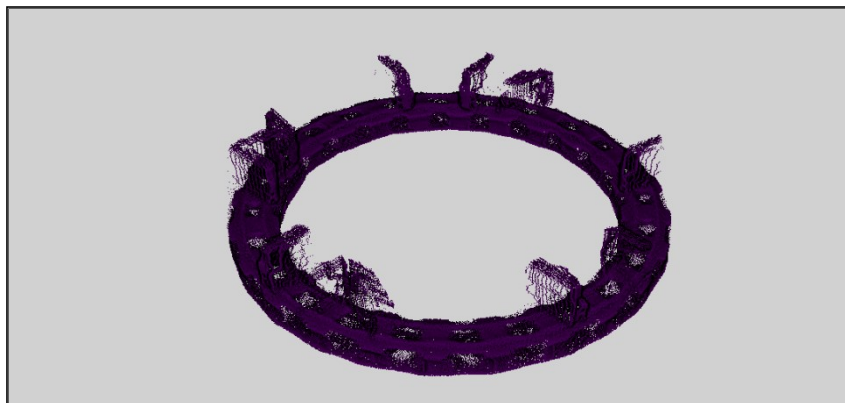


Figure 35 : Nuage de points « Support-Mouvement »

Premièrement, le rôle des différents nuages dans l'analyse doit être défini. Un *nuage de points de comparaison* et un *nuage de points de référence* doivent être renseignés.

- Le *nuage de comparaison* est celui sur lequel les distances seront calculées. CLOUDCOMPARE calculera les distances de chacun de ses points par rapport au *nuage de référence* et un champ scalaire contenant l'information de distance sera généré et hébergé dans ce *nuage de comparaison*.
- Le *nuage de référence* est le nuage depuis lequel les distances seront calculées.

Dans notre cas, les deux nuages ont la même densité de points. Le « *nuage support-mouvement* » sera le *nuage de comparaison* et le « *nuage support* », d'étendue plus large, sera le *nuage de référence* (voir Figure 36).

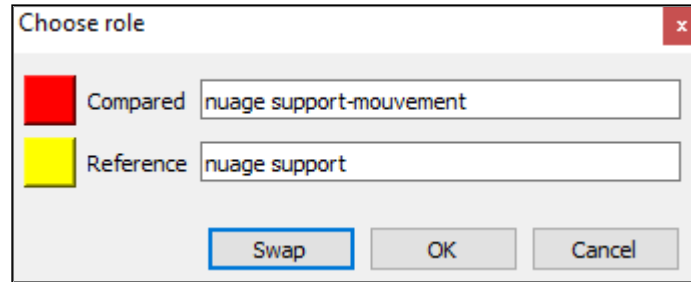


Figure 36 : Sélection du nuage de points de référence et de comparaison dans l'outil de calcul de distance de CLOUDCOMPARE

Deuxièmement, une approximation des distances sera calculée par le logiciel lors de l'ouverture de la boîte de dialogue de l'outil. Ces distances approximatives permettent au logiciel d'estimer le meilleur « *Octree level* » pour effectuer le calcul des distances réelles. L'« *Octree Level* » est le niveau de subdivision des octrees auquel le calcul de distance est effectué. Par default, il est défini automatiquement par CLOUDCOMPARE, mais nous le renseignons manuellement. Nous choisirons un niveau d'octree de 10, ce qui correspond à une division de notre *Bounding Box* en 264 849 cellules dans lesquelles les distances seront calculées.

Troisièmement, le mode de calcul des distances entre nuages doit être renseigné au logiciel. Quand aucun modèle local n'est utilisé, la distance entre nuages est simplement la distance du plus proche voisin. Le problème est que le plus proche voisin est rarement le point le plus proche sur la surface représentée par le nuage. Cela est particulièrement vrai si le nuage de référence a une faible densité ou présente de écarts importants entre les points (voir Figure 37).

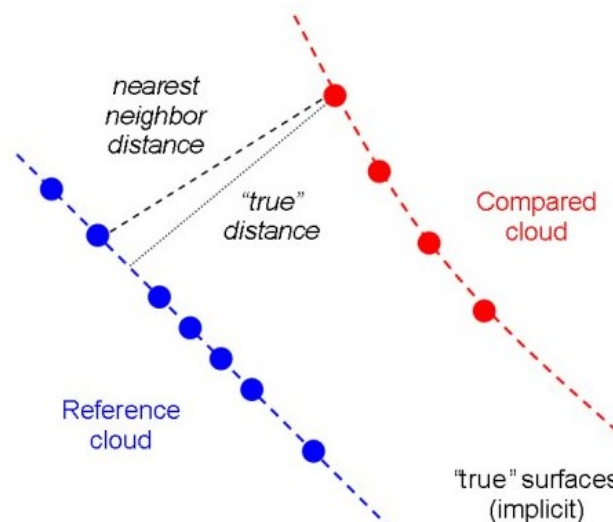


Figure 37 : Illustration de l'erreur sur le calcul des distances avec la méthode du « plus proche voisin » (source : <http://www.cloudcompare.org>).

Dans ce cas, il peut être judicieux d'utiliser une « stratégie de modélisation locale » qui consiste à calculer un modèle local autour des points de manière à approcher la surface réelle et à obtenir une meilleure estimation de la distance « réelle » entre nuages (voir Figure 38). Pour cela, trois modèles sont proposés par le logiciel. Tous trois sont basés sur un plan d'ajustement par moindres carrés qui passe par le point le plus proche et ses voisins. On distingue donc les modes :

1. *Plan des moindres carrés* : Ce mode de calcul utilise directement le plan calculé pour déterminer les distances réelles entre nuages.
2. *La triangulation 2D1/2* : Ce mode de calcul utilise la projection des points sur le plan pour calculer une triangulation de Delaunay (en utilisant les points 3D originaux comme sommets pour le maillage afin d'obtenir un maillage 2.5D).
3. *Quadric* : Ce mode de calcul utilise une fonction quadratique à 6 paramètres.

Classés dans cet ordre, les modèles locaux sont triés par « fidélité » croissante à la géométrie locale (et aussi au temps de calcul nécessaire). Notons que pour faire un choix, il faut également prendre en considération si la géométrie locale est principalement lisse ou avec des bords tranchants. En effet, la triangulation de Delaunay est le seul modèle qui peut théoriquement représenter des arêtes vives (mais elle suppose que nous ayons des points sur les bords) et la fonction quadratique est la seule qui peut représenter des surfaces lisses ou courbes. Par défaut, il est recommandé d'utiliser le modèle quadratique, car il est le plus polyvalent. C'est donc ce dernier que nous utiliserons.

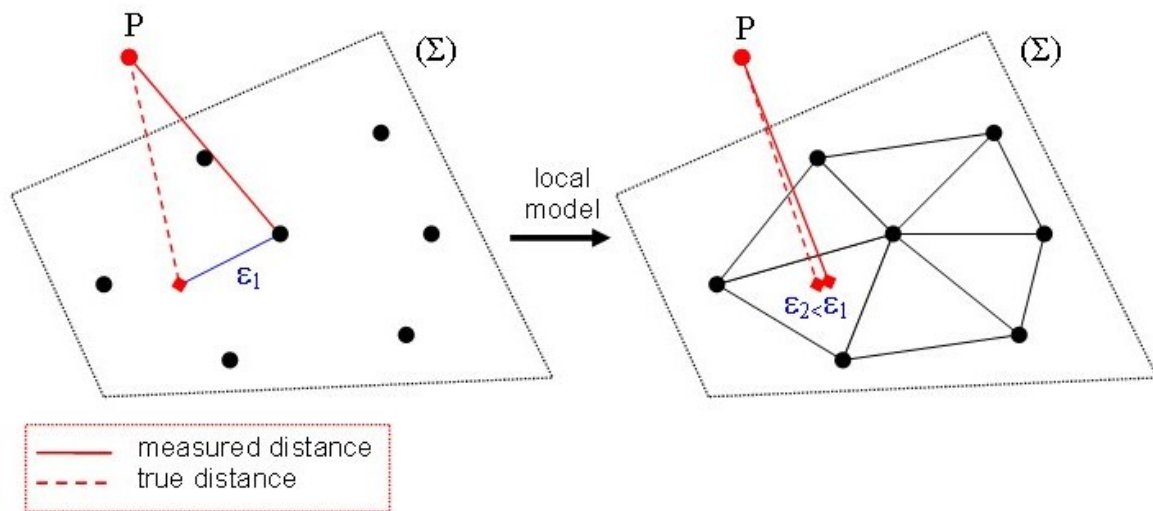


Figure 38 : Utilisation d'un modèle local pour le calcul des distances entre nuages de points
(Source : <http://www.cloudcompare.org>)

Nous disposons maintenant de suffisamment d'éléments pour lancer le calcul des distances réelles entre nos deux nuages de points.

• Résultats

Le résultat est donc notre « *nuage support-mouvement* » hébergeant le champ scalaire renseignant la distance au nuage de référence pour chacun des points qu'il contient. L'histogramme des distances est visible à la Figure 39.

