



UNIVERSITÉ PARIS 8 - VINCENNES À SAINT-DENIS

Master 1 Informatique

**Mémoire projet tuteuré
Qualification de caméras RGB-D**

**Yasmine BOUDJEMAÏ
Mélanie DE JESUS CORREIA**

Date de soutenance : le 10/01/2020

Organisme : Université Paris 8 Vincennes-Saint-Denis
Tuteur : Farès BELHADJ

Dédicaces

Remerciements

Nous souhaiterions adresser nos sincères remerciements à notre tuteur de projet **M.Farès BELHADJ** pour nous avoir accordées sa confiance et de nous avoir accompagnées tout au long de ce projet.

Toute notre gratitude et toute notre reconnaissance vont à l'encontre de nos deux familles respectives pour leur soutien quotidien inestimable.

Résumé

Dès lors de sa sortie, de nombreux domaines ont connu un essor important à l'aide des technologies RGB-D. À titre d'exemple, on peut citer Natural-Pad¹ qui se sert de la Kinect pour la capture de mouvements pour des besoins de rééducation.

Dans des applications où la précision demeure nécessaire pour de meilleurs résultats, il est indispensable de savoir choisir le capteur qui convient le mieux aux exigences du contexte de l'application. Dans ce document, nous expliquons comment nous comptons réaliser notre application afin de qualifier un ensemble de caméras RGB-D et ce que nous avons pu développé entre-temps. Nous exposons les résultats produits par la première version de notre application, puis discutons de la fiabilité des résultats obtenus.

Dans le premier chapitre, nous proposons un état de l'art des technologies de la caméra RGB-D. Dans le second, nous une maquette de notre future application. En outre, nous détaillons les étapes par lesquelles nous sommes passés pour élaborer le prototype de l'outil répondant à la demande du projet tuteuré, ainsi que les résultats obtenus. Dans le troisième, nous montrons comment procéder pour utiliser un modèle bien précis de caméra afin de démarrer les tests avec ce dernier. Enfin, en conclusion, nous exposons succinctement le bilan du travail réalisé, de même que les perspectives envisagées pour parvenir à nos objectifs.

1. Entreprise, basée à Montpellier, spécialisée dans la création de jeux vidéo ludiques au service des besoins médicaux.

Table des matières

Introduction	1
1 État de l’art	3
1.1 Description d’une caméra RGB-D	3
1.2 Modèles existants	3
1.3 Erreur de mesure	4
1.3.1 Types et sources d’erreurs	5
1.3.2 Instruments de mesure de distances	5
1.3.3 Mesure d’erreurs par une approche théorique	7
1.4 Technologies utilisées pour la depth	8
1.4.1 Infrarouge	8
1.4.2 Stéréo-vision	10
1.5 Kinect V2 et Kinect V3	12
1.5.1 Kinect V2	12
1.5.2 Kinect V3	12
2 Modèle logiciel pour la qualification	13
2.1 Description de l’application	13
2.2 Application développée	13
2.2.1 Première partie : Affichage de la scène captée	13
2.2.2 Seconde partie : Interface graphique	13
3 Cas pratiques de qualification	15
3.1 Conception du modèle réel et du modèle virtuel	15
3.1.1 Conception du modèle réel	15
3.1.2 Conception du modèle virtuel	16
3.2 Preuve de concept (Proof of concept - POC)	16
3.3 Résultats et critique	17
3.3.1 Résultats obtenus	17
3.3.2 Critiques	18

4 Conclusion et Perspectives

19

Table des figures

1.1	Composants caméra RGB-D (Ici Asus Xtion Pro Live).	6
1.2	Composants caméra RGB-D (Ici Asus Xtion Pro Live).	9
1.3	(De gauche à droite) image renvoyée par la caméra RGB, image IR, depth map (une seule caméra), depth map (deux caméras). (figure tirée de l'article de <i>F.Alhwarin, A.Ferrein</i> et <i>I.Scholl</i> ayant comme titre <i>IR Stereo Kinect : Improving Depth Images by Combining StructuredLight with IR Stereo</i>). .	9
1.4	Schéma illustrant la relation entre la distance focale et le champ de vision (fov).	11
2.1	Interface graphique du programme.	14
3.1	Modèle-étalon.	15
3.2	Modèle 3D.	16

Introduction

La caméra RGB-D dans ses débuts, notamment la Kinect, a beaucoup servi le monde du jeux-vidéo. Cependant, depuis un certain nombre d'années, elle contribue activement et fondamentalement dans le développement de multiples applications dans diverses disciplines telles que la médecine, la robotique et plus généralement l'aide à la personne.