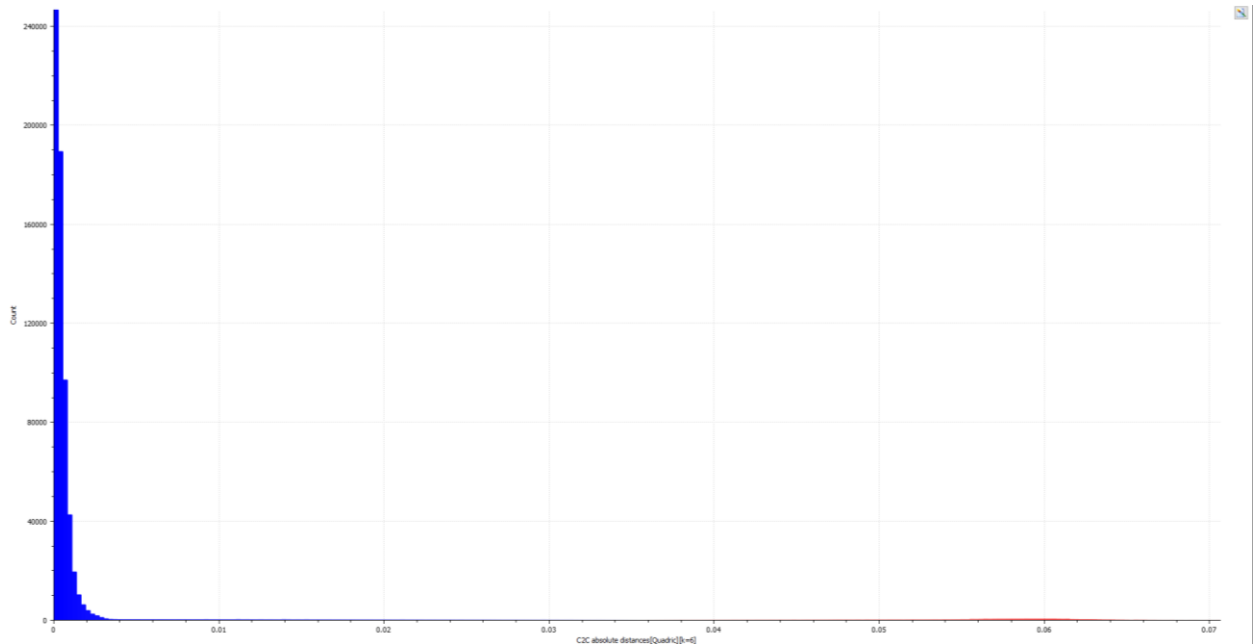


Figure 39 : Histogramme de distance absolue entre nos deux nuages de points avec utilisation d'un modèle local quadratique.

Nous observons qu'une grande partie des points se situent sous la barre des 0,004 m de distance avec le « *nuage support* ». Néanmoins, il est difficile de distinguer, directement sur l'histogramme, les points appartenant au support, les points appartenant au bruit de mesure et les points appartenant à l'objet en mouvement. Pour cela, une analyse basée sur les distances doit être faite.

Ainsi, nous avons exporté l'ensemble des données de l'histogramme au format CSV afin de les traiter dans MATLAB. Le CSV en sortie de CLOUDCOMPARE possède 4 attributs et 256 lignes. On y retrouve le numéro des classes de distance, le nombre de points qu'elles contiennent et la limite inférieure et la limite supérieure des classes. Les données étant triées par classe de distance, nous avons déterminé le milieu de chacune de ces classes et nous avons calculé les sommes cumulées des points dans chacune de ces dernières. Les résultats de cette opération sont présents dans le Tableau n° 5.

Tableau 5 : Résultats du calcul de centre des classes et des sommes cumulées de points par classe.

Class	Nb of point	Class start (m)	Class end (m)	Sum cum.	Mid class
1	246620	0	0.00027602	246620	0.00013801
2	189353	0.00027602	0.00055204	435973	0.00041403
3	97071	0.00055204	0.00082805	533044	0.00069004
4	42592	0.00082805	0.00110407	575636	0.00096606
...	...	...	...	...	...
254	4	0.06983245	0.07010847	674370	0.06997046
255	1	0.07010847	0.07038448	674371	0.07024648
256	3	0.07038448	0.0706605	674374	0.07052249

Nous pouvons réaliser un graphique (voir Figure 40) sur base des résultats calculés, à savoir les milieux de classe de distance en abscisse et la somme cumulée des nombres de points en ordonnée.

On obtient alors le graphique en bleu. Ce graphique permet de visualiser la différence, en termes de distance, entre les points du support, les points du bruit de mesure et les points de l'objet en mouvement. En effet, on peut identifier 3 zones sur ce graphe.

La première correspond aux points appartenant au support et présentant des distances très faibles, proches de la précision à laquelle les mesures sont effectuées. Cette zone représente la majeure partie des points et nous pouvons considérer qu'elle s'étend aux points dont les distances entre le « nuage support » et le « nuage support – mouvement » sont comprises entre 0 à 0,001 m. On retrouve dans cette catégorie 85,36 % de nos points.

La deuxième correspond au bruit de mesure, il s'agit de la zone d'inflexion de la courbe. On retrouve, dans cette catégorie, des points issus du bruit de mesure, mais aussi des points appartenant au support et d'autres à l'objet en mouvement.

La troisième zone correspond aux points n'appartenant qu'au support. Il s'agit de la partie supérieure de la courbe. La frontière entre la deuxième et troisième zone est difficile à estimer, car nous ne disposons d'aucune information pour la tracer. Nous développerons dans ce point une méthode nous permettant de l'estimer.

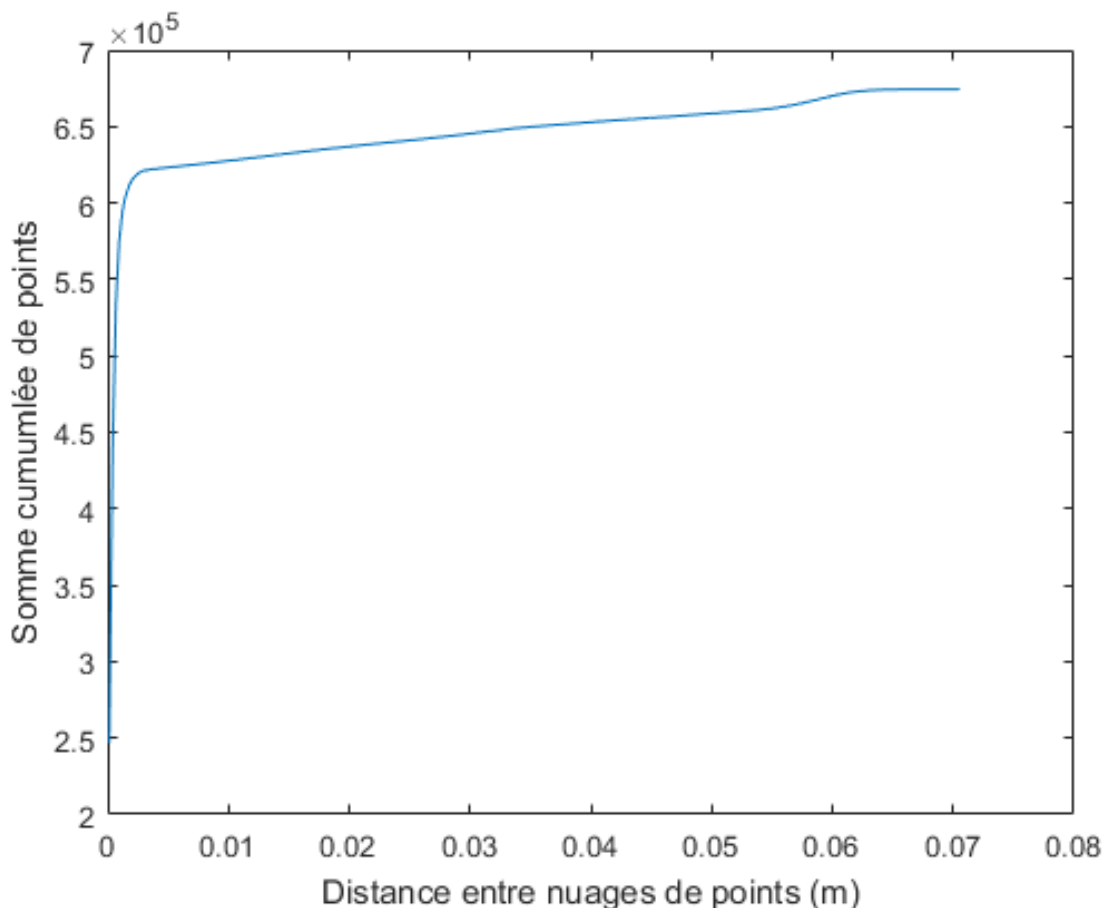


Figure 40 : Graphique représentant les milieux de classe de distance en abscisse et la somme cumulée des nombres de points correspondante en ordonnée.

A l'aide de l'application *Curve Fitting* de MATLAB nous avons pu déterminer les paramètres de la tangente à la partie supérieure de la courbe. Il s'agit d'une polynomiale de régression du premier degré dont les paramètres sont les suivants.

```
Linear model Poly1:
f(x) = p1*x + p2
Coefficients (with 95% confidence bounds):
p1 = 8.043e+05 (7.828e+05, 8.259e+05)
p2 = 6.202e+05 (6.193e+05, 6.21e+05)

Goodness of fit:
SSE: 3.236e+09
R-square: 0.989
Adjusted R-square: 0.9889
RMSE: 3570
```

L'intersection de cette tangente avec l'axe des Y (ordonnée à l'origine) nous donne alors une indication du nombre de points appartenant au support, à savoir 620 200 points. Ces points correspondent aux 620 200 points les plus proches du support. Connaissant ce nombre de points, il suffit de parcourir nos données pour trouver le centre de la catégorie de distance se rapprochant le plus de cette somme cumulée de points. Dans notre cas, nous trouvons la valeur de distance de 0,0028 m (voir Figure 41). Nous considérerons cette valeur de distance comme la frontière entre les points appartenant au bruit de mesure et les points appartenant à l'objet en mouvement. Cette partie a été réalisée à l'aide du script MATLAB présent à l'Annexe 9.

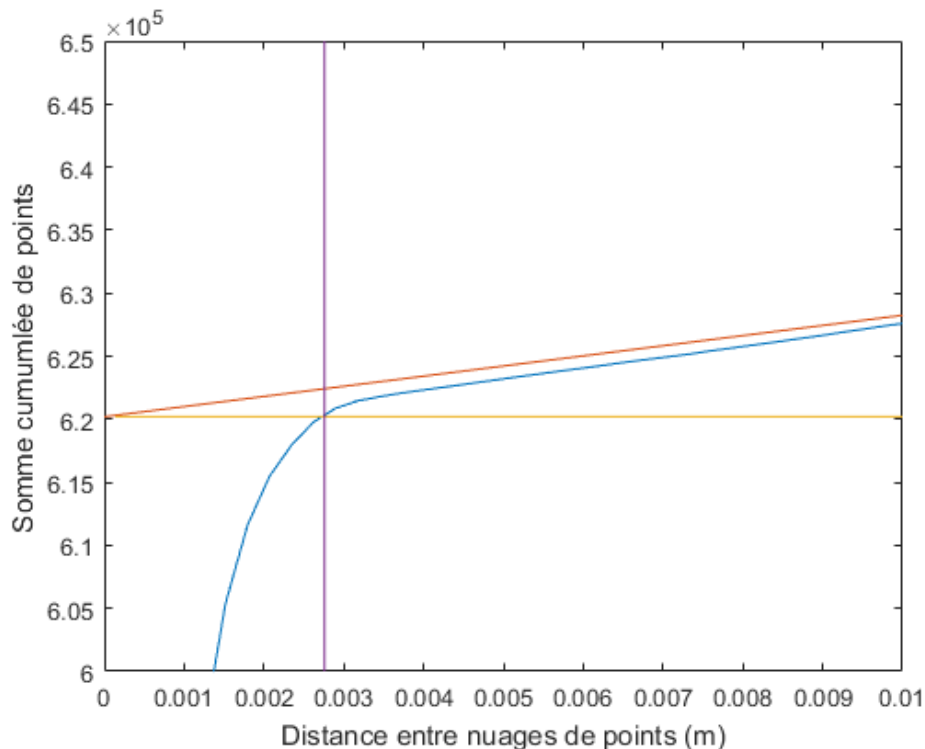


Figure 41 : Graphique représentant les milieux de classe de distance en abscisse et la somme cumulée des nombres de points correspondante en ordonnée (Zoom).

Ainsi, nous obtenons les résultats présents dans le Tableau 6.

Tableau 6 : Résultats de la méthode d'extraction du mouvement basée sur la distance entre le nuage de points "support" et le nuage de points "support - mouvement".

	% des points du nuage « support-mouvement »	Intervalles de distance (m)
Support	85.36 %	$0.0000 < i \leq 0.0010$
Bruit	6.62 %	$0.0010 < i \leq 0.0028$
Mouvement	8.02 %	$0.0028 < i < \max$

Pouvant maintenant identifier la frontière entre le bruit de mesure et les points du mouvement, nous classifions notre nuage de points et isolons ainsi les points de l'objet en mouvement (voir Figure 42 et Figure 43).

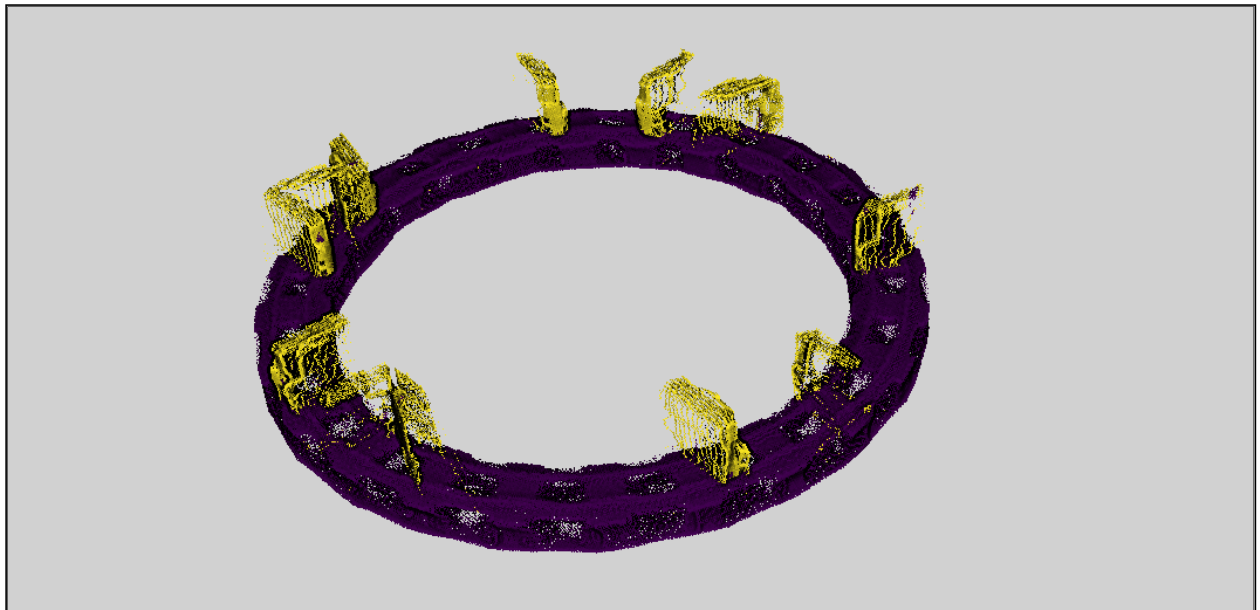


Figure 42 : Classification du nuage de points "support - mouvement"

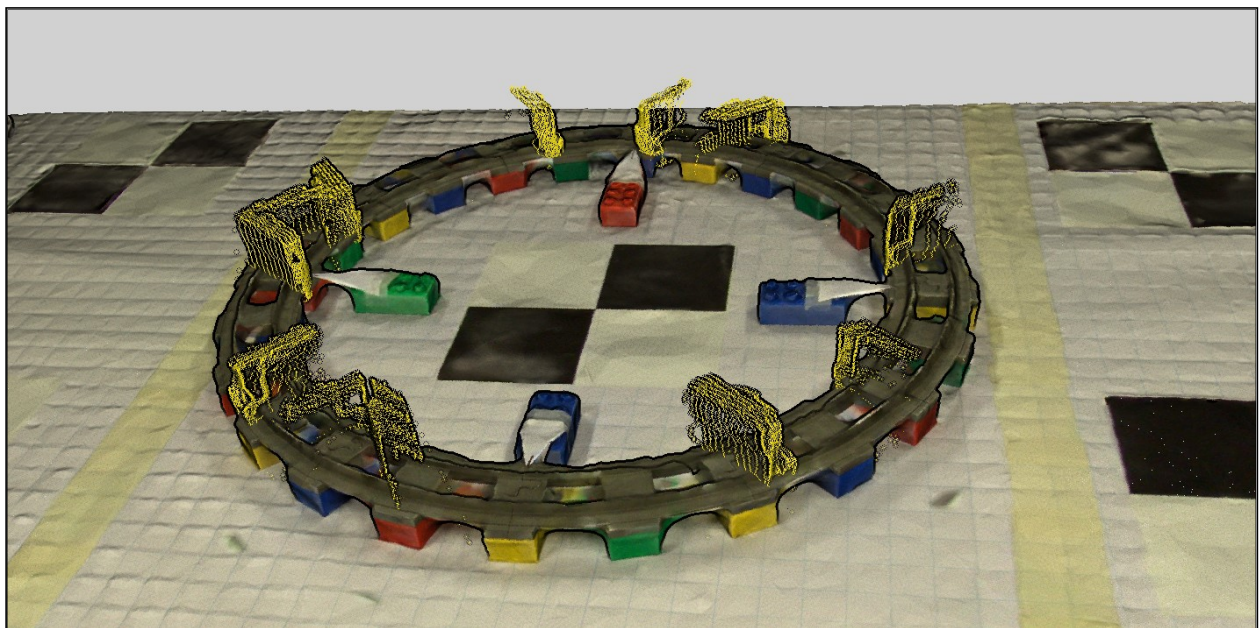


Figure 43 : Classification du nuage de points "support - mouvement" et superposition du nuage de points « support »

6 DISCUSSION

Après avoir testé notre travail par la mise en pratique de trois méthodes développées sur base de nos connaissances théoriques, nous pouvons maintenant analyser nos résultats afin de valider ou d'invalidier nos hypothèses de recherche. Notons qu'il est important de préciser que les analyses et les conclusions qui vont être développées dans ce point sont spécifiques aux méthodes développées et ne peuvent pas faire office de conclusion générale quant à l'utilisation de la photogrammétrie ou de la lasergrammétrie en matière de capture de mouvement. Ces conclusions doivent servir à orienter des recherches futures dans le domaine et non pas tirer un trait définitif sur l'usage de l'une ou l'autre technique d'acquisition.

Pour rappel, les hypothèses de recherche étaient les suivantes :

- **H1** : *« Un système optique et multi-capteurs de capture du mouvement, basé sur la photogrammétrie, permet de développer une méthode d'acquisition, de traitement et de rendu 3D d'objets en mouvement, à une précision millimétrique. »*
 - **H1-A** : *« Ce système de capture du mouvement optique multi-capteurs n'inclut pas des marqueurs posés sur l'objet en mouvement. »*
 - **H1-B** : *« Ce système de capture du mouvement optique multi-capteurs inclut des marqueurs posés sur l'objet en mouvement. »*
- **H2** : *« Un système de capture du mouvement basé sur la lasergrammétrie permettrait de développer une méthode d'acquisition, de traitement et de rendu 3D d'objets en mouvement, à une précision millimétrique. »*

Nous allons maintenant discuter ces hypothèses, nous commencerons par l'hypothèse H1.

En formulant cette hypothèse, le but était d'évaluer les possibilités d'utilisation de la photogrammétrie, en tant que technique d'acquisition, dans un dispositif de capture du mouvement. Comme nous l'avons mentionné, nous avons très vite été confronté à la limitation des ressources disponibles pour la recherche, ainsi qu'à certaines contraintes techniques que nous avons détaillées. Ces contraintes et ressources nous ont limités au développement de seulement deux méthodes d'acquisition utilisant la photogrammétrie. Ces deux méthodes d'acquisition, l'une avec marqueurs et l'autre sans marqueurs placés sur l'objet, ont été développées sur base de nos connaissances théoriques et sur base des informations issues de l'état de l'art que nous avons effectué. Elles ont pour but de tester les deux sous-hypothèses de recherches H1-A et H1-B.

Pour valider l'entièreté de notre première hypothèse, nous devons valider chacun des éléments qui la constituent.

Un premier élément de l'hypothèse H1 interroge la possibilité d'une méthode de capture, de traitement et de rendu 3D du mouvement d'un objet à une précision millimétrique. Nous avons tenté de répondre à cette question au travers des deux méthodes d'acquisition mises en place. Comme nous l'avons déjà mentionné, la précision en photogrammétrie peut s'évaluer à l'aide de deux erreurs mesurables sur le modèle final : l'erreur RMS de reprojection (exprimée en pixels) et l'erreur RMS de distance au rayon (exprimée dans l'unité de mesure du système de coordonnées utilisé).