Ce projet tuteuré a pour objet la conception d'outils pour la qualification de caméras RGB-D. Il vise à comparer différentes caméras afin de sélectionner la meilleure en fonction du besoin ciblé.

L'application résultante de ce projet devra être en mesure de détecter la colonne vertébrale de la personne posant de dos à la caméra RGB-D, principalement ceux atteints par la scoliose. Ainsi, nous nous attendons à ce qu'elle puisse modéliser la colonne en partant d'un modèle 3D et produire par la suite un rapport des mesures effectuées avec chacune des caméras.

Cependant, comme dans tout projet, des contraintes s'imposent. Parmi ces dernières, nous pouvons citer :

- La localisation de la colonne vertébrale. Pour cela, nous pensons utiliser le Machine Learning pour vérifier si le patient est atteint de scoliose. Si tel est le cas, arriver à détecter les positions des contours de la colonne vertébrale.
- Parvenir à manipuler le modèle 3D afin qu'il corresponde à la colonne vertébrale du patient.

Chapitre 1

État de l'art

Dans ce chapitre, nous allons définir ce qu'est une caméra RGB-D, nous énumérons certains des différents modèles existants, nous abordons brièvement les différentes techniques utilisées pour la récupération de la profondeur par une caméra. Enfin, nous discutons les différences recensées sur la **kinect** V2 et V3.

1.1 Description d'une caméra RGB-D

La caméra RGB-D, aussi appelée capteur RGB-D, est une caméra fournissant en même temps une image couleur et une carte de profondeur caractérisant la distance des objets vus dans l'image. Cela est rendu possible grâce à un capteur RGB et un capteur de profondeur (D pour Depth). C'est principalement ces captures qui vont nous intéresser tout au long des qualifications.

1.2 Modèles existants

Il existe différents modèles de caméras RGB-D. Parmi elles, nous pouvons citer la Kinect et ses différentes versions, Asus Xtion Pro Live, BlasterX Senz3D, Orbbec, Intel RealSens D415, ...

La **Kinect** a fait son apparition en septembre 2008. Elle a été conçue par Microsoft et était destinée pour la console de jeu XBox 360. Elle permettait aux utilisateurs d'interagir avec la console à l'aide d'une NUI¹ en utilisant les mouvements gestuels et une reconnaissance vocale. Elle sera plus tard

1. Natural User Interface (Interface Utilisateur Naturelle), se réfère à une interface utilisateur invisible.

utilisée dans les domaines de la recherche et du développement pour différents secteurs comme le domaine de la médecine, l'industrie automobile, la robotique, l'éducation,

Asus Xtion Pro Live est le modèle de référence que nous utilisons afin d'effectuer les qualifications. Elle utilise la technologie PrimeSense² pour la détection de mouvements.

BlasterX Senz3D a fait son apparition en Septembre 2016. Conçue par Creative, elle a été présentée comme une webcam intelligente basée sur ce qu'il y a de mieux en matière de savoir-faire et de technologie. Elle possède trois lentilles pour capturer les données visuelles : une caméra RVB, une caméra infra-rouge et un projecteur laser. Ces dernières collaborent avec la technologies Intel RealSense pour réagir aux expressions faciales et aux gestes corporels des utilisateurs.

Orbbec Astra Pro fait partie de la série Astra. Elle offre une vision par ordinateur qui permet des dizaines de fonctions telles que la reconnaissance des visages, la reconnaissance des gestes, le suivi du corps humain, la mesure tridimensionnelle, la perception de l'environnement et la reconstruction de cartes en trois dimensions. De plus, elle offre une réactivité haut de gamme, une mesure de la profondeur, des dégradés fluides et des contours précis ainsi que la possibilité de filtrer les pixels de profondeur de faible qualité.