Pour la première méthode, à savoir le système de capture de mouvement optique multi-capteurs n'incluant pas des marqueurs posés sur l'objet en mouvement, nous avons montré qu'il était possible d'atteindre des erreurs RMS de reprojection inférieures au pixel et des erreurs RMS de distance au rayon de maximum 0,002 m. Nous avons fait cela en utilisant une résolution de prises

de vue limitée à 1080 pixels de largeur et en altérant nos mesures, comme nous l'avons déjà mentionné, avec des erreurs dues au caractère réfléchissant de l'objet ou dues à la synchronisation des prises de vues via une méthode visuelle et manuelle. Nous sommes donc en mesure de dire que cette méthode est suffisamment précise, et qu'elle pourrait même gagner en précision avec une plus grande résolution des prises de vue, sur des objets moins réfléchissants et en utilisant des méthodes de synchronisation plus performantes.

Pour ce qui est de la précision millimétrique, l'hypothèse H1 peut donc être validée dans le cas de la sous-hypothèse H1-A.

Pour la seconde méthode, à savoir le système de capture de mouvement optique multi-capteurs incluant des marqueurs posés sur l'objet, nous avons montré qu'il était possible d'atteindre des erreurs RMS de reprojection inférieures au pixel et des erreurs RMS de distance au rayon inférieures au millimètre. Étant donné que le but de cette méthode était de tester l'utilisation des QR Codes pour géoréférencer notre modèle et extraire les positions de marqueurs placés sur l'objet, nous avons voulu éviter toute source d'erreur qui aurait pu venir altérer notre verdict. Ainsi, nous n'avons pas limité la résolution des images à 1080 pixels de largeur, mais avons utilisé une résolution maximale de 3840 pixels de largeur. De plus, les prises de vue ont été faites sur une simulation du mouvement de l'objet à intervalles réguliers et non sur un objet réellement en mouvement, nous avons ainsi pu nous passer des erreurs de synchronisation. Étant donné que nous cherchions uniquement à déterminer la position du centre des QR codes par photogrammétrie, la précision des positions de l'objet en mouvement n'est pas altérée par le bruit de mesure liée à la nature de la surface de l'objet. La seule source d'erreur est donc l'algorithme de lecture et détection de centre des QR Codes dont nous avons ainsi pu évaluer la précision. Une précision qui répond à nos attentes.

Pour ce qui est de la précision millimétrique, l'hypothèse H1 peut donc être validée dans le cas de la sous-hypothèse H1-B.

Le deuxième élément important de notre hypothèse H1 est que le système de capture développé permet de mettre en place une méthode de capture, de traitement et de rendu 3D du mouvement d'un objet. On parle donc d'une méthode entièrement fonctionnelle du début à la fin, de l'acquisition au rendu 3D du mouvement de l'objet. La réponse à cet élément de l'hypothèse est dans notre cas, différente en fonction de la sous-hypothèse envisagée.

Dans le cas du système optique multi-capteurs n'incluant pas des marqueurs posés sur l'objet, nous avons vu que le rendu d'une animation 3D de l'objet dans le logiciel 3DS MAX était impossible étant donné la nature de la surface de l'objet que nous avons essayé de représenter. Le logiciel était incapable de trouver des similitudes entre des modèles 3D représentant le même objet à des instants différents de son mouvement. Un rendu 3D fluide du mouvement de l'objet était donc impossible avec les conditions d'expérience que nous avons choisies. De meilleurs résultats auraient été atteints en utilisant un objet de texture plus mate et moins sensible au bruit de mesure. Bien qu'une animation 3D du mouvement n'a pas été réalisée, nous avons quand même pu décomposer le mouvement en étapes successives, ce qui permet déjà une première représentation 3D de ce dernier.

Pour ce qui est de la mise en place d'une méthode complète, de l'acquisition au rendu 3D, l'hypothèse H1 peut donc être validée dans le cas de la sous-hypothèse H1-A, moyennant une nuance vis-à-vis du rendu 3D final qui ne permet pas une animation du mouvement, mais bien, une décomposition en instants successifs du mouvement.

Dans le cas du système de capture de mouvement optique multi-capteurs incluant des marqueurs posés sur l'objet, nous avons vu que le rendu d'une animation 3D de l'objet dans le logiciel 3DS MAX était possible. Cela est en grande partie dû aux conditions de mesures plus optimales

(nombre et qualité des prises de vues utilisées pour produire le modèle 3D de l'objet au repos). Le logiciel était donc capable d'animer le modèle 3D de l'objet en le rattachant aux trajectoires calculées des marqueurs. Un rendu 3D fluide du mouvement de l'objet était donc possible, bien que ce dernier ne soit qu'une animation dans les trois dimensions de l'espace d'un modèle de l'objet fixe. Ce n'est donc pas, à strictement parler, une animation de l'objet en mouvement. Néanmoins, cette méthode nous permet d'obtenir une modélisation 3D réaliste de l'objet en mouvement, et ce à des pressions acceptables. Rappelons néanmoins que cette méthode a pour prérequis que l'objet ne se déforme pas pendant son mouvement. On ne pourra donc pas utiliser cette méthode pour effectuer un monitoring de déformation 3D par exemple. De meilleurs résultats auraient pu être atteints en utilisant un objet de texture plus mate, moins sensible au bruit de mesure.

Pour ce qui est de la mise en place d'une méthode complète, de l'acquisition au rendu 3D, l'hypothèse H1 peut donc être validée dans le cas de la sous-hypothèse H1-B, moyennant une restriction de l'utilisation de cette méthode au monitoring de mouvement d'objet ne se déformant pas dans le temps.

Un système optique et multi-capteurs de capture du mouvement, basé sur la photogrammétrie, permet de développer une méthode d'acquisition, de traitement et de rendu 3D d'objets en mouvement, à une précision millimétrique.

Ainsi, l'hypothèse H1 peut être validée.

Continuons notre discussion avec l'hypothèse H2.

Pour cette hypothèse, l'objectif a donc été de savoir s'il était possible d'extraire le mouvement d'un objet depuis un nuage de points issu d'un scanner laser. Plus précisément, nous avons cherché à faire la distinction entre les points faisant partie du mouvement et les points qui ne font pas partie du mouvement sur base de la distance entre un nuage de points représentant le support de l'objet et un nuage de points représentant l'objet en mouvement sur son support. Pour cela, la méthode décrite au point 5.4 a été mise en place et les résultats de cette méthode ont été décrits.

Ainsi, notre méthode nous a permis d'isoler les points du support des points du mouvement. Néanmoins, dû au bruit de mesure, mais aussi à la texture de l'objet utilisé, une grande partie des points de notre nuage (6,62 %) n'a pas pu être classée, et a par défaut été assignée au bruit de mesure.

La méthode mise en place a donc permis d'identifier et d'extraire les points appartenant au mouvement de l'objet. Néanmoins, en affichant ces points dans CLOUDCOMPARE, on constate qu'il serait très difficile de reconstituer une animation complète du mouvement de l'objet uniquement sur base des points extraits. Aller plus loin dans cette direction nécessiterait des études plus poussées en matière de segmentation et de classification de nuage de points. Cette piste pourrait être explorée dans un autre travail de fin d'étude et ainsi rejoindre des sujets de recherche semblables déjà étudiés à l'Unité de Géomatique.

Dans l'ensemble, nous pouvons donc affirmer qu'un système de capture du mouvement basé sur la lasergrammétrie permettrait de développer une méthode d'acquisition, de traitement et de rendu 3D d'objets en mouvement, à une précision millimétrique, moyennant des études poussées en matière de segmentation et classification de nuage de points 3D.

Ainsi, l'hypothèse H2 peut être également validée.

7 CONCLUSION

Nous concluons ce travail de fin d'études en replaçant les résultats obtenus dans le contexte initial de notre recherche ainsi que dans le contexte global de la capture du mouvement. Nous ferons cela pour les deux techniques d'acquisition étudiées : la photogrammétrie et la lasergrammétrie.

Pour ce qui est de la photogrammétrie, nous avons mis en évidence qu'une méthode de capture du mouvement précise était possible moyennant certains compromis, que ce soit avec ou sans marqueurs placés sur l'objet. Dans le contexte initial de la recherche, à savoir l'application à une œuvre d'art cinétique comme l'œuvre « 24 boules sur 24 sommets inclinés » de Pol Bury, il est important de préciser que la mise en œuvre aurait été bien plus complexe étant donné le nombre important d'objets en mouvement et les difficultés variables en fonction de la méthode utilisée.

Également, nous garderons à l'esprit que les conditions d'expérimentation et les ressources disponibles étaient insuffisantes pour mettre en place une méthode de capture de mouvement applicable à n'importe quel type d'objet. De plus, il est important de rappeler les nombreux compromis effectués, ainsi les contraintes que nous nous sommes imposés pour assurer le bon déroulement des expériences. Parmi elles, la répétabilité du mouvement reste la plus contraignante et la plus indispensable étant donné le nombre extrêmement limité de capteurs dont nous disposions.

Ainsi, ce constat permet de mettre le doigt sur un élément essentiel de l'utilisation de la photogrammétrie en matière de capture de mouvement. Faute de moyens financiers, la recherche scientifique en la matière est souvent en retard quand on la compare aux avancées technologiques déjà présentes dans l'industrie cinématographique ou du jeu vidéo par exemple. Nous pouvons donc dire que la recherche sur l'utilisation de la photogrammétrie en capture du mouvement est, certes, intéressante, mais souvent dépassée quand on la compare aux techniques déjà développées dans le privé.

Pour ce qui est de la lasergrammétrie, nous avons mis en évidence qu'une méthode précise de capture du mouvement était possible moyennant des études poussées en matière de segmentation et de classification de nuage de points 3D.

Dans le contexte initial de notre recherche, nous ne disposons pas encore assez d'éléments pour aboutir à une méthode d'acquisition, de traitement et de rendu 3D opérationnel, mais nous pouvons d'ores et déjà affirmer qu'une extraction des objets en mouvement d'un nuage de points est possible.

Dans le contexte global de la capture du mouvement, l'intérêt scientifique de l'utilisation de la lasergrammétrie en tant que technique d'acquisition est donc tout autre. L'utilisation de cette technique n'en est qu'à ces débuts et les perspectives sont nombreuses dans des thèmes de recherche déjà abordés par les spécialistes des nuages de points.

Ainsi, les perspectives d'utilisation de la lasergrammétrie en matière de capture de mouvement forment un intérêt certain pour cette technique et représentent un challenge scientifique digne d'investissements futurs.