Intel RealSense D415 a un champ de vision standard bien adapté aux applications de haute précision telle que la numérisation 3D. Elle comprend le processeur Intel RealSense Vision D4 offrant une résolution en profondeur élevée, des capacités de longue portée, une technologie d'obturation globale et un large champ de vision. Grâce à ces deux dernières, la caméra offre une perception précise de la profondeur lorsque l'objet est en mouvement ou que l'appareil est en marche. De plus, elle couvre un champ de vision plus large, minimisant ainsi les angles morts.

1.3 Erreur de mesure

L'erreur de mesure, autrefois appelée l'erreur absolue, est un processus qui permet d'évaluer l'écart entre la valeur mesurée et la valeur de référence qui est soit exacte ou connue. L'objectif de la mesure d'erreurs est de jauger à quel degré les deux valeurs sont proches. Dans cette section, nous citons les types ainsi que quelques sources d'erreurs, ensuite, nous faisons un détour sur

2. Connu principalement pour sa licence de conception matérielle et de puce employée dans le mécanisme de détection de mouvements de la Kinect XBox360. Pour plus d'informations, le lecteur peut se référer à ce lien <https://www.crunchbase.com/organization/primesense#section-web-traffic-by-similarweb>.

quelques instruments de mesure de distances ainsi que les moyens théoriques employés à cet effet.

1.3.1 Types et sources d'erreurs

Il existe deux types d'erreurs : aléatoires et systématiques.

Les erreurs aléatoires sont des erreurs dont les valeurs sont incohérentes et imprévisibles même en répétant les observations et en conservant les mêmes paramètres.

Les erreurs dites systématiques, à l'inverse des erreurs aléatoires, sont des erreurs reproductibles, elles demeurent constantes tant que les paramètres restent inchangés.

Ces erreurs peuvent résulter de différentes sources ; elles peuvent être dues à des facteurs environnementaux, à la résolution de l'instrument utilisé, au calibrage de l'appareil employé ou même à des erreurs humaines.

1.3.2 Instruments de mesure de distances

Diverses instruments de mesures de distance sont mis à porté des professionnels tels que les ingénieurs en génie civil, les architectes, les métreurs, les dessinateurs, les topographes, les menuisiers, les maçons, les agents immobiliers, les électriciens, les décorateurs, les experts. Nous nommons certains d'entre eux ci-après.

Distancemètre laser Leica DISTO D8 : cet appareil est pourvu d'un localisateur numérique, d'un écran couleur haute résolution 2,4", d'un capteur d'inclinaison 360° et de lma technologie Bluetooth.

Parmi ses caractéristiques techniques, on peut citer :

- Il possède une porté allant de 0.05m jusqu'à 200m.
- Une précision de $\pm 1mm$.
- Des opérations arithmétiques ainsi que des fonctions mathématiques sont intégrées à l'instrument telles que la fonction Pythagore, la fonction Trapèze ou encore la fonction triangle.
- Image en temps réel
- Un Plug-In d'AutoCAD embarqué dans l'appareil qui permet de dessiner et de planifier avec le logiciel d'AutoCAD.
- Il inclut un logiciel de transfert de données pour le distancemètre **Leica Disto Transfert**.



FIGURE 1.1 – Composants caméra RGB-D (Ici Asus Xtion Pro Live).

- Avec la technologie Bluetooth intégrée, l'utilisateur peut transférer les mesures effectuées vers un ordinateur portable ou un PC sans avoir recours aux câbles. Les données ainsi transférées peuvent être éditées à l'aide de logiciels tels que Word, Excel, AutoCAD ou encore d'autres.

Odomètre vérifiable PCE-MW 2 : Instrument permettant de mesurer différentes zones (ie : chantiers, zones industrielles, zones sportives, ...), des câblages, des tuyauteries. Parmi ses caractéristiques, on peut mentionner :

- Pèse 2200g
- Mesure pouvant aller jusqu'à 99999,9 mètres.
- Permet d'effectuer une addition en se déplaçant vers l'avant et une soustraction en bougeant vers l'arrière.
- Possède une précision de $\pm 1\%$
- La roue est en plastique et permet de résister aux huiles.

Distancemètre laser PCE-LRF 600 : Le design de cet instrument est bien conçu puisqu'il est léger et il permet de le maintenir avec une seule main. De plus, il est facile d'utilisation. De ses caractéristiques, on peut évoquer :

- Il pèse 165g.
- Il a une portée allant de 15m à 600m.
- Une précision de $\pm 1m / \pm 0.1\%$.

- Il possède un laser de type classe 1.
- Résistant aux éclaboussures d'eau
- Il permet une mesure des distances dans diverses situations météorologiques.

1.3.3 Mesure d'erreurs par une approche théorique

A ce jour, on recense de multiples méthodes exploités à cet effet. Nous nous contentons de citer ci-dessous celles qui nous paraissent êtres les plus pertinentes.

RMSE/RMSD : Root Mean Square Error ou encore Root Mean Square Deviation que l'on pourrait traduire par ; la racine carré de la moyenne des erreurs au carré. Elle demeure parmi les plus fréquentes en la matière. C'est une mesure de l'exactitude, qui permet de comparer les erreurs de prévision de différents modèles pour un ensemble de données précis et non entre des ensembles de données. Sa formule s'exprime comme suit :

$$\sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (x_i - \hat{x}_i)^2} \quad (1.1)$$

De par la formule, on peut déduire que le résultat sera toujours positif et qu'obtenir une valeur de 0 (qui n'est généralement non atteint en pratique), témoignerait d'une exactitude parfaite. Ainsi, obtenir des valeurs avoisinant le zéro est positif pour les mesures.

MSE : Mean Square Error que l'on pourrait traduire par ; l'erreur quadratique moyenne. Sa formule est comme suit :

$$\frac{1}{N} \cdot \sum_{i=1}^N (x_i - \hat{x}_i)^2 \quad (1.2)$$

MAE : Mean Absolute Error qui pourrait être traduit par ; erreur moyenne absolue. Son équation est comme ci-dessous :

$$\frac{1}{N} \cdot \sum_{i=1}^N |x_i - \hat{x}_i| \quad (1.3)$$

MAPE : Mean Absolute Percentage Error qui pourrait être traduit par ; le pourcentage de l'erreur moyenne absolue. Elle fait partie également des moyens de mesures d'erreur les plus fréquents. A noter que la

valeur observée ne doit être aucunement nulle. La formule est comme ci-après :

$$\frac{1}{N} \cdot \sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (1.4)$$

Avec :

x_i , la valeur observée à la i ème observation.

\hat{x}_i , la valeur exacte à la i ème observation.

N , représente le nombre d'observations

1.4 Technologies utilisées pour la depth d'une caméra RGB-D

Des méthodes employées pour la récupération de la depth, on peut en citer deux principales.

1.4.1 Infrarouge

Le projecteur infrarouge de certaines caméras le possédant qu'on peut voir sur la figure 1.2 projette un spectre infrarouge sur la scène captée. Le motif produit sur cette dernière sera capté par la caméra infrarouge et sera par la suite comparé à la base de données de motifs de référence stockés au préalable dans la caméra. Ces derniers seront indispensables pour la mesure de la profondeur de chaque pixel.

Par la suite, les valeurs obtenues seront corrélées à un capteur RGB qu'on peut apercevoir sur la figure. Ces données pourront être représentées par un nuage de points³.

Toutefois, ce type de caméras possède certaines restrictions. Parmi ces dernières, on peut citer :

- Distances de mesure limitées.
- Problèmes de calculs des informations de profondeur à l'encontre de surfaces brillantes, très mates, transparentes, réfléchissantes ou encore envers des objets absorbants.
- Interférence des motifs (patterns) infrarouges si présence de plusieurs caméras RGB-D de même type. En effet, chaque capteur visualisera ses propres motifs ainsi que ceux des autres caméras présentes et

3. Est une représentation des points de coordonnées tridimensionnelles dont chacun peut avoir des attributs qui lui sont propres.