8 BIBLIOGRAPHIE

Articles de périodique:

- Bodenheimer, B., Rose, C., Rosenthal, S., & Pella, J. (1997). The process of motion capture: Dealing with the data. *Computer Animation and Simulation*, 3–18.
https://doi.org/10.1007/978-3-7091-6874-5_1
- Buch, N., Velastin, S. A., & Orwell, J. (2011). A review of computer vision techniques for the analysis of urban traffic. *IEEE Transactions on Intelligent Transportation Systems*, 920–939.
<https://doi.org/10.1109/TITS.2011.2119372>
- Corazza, S., Mündermann, L., Chaudhari, A. M., Demattio, T., Cobelli, C., & Andriacchi, T. P. (2006). A markerless motion capture system to study musculoskeletal biomechanics: Visual hull and simulated annealing approach. *Annals of Biomedical Engineering*, 1019–1029.
<https://doi.org/10.1007/s10439-006-9122-8>
- Gruen, A. (1997). Fundamentals of videogrammetry — A review. *Human Movement Science*, 155–187.
[https://doi.org/10.1016/S0167-9457\(96\)00048-6](https://doi.org/10.1016/S0167-9457(96)00048-6)
- Han, S., & Lee, S. (2013). A vision-based motion capture and recognition framework for behavior-based safety management. *Automation in Construction*, 131–141.
<https://doi.org/10.1016/j.autcon.2013.05.001>
- Henry, P., Krainin, M., Herbst, E., Ren, X., & Fox, D. (2012). RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments. *International Journal of Robotics Research*, 647–663.
<https://doi.org/10.1177/0278364911434148>
- Hepach, R., Vaish, A., & Tomasello, M. (2015). Novel paradigms to measure variability of behavior in early childhood: posture, gaze, and pupil dilation. *Frontiers in Psychology*, 6.
<https://doi.org/10.3389/fpsyg.2015.00858>
- Kim, S., & Nussbaum, M. A. (2013). Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. *Ergonomics*, 314–326.
<https://doi.org/10.1080/00140139.2012.742932>
- Knyaz, V. A. (2015). Scalable photogrammetric motion capture system “mosca”: Development and application. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43–49.
<https://doi.org/10.5194/isprsarchives-XL-5-W6-43-2015>
- Kotsialos, A., & Papageorgiou, M. (2001). The importance of traffic flow modeling for motorway traffic control. *Networks and Spatial Economics*, 179–203.
<https://doi.org/10.1023/A:1011537329508>

- Moeslund, T. B., Hilton, A., & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 90–126.
<https://doi.org/10.1016/j.cviu.2006.08.002>
- Nguyen, K. D., Chen, I. M., Luo, Z., Yeo, S. H., & Duh, H. B. L. (2011). A wearable sensing system for tracking and monitoring of functional arm movement. *IEEE/ASME Transactions on Mechatronics*, 213–220.
<https://doi.org/10.1109/TMECH.2009.2039222>
- Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, 976–990.
<https://doi.org/10.1016/j.imavis.2009.11.014>
- Ren, J., Zhang, C., Zhang, L., Wang, N., & Feng, Y. (2018). Automatic Measurement of Traffic State Parameters Based on Computer Vision for Intelligent Transportation Surveillance. *International Journal of Pattern Recognition and Artificial Intelligence*, 1–21.
<https://doi.org/10.1142/S0218001418550030>
- Seo, J., Alwasel, A., Lee, S., Abdel-Rahman, E. M., & Haas, C. (2017). A comparative study of in-field motion capture approaches for body kinematics measurement in construction. *Robotica*.
<https://doi.org/10.1017/S0263574717000571>
- Yabukami, S., Kikuchi, H., Yamaguchi, M., Arai, K. I., Takahashi, K., Itagaki, A., & Wako, N. (2000). Motion capture system of magnetic markers using three-axial magnetic field sensor. *IEEE Transactions on Magnetics*, 3646–3648.
<https://doi.org/10.1109/20.908928>
- Zhu, H., Yu, Y., Zhou, Y., & Du, S. (2016). Dynamic human body modeling using a single RGB camera. *Sensors (Switzerland)*, 1–16.
<https://doi.org/10.3390/s16030402>

Actes d'un symposium ou d'une conférence :

- Bonenfant, M., Laurendeau, D., Fortin-Cote, A. (2017). A Computer Vision System for Virtual Rehabilitation. *14th Conference on Computer and Robot Vision (CRV)*, 269–276.
<https://doi.org/10.1109/CRV.2017.30>
- Stanton, C., Bogdanovych, A., & Ratanasena, E. (2012). Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning. *Australasian Conference on Robotics and Automation, ACRA*, 3–5.
- Starbuck, R., Seo, J., Han, S., & Lee, S. (2014). A Stereo Vision-based Approach to Marker-less Motion Capture for On-Site Kinematic Modeling of Construction Worker Tasks. *The Sixth Annual International Conference on Computing in Civil and Building Engineering*, 455–462.
<https://doi.org/10.1061/9780784413616.053>

Urtasun, R., Fleet, D. J., & Fua, P. (2005). Monocular 3-D tracking of the golf swing.
Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 932–938.
<https://doi.org/10.1109/CVPR.2005.229>

Ressources électroniques : sites web

Autodesk : <https://www.autodesk.be/fr/products/3ds-max/overview>

Cloud Compare : <http://www.cloudcompare.org/>

Context Capture : <https://www.bentley.com/en/products/brands/contextcapture>

Leica : <https://leica-geosystems.com/>

National instrument : <http://www.ni.com/white-paper/14103/fr>

9 ANNEXES

Listes des annexes

Annexe n°1 : Caractéristiques techniques des ressources matérielles.	64
Annexe n°2 : Bruit de mesure : données et régression polynomiale.....	67
Annexe n°3 : Calcul de distance et précisions : script MATLAB	68
Annexe n°4 : Classification des prises de vues : script MATLAB	69
Annexe n°5 : Rapport d'aérotriangulation pour l'orientation relative : exemple	70
Annexe n°6 : Rapport d'aérotriangulation pour l'orientation absolue : exemple.....	71
Annexe n°7 : Export XML des coordonnées des QR Codes : tableau de synthèse	72
Annexe n°8 : Trajectoires des marqueurs (Spline Cubic 3D) : script MATLAB	73
Annexe n°9 : Limite entre mouvement et bruit de mesure : script MATLAB.....	74

Annexe n°1 : Caractéristiques techniques des ressources matérielles.**LEICA TPS 1205****7.1 Angle Measurement****Accuracy**

Type	std. dev. Hz, V, ISO 17123-3		Display least count	
	["]	[mgon]	["]	[mgon]
1201+	1	0.3	0.1	0.1
1202+	2	0.6	0.1	0.1
1203+	3	1.0	0.1	0.5
1205+	5	1.5	0.1	0.5

Characteristics

Absolute, continuous, diametric.

7.2 Distance Measurement with Reflectors**Range**

Reflector	Range A		Range B		Range C	
	[m]	[ft]	[m]	[ft]	[m]	[ft]
Standard prism (GPR1)	1800	6000	3000	10000	3500	12000
Three standard prisms (GPR1)	2300	7500	4500	14700	5400	17700
360° prism (GRZ4, GRZ122)	800	2600	1500	5000	2000	7000
360° Mini prism (GRZ101)	450	1500	800	2600	1000	3300
Mini prism (GMP101)	800	2600	1200	4000	2000	7000
Reflector tape (GZM31) 60 mm x 60 mm	150	500	250	800	250	800
Machine Automation power prism (MPR122)  For Machine Control purposes only!	800	2600	1500	5000	2000	7000

Shortest measuring distance: 1.5 m

Accuracy

Accuracy refers to measurements to standard prisms.

EDM measuring mode	std. dev. ISO 17123-4, standard prism	std. dev. ISO 17123-4, tape	Measurement time, typical [s]
Standard	1 mm + 1.5 ppm	5 mm + 2 ppm	2.4
Fast	3 mm + 1.5 ppm	5 mm + 2 ppm	0.8
Tracking	3 mm + 1.5 ppm	5 mm + 2 ppm	< 0.15

Beam interruptions, severe heat shimmer and moving objects within the beam path can result in deviations of the specified accuracy.

The display resolution is 0.1 mm.

Characteristics

Principle:	Phase measurement
Type:	Coaxial, visible red laser
Carrier wave:	658 nm
Measuring system:	System analyser basis 100 MHz - 150 MHz

7.3**Distance Measurement without Reflectors****Range**

Type	Kodak Gray Card	Range D		Range E		Range F	
		[m]	[ft]	[m]	[ft]	[m]	[ft]
R400	White side, 90 % reflective	200	660	300	990	>400	>1310
R400	Grey side, 18 % reflective	150	490	200	660	>200	>660
R1000	White side, 90 % reflective	800	2630	1000	3280	>1000	>3280
R1000	Grey side, 18 % reflective	400	1320	500	1640	>500	>1640

Range of Measurement: 1.5 m - 1200 m
 Display unambiguous: up to 1200 m

Atmospheric conditions

D: Object in strong sunlight, severe heat shimmer
 E: Object in shade, sky overcast
 F: Underground, night and twilight

Accuracy

Standard measuring	std. dev. ISO 17123-4	Measure time, typical [s]	Measure time, maximum [s]
0 m - 500 m	2 mm + 2 ppm	3 - 6	12
>500 m	4 mm + 2 ppm	3 - 6	12

Object in shade, sky overcast. Beam interruptions, severe heat shimmer and moving objects within the beam path can result in deviations of the specified accuracy. The display resolution is 0.1 mm.