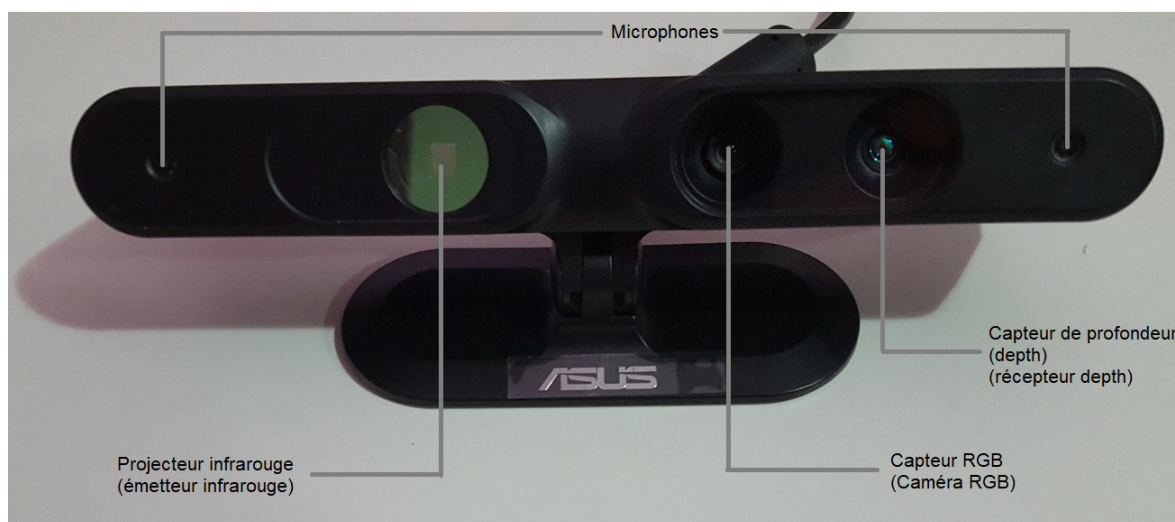


FIGURE 1.2 – Composants caméra RGB-D (Ici Asus Xtion Pro Live).

ne saura distinguer les siens des autres qui se chevauchent. De cela découle une perte considérable d'informations de profondeur comme nous pouvons l'observer sur la figure 1.3.



FIGURE 1.3 – (De gauche à droite) image renvoyée par la caméra RGB, image IR, depth map (une seule caméra), depth map (deux caméras). (figure tirée de l'article de *F.Alhwarin, A.Ferrein et I.Scholl* ayant comme titre *IR Stereo Kinect : Improving Depth Images by Combining StructuredLight with IR Stereo*).

Cependant, des travaux de recherches ont été conduits afin d'y remédier. Parmi eux, on peut citer le travail réalisé par l'équipe de *Rafibakhsh* ([**RAFIBAKHSH15**]) qui recommande de laisser un angle de 35° entre deux caméras suspendues à la même hauteur en considérant les scènes captées dans de bonnes conditions et une interférence très faible. *Maimone* et *Fuchs* [**MaimoneFuchs15**] proposent un algorithme de remplissage et de lissage en modifiant le filtre médian aux

zones trouées à l'exception des bords. Quant à *F. Kenton Musgrave, Craig E.* et *Robert S. Mace* [**KentonCraigMace12**], ils appliquent une certaine quantité minimale de mouvements (en utilisant des composants matériels supplémentaires) à certains capteurs de sorte que chacun puisse voir son propre motif infrarouge de façon nette et une version floue des motifs de ses voisins.

1.4.2 Stéréo-vision

Une caméra stéréoscopique est un appareil qui contient deux voire plus de capteurs d'images. Ceci nous rappelle la vision binoculaire humaine. En effet, ce mécanisme permet au système nerveux central de percevoir simultanément les images issues de chaque œil envoyées sous forme de signaux. Ainsi, il sera en mesure de se servir de ces différences (entre les deux images) pour permettre une vision stéréoscopique pour la perception de relief et une mesure des distances en utilisant la triangulation⁴. Le concept que nous venons d'expliquer est appelé disparité stéréoscopique⁵.

La stéréo-vision sert principalement à reconstituer la scène observée sous forme de modèle 3D.

Carte de disparités

Le principe de la disparité est une approche du mécanisme humain vu précédemment. Il consiste en la différence entre les coordonnées pixels d'un point bidimensionnel d'une image et celles de son correspondant (présent sur une autre image prise au même moment). Ainsi, en appliquant le même traitement sur tous les pixels correspondants, on obtient la carte des disparités. Une carte est dite éparsée lorsqu'une disparité est associée à quelques pixels et lorsque cette dernière est associée à chaque pixel, on dit d'elle qu'elle est dense.

Une formule pour calculer la profondeur en fonction de la disparité s'obtient comme suit :

$$z = \frac{B \cdot f}{d} \quad (1.5)$$

Avec :

4. Approche géométrique permettant une mesure des distances. Le lecteur peut consulter cette page web pour de plus amples informations : <https://fr.wikipedia.org/wiki/Triangulation>.

5. Différence dans la localisation d'un objet perçu par l'œil gauche et l'œil droit résultant de la séparation horizontale des yeux dont certaines caméras essaient de s'approprier la technique afin de récupérer la profondeur.

z représentant la profondeur.

B *Baseline* représente la distance séparant les deux capteurs.

f La distance focale en pixels calculée comme montré ci-après en nous servant de la figure 1.4 .

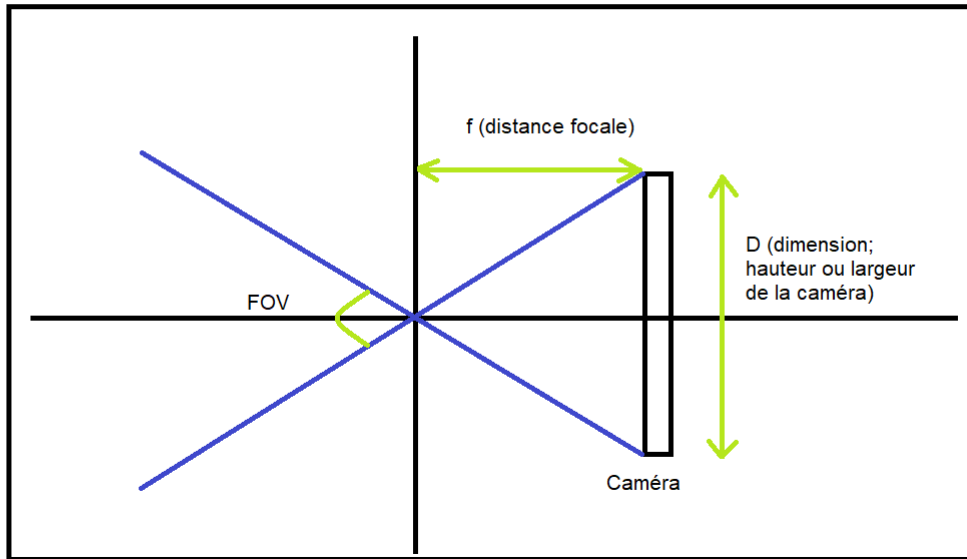


FIGURE 1.4 – Schéma illustrant la relation entre la distance focale et le champ de vision (fov).

$$\tan\left(\frac{fov}{2}\right) = \frac{D}{2.f} \Leftrightarrow f = \frac{D}{2. \tan\left(\frac{fov}{2}\right)} \quad (1.6)$$

Avec :

D dimension qui peut représenter soit la largeur soit la hauteur en fonction du fov de la caméra employé.

fov (Field Of View) représente l'angle du champ de vision (soit horizontal

ou vertical).

f représente la distance focale.

1.5 Kinect V2 et Kinect V3

1.5.1 Kinect V2

Kinect V2 se sert de la méthode TOF (ie : Time Of Flight ou autrement dit Temps De Vol) pour générer la carte de profondeur. Cette technique se base sur la différence de temps entre l'émission d'un faisceau lumineux et son retour après réflexion sur un objet.

La distance est calculée comme suit :

$$d = c \cdot \frac{\Delta t}{2}, \text{ avec } c \text{ la vitesse de la lumière dans l'air.}$$

Elle permet une bien meilleure précision même dans le noir que sa version précédente (**Kinect V1**).

1.5.2 Kinect V3

Kinect V3, baptisée **Microsoft Azure Kinect V3**, tout comme la précédente version, se base aussi sur la technologie TOF. Elle comprend un capteur RGB de 12 Mp, un capteur de depth de 1Mp avec un fov réglable en large ou réduit ainsi que 7 microphones intégrés.

Kinect V2 Vs Kinect V3

Kinect V3 est beaucoup plus légère et plus petite que la **V2**. Elle possède entre autre plus de microphones que la précédente. Elle a été conçue afin d'être principalement utilisée avec **Azure** , le service cloud de **Microsoft**. C'est sans doute ce qui la démarque le plus de sa prédécesseure. En effet, ce service lui permet d'effectuer une partie des calculs. En outre, elle bénéficie des **Cognitive Services**, autrement dit de l'intelligence artificielle pourra être incluse dans les applications créées.