Characteristics

Type: Coaxial, visible red laser
 Carrier wave: 658 nm
 Measuring system: System analyser basis 100 MHz - 150 MHz

Laser dot size

Distance [m]	Laser dot size, approximately [mm]
at 30	7 x 10
at 50	8 x 20

Source: Leica TPS1200+ User Manual Version 9.1; Leica Geosystems

LEICA P30

Précision du système													
Précision d'une mesure simple*													
Précision sur la portée	1,2 mm + 10 ppm sur toute la portée												
Précision angulaire	8" horizontal, 8" vertical												
Précision en 3D	3 mm à 50 m, 6 mm à 100 m												
Acquisition de cible**	Écart type 2 mm à 50 m												
Compensateur bi-axial	Compensateur « temps réel » intégré, sélection on / off, résolution 1", plage dynamique ± 5', précision 1,5"												
Système de mesure de distance													
Type	Temps de vol ultra-rapide continu amélioré par la technologie WaveForm Digitising (WFD)												
Longueur d'onde	1 550 nm (invisible) / 658 nm (visible)												
Classe laser	1 (selon la norme CEI 60825:2014)												
Divergence de faisceau	< 0,23 mrad (FWHM, angle total)												
Diamètre du faisceau à la sortie	≤ 3,5 mm (FWHM)												
Portée et réflectivité	Portée minimale 0,4 m Portée maximale en fonction de la réflectivité												
	<table><tr><td></td><td>120 m</td><td>180 m</td><td>270 m</td></tr><tr><td>P30</td><td>18 %</td><td>–</td><td>–</td></tr><tr><td>P40</td><td>8 %</td><td>18 %</td><td>34 %</td></tr></table>		120 m	180 m	270 m	P30	18 %	–	–	P40	8 %	18 %	34 %
	120 m	180 m	270 m										
P30	18 %	–	–										
P40	8 %	18 %	34 %										
Vitesse de numérisation	Jusqu'à 1 000 000 points par seconde												
Précision de surface modélisée*	0,4 mm rms à 10 m 0,5 mm rms à 50 m												
Champ visuel													
Horizontal	360°												
Vertical	290°												
Capacité de stockage	256 Go interne SSD ou clé USB externe												
Communications / Transfert de données	Ethernet Gigabit, réseau LAN sans fil (WLAN) intégré ou périphérique USB 2.0												
Écran intégré	Contrôle par écran tactile avec styllet, affichage graphique couleur VGA (640 x 480 pixels)												
Plomb laser	Classe laser 1 (CEI 60825:2014) Précision de centrage: 1,5 mm à 1,5 m Diamètre du point laser 2,5 mm à 1,5 m Désactivable												
Système d'imagerie													
Caméra interne													
Résolution	4 mégapixels pour chaque image couleur 17° x 17° ; 700 mégapixels pour les images panoramiques												
Taille des pixels	2,2 µm												
Vidéo	Vidéo en streaming avec zoom, adaptation automatique à la lumière ambiante												
Balance des blancs	Ensoleillé, nuageux, lumière froide, lumière chaude, personnalisé												
HDR	Avec mappage ton local / plage complète												
Caméra externe	Compatible Canon EOS60D et EOS70D												

Alimentation	
Alimentation	24 V CC, 100 – 240 V CA
Type de batterie	x2 interne: lithium-ion; externe: lithium-ion (connexion via port externe, utilisation simultanée, remplaçable à chaud)
Autonomie	Interne > 5,5 h (2 batteries) Externe > 7,5 h (température ambiante)

Environnement	
Température d'utilisation	-20° C à +50° C
Température de stockage	-40° C à +70° C
Humidité	95 %, sans condensation
Poussière/Humidité	Étanchéité aux particules solides et aux liquides IP54 (CEI 60529)

Données physiques	
Scanner	
Dimensions (P x L x H)	238 mm x 358 mm x 395 mm
Poids	12,25 kg nominal (sans batteries)
Batterie (interne)	
Dimensions (P x L x H)	40 mm x 72 mm x 77 mm
Poids	0,4 kg
Fixation	À l'endroit ou à l'envers

Panneau de contrôle	
Écran tactile couleur pour le pilotage du scanner.	
Pilotage à distance: carnet de terrain Leica CS10/CS15 ou iPad, iPhone, smartphones ou autres périphériques utilisant la connexion à distance; simulateur externe.	

Fonctions	
Fonctions topos et assemblage embarqué	Orientation rapide, définition de l'azimut, visée arrière connue, résection (4 et 6 paramètres), cheminement
Contrôle et Ajustement	Procédure terrain pour la vérification des paramètres d'angles, de compensation et du distancemètre
Acquisition de cible intégrée	Sélection de cible depuis la vidéo temps réel ou depuis un scan
Interface utilisateur intégrée	Possibilité de passer d'une interface simplifiée à une interface avancée
Touche « Cliquez – scannez »	Utilisation du scanner avec un bouton unique
Définition de la zone de numérisation	Sélection de la zone de numérisation depuis la vidéo ou un scan; préparation et lancement de scans successifs

Information pour commander	
Contactez Leica Geosystems ou ses représentants agréés.	

Sous réserve de modifications.
Toutes données de précision $\pm$ un sigma, sauf indication contraire.
* Albédo à 78 %
** Ajustement algorithmique pour cibles HDS 4.5" noir & blanc planes
Scanner: Classe laser 1 selon la norme CEI 60825:2014
Plomb laser: Classe laser 1 selon la norme CEI 60825:2014
iPhone et iPad sont des marques déposées de Apple Inc.

Scan duration (ScanStation P30):

Resolution [mm @ 10 m]	Estimated scan duration [HH:MM:SS] for a full dome scan
50	00:00:20
25	00:00:33
12.5	00:00:58
6.3	00:01:49
3.1	00:03:30
1.6	00:13:33
0.8	00:54:07

Source: Leica P30-P40 User Manual Version 2.0; Leica Geosystems

Canon 5D Mark III

Objectif : Canon EF 24 mm 1:4, focale fixe.

Taille du capteur : 36x24 mm

Résolution Vidéo maximale : 1920x1080 pixels

Fréquence vidéo maximale : 30 fps

Source: Canon®

Annexe n°2 : Bruit de mesure : données et régression polynomiale.

Tableau 7: bruit de mesure en fonction de la distance et de la couleur de la surface.

<i>Distance (m)</i>	<i>Sigma Black (mm)</i>	<i>Sigma Gray (mm)</i>	<i>Sigma White (mm)</i>	<i>Sigma Mean (mm)</i>
10	0.7	0.6	0.4	0.6
25	1.4	0.9	0.7	1.0
50	3.8	1.7	1.1	2.2
100	14.1	5.4	2.8	7.4

Caractéristiques de la régression polynomiale de degrés 2 passant par les mesures moyenne de bruit :

$$f(x) = p1 * x^2 + p2 * x + p3$$

Avec:

Coefficients (with 95% confidence bounds):

$p1 = 0.0007046$ (0.0003013, 0.001108)

$p2 = -0.002234$ (-0.04905, 0.04459)

$p3 = 0.5737$ (-0.3924, 1.54)

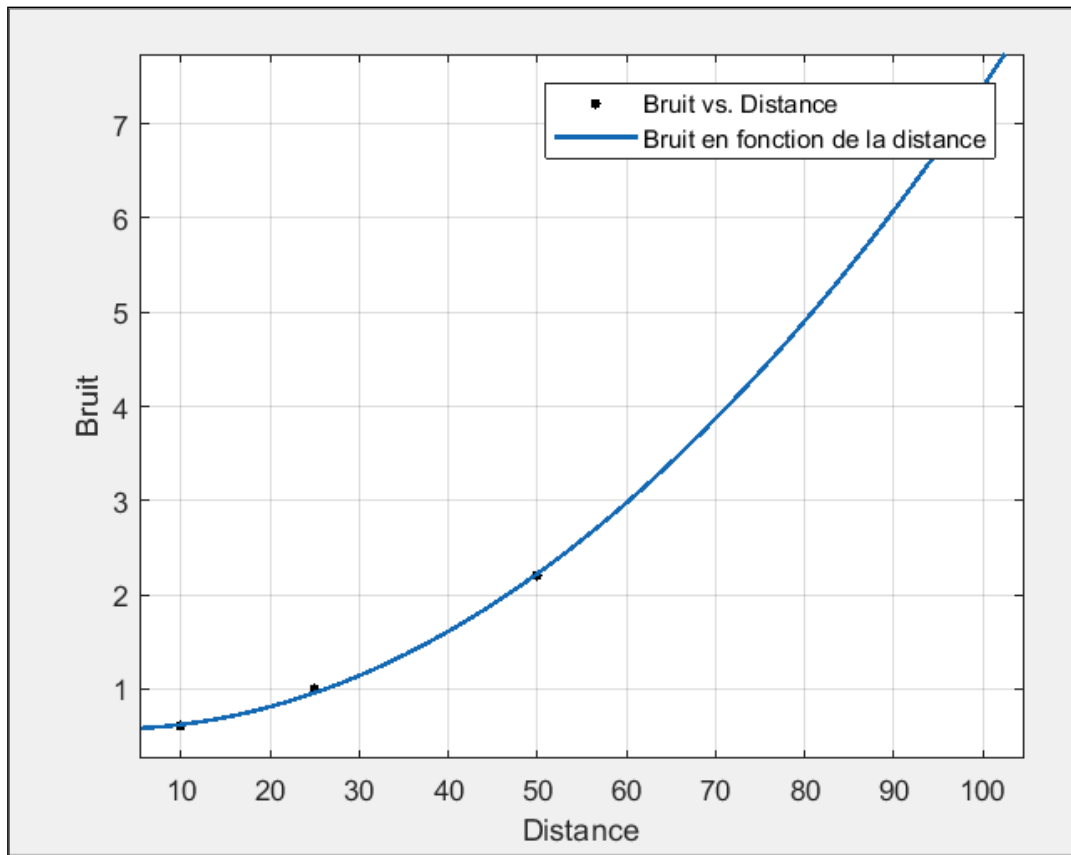
Goodness of fit:

SSE: 0.002787

R-square : 0.9999

Adjusted R-square: 0.9997

RMSE : 0,05279



Annexe n°3 : Calcul de distance et précisions : script MATLAB

Script:

```
%% Détermination des précisions de mesure pour méthodes d'acquisition
disp('LEICA TPS1205');
    disp('Distance de mesure (m)');
        d_tps=2 % Marge d'erreur de 0.5 m par rapport à la dist. min (= 1.5m)

    disp('Erreur distance sur miniprisme (m)');
        e_mini= 0.001+0.0015*(d_tps/1000)

    disp('Erreur distance sur tape reflechissant (m)');
        e_tape=0.005+0.002*(d_tps/1000)

disp('LEICA P30');
% Calcul de resolution à 10 m optimale pour une résolution sur objet de 1mmm
res_10=[0.0500,0.0250,0.0125,0.0063,0.0031,0.0016,0.0008]; % Resolution à 10 m
(en m) --> Test resolution 50, 25, 12.5 ,6.3, 3.1, 1.6, 0.8 mm
LEICA_P30=zeros(7,4);
for i=1:length(res_10)
    d_p30=100; % Initialisation de la distance de mesure(m)
    res_d=1; % Initialisation de resolution (m)
    e_ang_hz=0.0000387851; % Erreur angulaire HZ(rad) = 8"
    e_ang_v=0.0000387851; % Erreur angulaire V (rad) = 8"
    while res_d>=0.001
        d_p30=d_p30-0.005; % incrémentation degressive sinon ok dès 40
mais bruit de proximité important et pas quantifiable
        res_d=( res_10(i)/10)*d_p30; % Resolution à d_p30 (m)
        e_ang_d_hz=atan(e_ang_hz)*d_p30; % Erreur dist du à erreur
angulaire hz à d_p30 (m)
        e_ang_d_v=atan(e_ang_v)*d_p30; % Erreur dist du à erreur
angulaire v à d_p30 (m)
        e_dist=0.0012+0.010*(d_p30/1000); % Erreur de linéarité
        % BRUIT Fct Bruit basé sur donnée mail de Florent Fonction x^2
        % regression à la casio GRAPH 35+
        e_bruit_d=(0.0007046*(d_p30)^2+(-0.002234)*d_p30+0.5737)/1000;

        RMS_d=sqrt((e_ang_d_hz)^2+(e_ang_d_v)^2+(e_dist)^2+(e_bruit_d)^2);
        % Erreur
    end
    LEICA_P30(i,1)=res_10(i);
    LEICA_P30(i,2)=d_p30;
    LEICA_P30(i,3)=res_d;
    LEICA_P30(i,4)=RMS_d;
end
LEICA_P30 % Tableau : Resolution à 10 , Distance si resolution sur objet =
0.001 mm , Erreur RMS

disp('CANON 5D MARK III + 24mm');
    L_image=1920*0.001;% Largeur de image (Resolution vidéo de 1080p et
resolution voulue de 1mm)
    dist_focal=0.024; % Objectif focal fixe de 24 mm
    L_capteur=0.036; % Largeur du capteur
    d_canon=(dist_focal/L_capteur)*L_image% Distance max de prise de vue pour
1 mm de resolution
```

Annexe n°4 : Classification des prises de vues : script MATLAB

Script:

```
for i=1:31
    cd 'C:/Users/Pierre/Desktop/CLASSIF'
    folder=sprintf('C:/Users/Pierre/Desktop/CLASSIF/%d',i)
    mkdir (folder)
    for j=1:17
        cd 'C:/Users/Pierre/Desktop/CLASSIF/ALL_pose'
        name_jpg=sprintf('%d (%d).jpg',j,i);
        A=imread(name_jpg);
        cd (folder)
        imwrite(A,sprintf('%d_%d.jpg',i,j));
    end
end
```

I. Settings

A. Positioning mode/scene:

- Positioning mode: Automatic
- scene: vertical

B. Main settings:

- Key frame ratio: Normal
- QR codes extraction: Disabled
- Pair selection mode: Sequence
- Blockwise color equalization: Enabled
- Pair selection max\min distance: 3
- Component construction mode: OnePass

C. Estimation policies:

- Tri Point: Computes
- Feature: Computes
- Rotation: Computes
- Floor-group estimation mode: OnePass
- Foot length: Adjust
- Fisheye focal: Keep
- Fisheye zoom: Adjust
- Roll/dn tilt: Adjust
- Tangential distortion: Keep
- Aspect ratio: Keep
- Sliver: Keep
- Fish-eye direction: Keep
- Estimation group: PerPhotoGroup

II. Results

A. Global

- Positioning mode: Automatic
- scene: vertical

Errors

	Before serial triangulation					After serial triangulation				
Type	Number of points	Median reprojection error [px]	RMS of reprojection errors [px]	RMS of distances to rays [μm]	RMS of 3D horizontal vertical error [μm]	Number of points	Median reprojection error [px]	RMS of reprojection errors [px]	RMS of distances to rays [μm]	RMS of 3D horizontal vertical error [μm]
Automatic, 3x point	0					1198	0.33	0.59	0.003	

NB: Horizontal and vertical error are given according to the following spatial reference system: Curvilinear

C. Connections

	Before serial triangulation			After serial triangulation		
Number of tied pairs	Median number of tied pairs per photo			Median number of connected photos per photo		
45	6			8		

B. Per photo:

Photo		Before serial triangulation			After serial triangulation		
Photogroup	File name	Number of key points	RMS of reprojection errors [px]	RMS of distances to rays [μm]	Number of points	RMS of reprojection errors [px]	RMS of distances to rays [μm]
0	15_1.jpg	4755	0		106	0.56	0.003
0	15_2.jpg	4784	0		116	0.66	0.006
0	15_3.jpg	5337	0		132	0.65	0.003
0	15_4.jpg	5031	0		108	0.67	0.003
0	15_5.jpg	6036	0		84	0.38	0.004
0	15_6.jpg	4902	0		206	0.45	0.004
0	15_7.jpg	4569	0		315	0.57	0.006
0	15_8.jpg	5431	0		419	0.62	0.006
0	15_9.jpg	4576	0		468	0.64	0.006
0	15_10.jpg	5126	0		480	0.60	0.003
0	15_11.jpg	4961	0		354	0.51	0.003
0	15_12.jpg	4415	0		341	0.54	0.003
0	15_13.jpg	5039	0		238	0.51	0.003
0	15_14.jpg	4530	0		234	0.61	0.006

-Connections

C. Per point:

NB: Horizontal and vertical errors are given according to each control point respective spatial reference systems

- Automatic tie points:

		Before aerotriangulation				After aerotriangulation			
Median number of laser points per photo	1198	3	234	0.59	0.003	3	246	0.71	0.002
	4838								

--- Positioning mode: Use control points for adjustment

NB: Horizontal and vertical errors are given according to the following spatial reference system: Cartesian

- Connections

B. Per photo:

- Automatic tie points

Photo		Before service/gulison			After service/gulison		
Photogroup	File name points	Number of points	RMS of reprojection errors [µm]	RMS of distance: to ray [µ]	Number of points	RMS of reprojection errors [µm]	RMS of distance: to ray [µ]
0	15_1.jpg	4755	1.06	0.56	177	0.98	0.002
0	15_2.jpg	4784	1.16	0.66	183	0.77	0.002
0	15_3.jpg	5337	1.32	0.65	93	0.57	0.001

Annexe n°7 : Export XML des coordonnées des QR Codes : tableau de synthèse

Exemple d'export au format XML des coordonnées des QR Codes simples :

```
<?xml version="1.0" encoding="utf-8"?>
<SurveysData>
  <TiePoints>
    <TiePoint>
      <Name>QR Code #101</Name>
      <Description>CCQR:UTP:101:QR Code Objet 1</Description>
      <Position>
        <x>238.101426999316</x>
        <y>16.9124368886961</y>
        <z>99.3656297015178</z>
      </Position>
    </TiePoint>
    <TiePoint>
      <Name>QR Code #102</Name>
      <Description>CCQR:UTP:102:QR Code Objet 2</Description>
      <Position>
        <x>247.543989260076</x>
        <y>45.4461675796403</y>
        <z>99.3413945071299</z>
      </Position>
    </TiePoint>
    <TiePoint>
      <Name>QR Code #103</Name>
      <Description>CCQR:UTP:103:QR Code Objet 3</Description>
      <Position>
        <x>276.255522009716</x>
        <y>35.9396753221707</y>
        <z>99.0384951691163</z>
      </Position>
    </TiePoint>
    <TiePoint>
      <Name>QR Code #104</Name>
      <Description>CCQR:UTP:104:QR Code Objet 4</Description>
      <Position>
        <x>266.880599148428</x>
        <y>7.35953227840015</y>
        <z>99.0867475218065</z>
      </Position>
    </TiePoint>
  </TiePoints>
</SurveysData>
```

Tableau 8: synthèse des coordonnées extraites des QR Codes simples dans le système de coordonnées local (avec 1 à 7 : les instants du mouvement et 101 à 104 : les 4 marqueurs placés sur l'objet)

Instants	101			102			103			104		
	x	y	z	x	y	z	x	y	z	x	y	z
1	213,8372	432,3231	98,8926	224,3806	460,4767	98,6949	252,7280	449,7920	98,8417	242,1357	421,6801	99,0811
2	206,6472	359,1413	98,7794	209,0223	389,1349	98,5493	239,1373	386,7308	98,7197	236,6769	356,6749	99,0671
3	217,2569	287,4170	98,7523	212,1768	317,1161	98,4094	241,7495	322,4617	98,8200	247,1545	292,8225	98,9129
4	243,9243	222,0754	99,3413	231,8890	249,6604	98,9179	259,6048	261,8025	99,0305	271,4804	234,3767	99,3899
5	262,9537	151,7813	99,3991	255,2379	180,9058	99,3666	284,5592	188,6795	98,9821	292,1177	159,7581	99,3149
6	258,8655	74,9337	99,3830	260,5568	105,0715	99,1892	290,7461	103,4213	98,8320	289,0830	73,1500	99,0566
7	238,1014	16,9124	99,3656	247,5440	45,4462	99,3414	276,2555	35,9397	99,0385	266,8806	7,3595	99,0867