Chapitre 2

Modèle logiciel pour la qualification de caméra RGB-D

Dans ce passage, nous allons décrire notre outil que nous voudrions concevoir pour qualifier les caméras. Par la suite, nous montrons ce que nous avons pu commencé à faire pour débiter ce projet.

2.1 Description de l'application

2.2 Application développée

Notre application est composée d'une fenêtre qui comprend une partie affichage de la scène filmée en temps réel ainsi qu'une partie interface graphique qui nous permet de communiquer avec le programme. Nous parlerons un peu plus en détail de ces parties dans les sous-sections à venir.

2.2.1 Première partie : Affichage de la scène captée

Pour cet effet, nous sommes partis d'un programme fourni par *M. Farès BELHADJ* auquel nous avons rajouté des choses pour la conception de l'outil. Il comprend l'utilisation de la bibliothèque OpenGL, GL4Dummies ainsi que la bibliothèque OpenNI2 pour établir la connexion avec la caméra RGB-D pour nous fournir par la suite la carte de couleurs et la carte de profondeurs.

2.2.2 Seconde partie : Interface graphique

Pour l'interface graphique, nous avons exploité la bibliothèque Dear ImGui disponible sur GitHub sur le lien suivant <https://github.com/ocornut/>

imgui. C'est une bibliothèque assez facile d'utilisation une fois l'étape de l'intégration dans le contexte OpenGL effectué. Cette interface contient un certain nombre de composants. Parmi les composants principaux, nous pouvons citer une liste déroulante qui nous permet de sélectionner le modèle de caméra sur laquelle nous souhaitons réaliser les tests de mesures. Ainsi, une fois le choix de caméra effectué, le programme met automatiquement à jour tous les paramètres présents dans les champs suivants la liste en récupérant les données à partir d'un fichier que nous avons nommé *infoCameras.txt*. Ainsi, ce dernier pourra être exploité pour tout rajout de nouvelles caméras et/ou modification sur les données uniquement. Autres les composants cités ci-dessus, nous noterons la présence de deux cases à cocher comme montré dans la figure 2.1 La première permet une détection du modèle et l'autre de passer les tests de mesures.

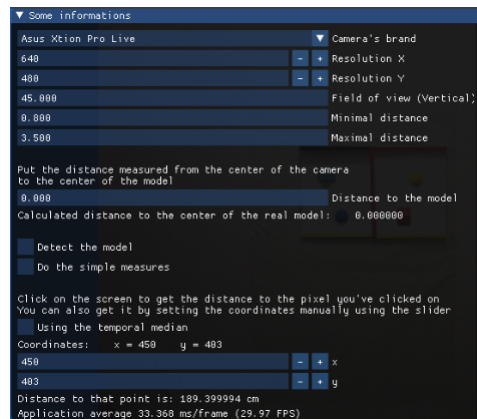


FIGURE 2.1 – Interface graphique du programme.

Chapitre 3

Cas pratiques de qualification

3.1 Conception du modèle réel et du modèle virtuel

3.1.1 Conception du modèle réel

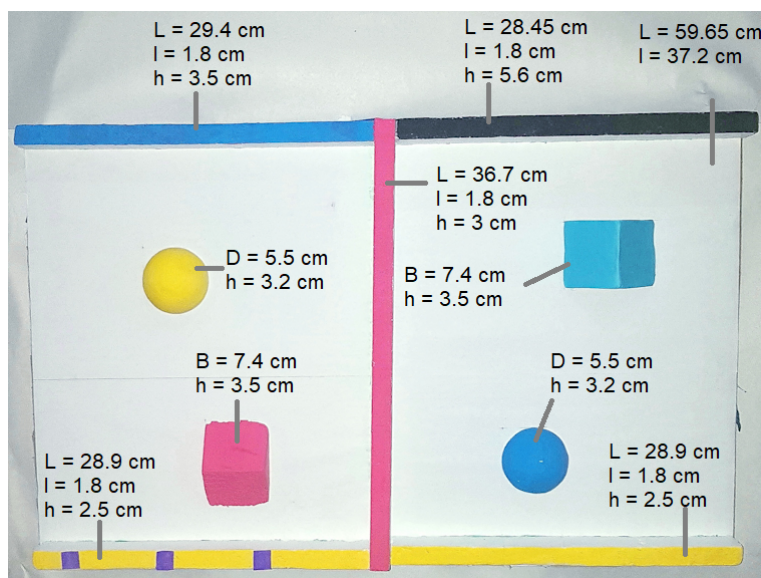


FIGURE 3.1 – Modèle-étalon.