Annexe n°8 : Trajectoires des marqueurs (Spline Cubic 3D) : script MATLAB

Script:

```

csv = csvread('coord.csv');

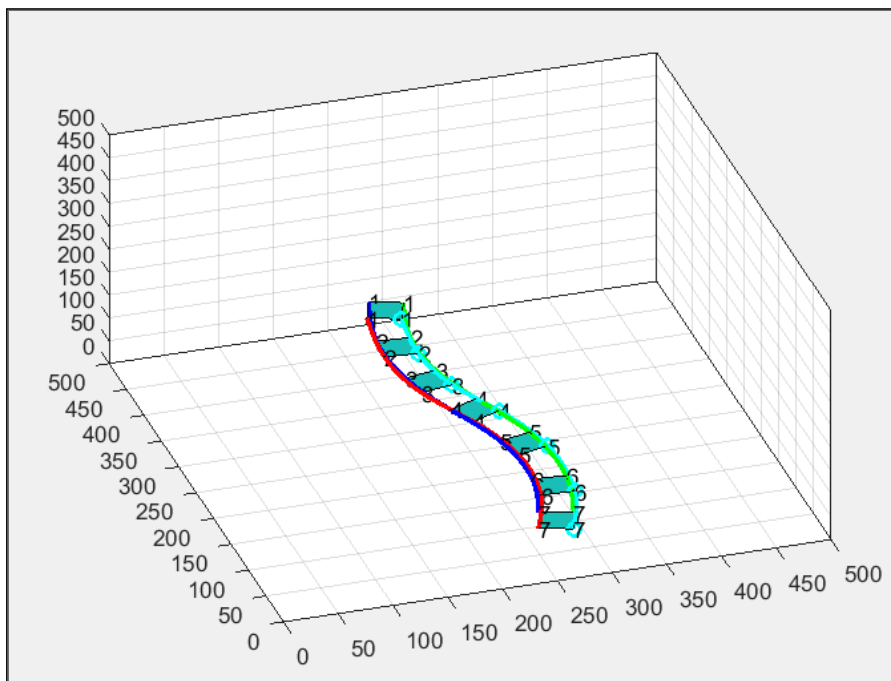
%Spline cubic 3d
npts = 7;

%101
xyz_101 = csv(:,2:4)
xyz_102 = csv(:,5:7)
xyz_103 = csv(:,8:10)
xyz_104 = csv(:,11:13)
scatter3(xyz_101(:,1),xyz_101(:,2),xyz_101(:,3),'r','LineWidth',2);
scatter3(xyz_102(:,1),xyz_102(:,2),xyz_102(:,3),'b','LineWidth',2);
scatter3(xyz_103(:,1),xyz_103(:,2),xyz_103(:,3),'g','LineWidth',2);
scatter3(xyz_104(:,1),xyz_104(:,2),xyz_104(:,3),'c','Linewidth',2);
axis([0 500 0 500 0 500])
text(xyz_101(:,1),xyz_101(:,2),xyz_101(:,3),[repmat(' ',npts,4), num2str((1:npts)')])
text(xyz_102(:,1),xyz_102(:,2),xyz_102(:,3),[repmat(' ',npts,4), num2str((1:npts)')])
text(xyz_103(:,1),xyz_103(:,2),xyz_103(:,3),[repmat(' ',npts,4), num2str((1:npts)')])
text(xyz_104(:,1),xyz_104(:,2),xyz_104(:,3),[repmat(' ',npts,4), num2str((1:npts)')])
ax = gca;
ax.XTick = [0 50 100 150 200 250 300 350 400 450 500];
ax.YTick = [0 50 100 150 200 250 300 350 400 450 500];
ax.ZTick = [0 50 100 150 200 250 300 350 400 450 500];
box on
hold on
fnplt(cscvn(xyz_101(:,[1:end 1]))','r',2)
fnplt(cscvn(xyz_102(:,[1:end 1]))','b',2)
fnplt(cscvn(xyz_103(:,[1:end 1]))','g',2)
fnplt(cscvn(xyz_104(:,[1:end 1]))','c',2)
hold off

for i=1:npts
patch([xyz_101(i,1) xyz_102(i,1) xyz_103(i,1) xyz_104(i,1)],[xyz_101(i,2) xyz_102(i,2)
xyz_103(i,2) xyz_104(i,2)],[xyz_101(i,3) xyz_102(i,3) xyz_103(i,3) xyz_104(i,3)],[1 1
1 1])
end

```

Résultat : la trajectoire de chacun des capteurs dans le système de coordonnées local défini par les QR Codes géoréférencés.



Annexe n°9 : Limite entre mouvement et bruit de mesure : script MATLAB

Script:

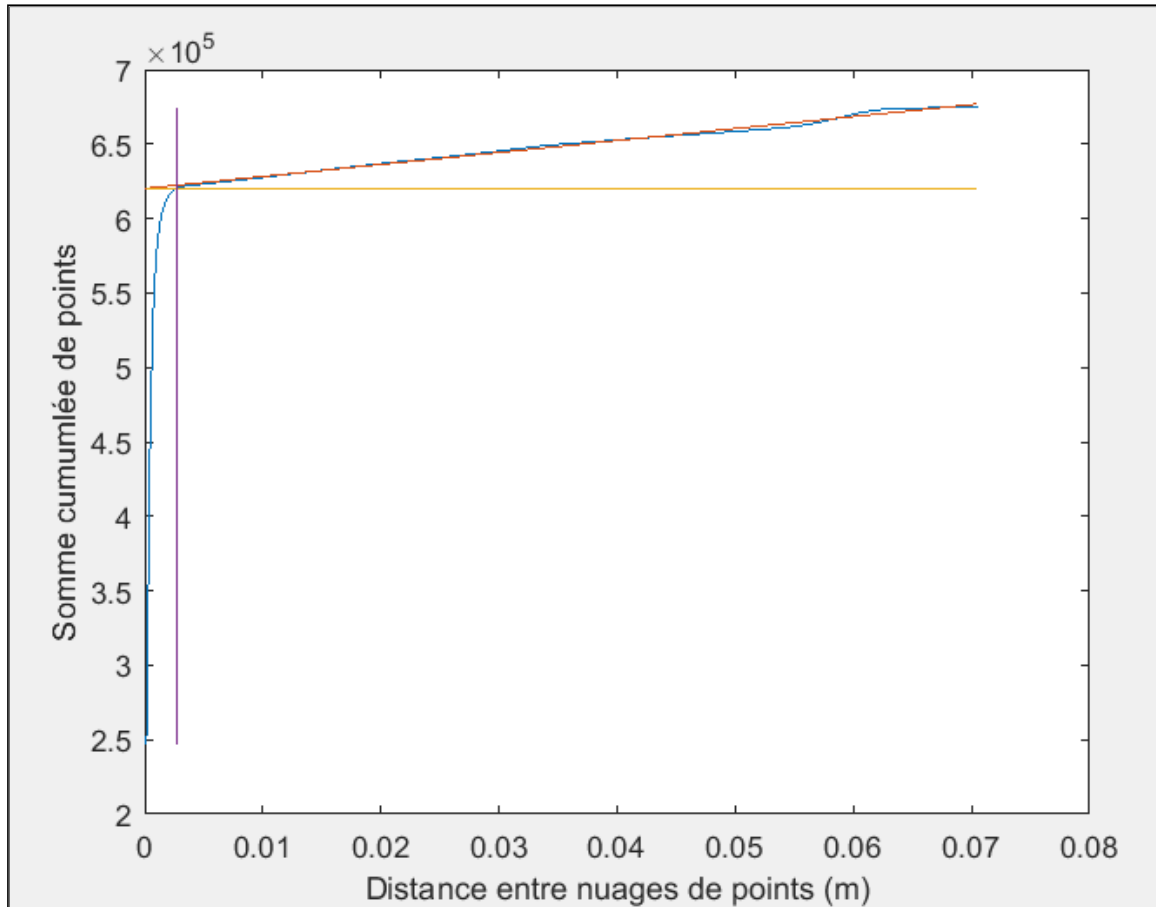
```
A=csvread('Histo_cum.csv');
X=A(:,2); %distance
Y=A(:,1); %n° point cumulé
x = linspace(0,max(X),100);

%Curve fitting application
p1=8.043e05;
p2=6.202e05;
y=p1*x+p2;

% Finding distance on plot
i=1;
while Y(i,1)<p2
    i=i+1;
    abs=(X(i,1)+X(i-1,1))/2
end
i
p2
abs

%Plot
plot(X,Y,x,y,[min(X) max(X)], [p2 p2], [abs abs], [min(Y) max(Y)])
ylabel('Somme cumulée de points'),
xlabel('Distance entre nuages de points (m)'),
ylim([600000 650000])
xlim([0 0.01])
```

Résultat :



Mots-clés

Capture de mouvement, Lasergrammétrie, Photogrammétrie, QR Code, Animation 3D, Context Capture, CloudCompare, 3DS Max.

Keywords

Motion Capture, Lasergrammetry, Photogrammetry, QR Code, 3D Animation, Context Capture, CloudCompare, 3DS Max.

Résumé

Les méthodes 3D de capture du mouvement sont en permanente évolution et utilisent constamment de nouvelles techniques d'acquisition dans le but d'atteindre des rendus toujours plus précis et réalistes. Parmi ces techniques d'acquisition 3D en pleine expansion, la photogrammétrie et la lasergrammétrie font leurs premiers pas en matière de capture de mouvement.

Ce travail de fin d'études a pour objectif d'établir les perspectives d'utilisation de ces deux techniques d'acquisition en matière de capture de mouvement. Pour cela, trois hypothèses de recherches sont élaborées sur base d'une revue de la littérature existante ainsi que sur base de tests et d'essais effectués par nos soins préalablement en laboratoire. Le cœur du travail consiste en la traduction de ces trois hypothèses en trois méthodes d'acquisition, de traitement et de rendu 3D.

Les perspectives de la photogrammétrie sont évaluées au travers de deux méthodes : l'une n'utilisant pas de marqueurs placés sur l'objet en mouvement et l'autre utilisant des marqueurs (QR Codes) placés sur l'objet en mouvement. Les perspectives de la lasergrammétrie sont, quant à elles, évaluées au travers d'une méthode d'extraction de points appartenant au mouvement par analyse de distance entre nuages de points.

La mise en œuvre de ces trois méthodes et la synthèse des résultats obtenus permet de valider ou invalider chacune des trois hypothèses établies, ainsi que de tirer des conclusions sur les perspectives de l'utilisation de la photogrammétrie et de la lasergrammétrie en matière de capture de mouvement.

Resume

3D motion capture methods are constantly evolving and are constantly using new acquisition techniques in order to achieve even more accurate and realistic renderings. Among these rapidly expanding 3D acquisition techniques, photogrammetry and lasergrammetry are taking their first steps into motion capture.

This document aims to establish the prospect of these two techniques of acquisition in motion capture. For this, three research hypotheses are developed based on a review of the existing literature as well as on the basis of tests carried out by us previously in laboratory. Most of the work consists of the translation of these three hypotheses into three methods of acquisition, processing and 3D rendering.

Prospects of photogrammetry are evaluated through two methods: one using no markers placed on the moving object and the other using markers (QR Codes) placed on the moving object. Prospects of lasergrammetry are evaluated through a method of extracting points belonging to the movement with the help of distance analysis between points clouds.

The implementation of these three methods and the synthesis of the results obtained allows us to validate or invalidate each of the three hypotheses established and to draw conclusions on the prospects of photogrammetry and lasergrammetry in motion capture.