Pour fabriquer le modèle réel, nous avons décidé d'opter pour le polystyrène pour les composants du modèle et du carton pour sa base. Cette

dernière sera par la suite recouverte de papier blanc. Les composants seront peints de couleurs distinctes suivant leur hauteur (les éléments qui possèdent une même hauteur seront peints de la même couleur plus particulièrement les rebords du modèle) sauf les demi-sphères et "pyramides".

Nous avons opté pour le polystyrène car c'est un matériau facilement maniable et que l'on peut trouver assez facilement.

3.1.2 Conception du modèle virtuel

La réalisation du modèle virtuel s'est faite à l'aide du logiciel 123D Design qui est assez facile d'utilisation. Chaque élément du modèle est conçu en fonction des vraies mesures faites sur le modèle réel.

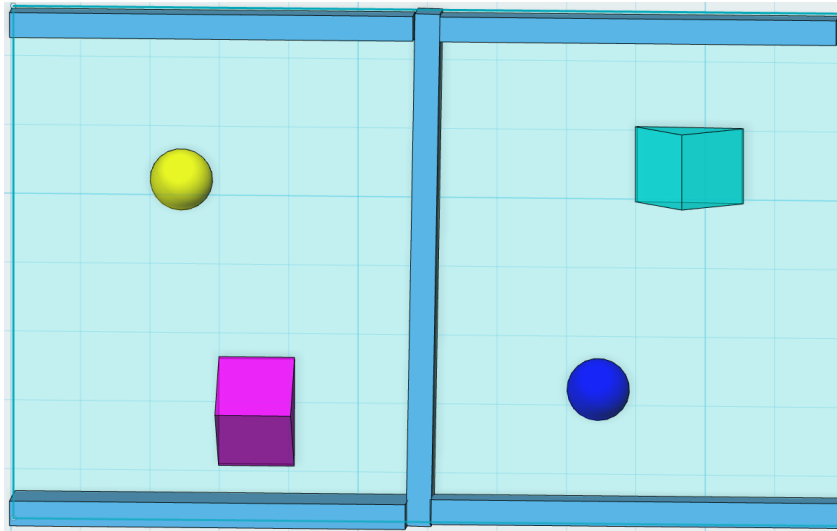


FIGURE 3.2 – Modèle 3D.

3.2 Preuve de concept (Proof of concept - POC)

Dans cette section, nous montrons comment nous avons fait pour configurer la caméra **Asus Xtion Pro Live** afin que le programme puisse effectuer les tests avec cette dernière. Tout d'abord, nous avons édité le fichier *infoCameras.txt* cité dans le chapitre précédent afin de pouvoir y ajouter les informations relatives à cette caméra.

Par la suite, nous n'avons pas eu besoin de rajouter des lignes de code pour la reconnaissance de la caméra (initialisation et récupération de la carte de

couleurs et de la carte de profondeurs) car ce modèle de caméra est supporté par la bibliothèque OpenNI2.

Enfin, il nous suffira de sélectionner, à partir de l'interface de l'application, le modèle du capteur branché et cocher sur les cases *Detect the model* pour une détection du modèle-étalon et *Do the simple measures* pour démarrer les tests de mesures et rédiger le rapport comme expliqué précédemment.

3.3 Résultats et critique

3.3.1 Résultats obtenus

Ces résultats sont extraits du rapport rédigé par le programme après réalisation des tests avec la caméra Asus Xtion Pro Live avec une distance de 1.936m (distance entre la caméra et le modèle). Nous avons créé la base de données d'images templates suivant cette distance.

object	RMSE	Bias	MAPE
Whole model	0.645805	0.464625	0.445257
Blue piece	0.163108	0.0310855	0.0294527
Black piece	0.140734	0.0231826	0.0215494
Vertical magenta piece	0.301483	0.130255	-0.130255
Yellow piece 1	0.132736	0.019909	0.018747
Yellow piece 2	0.122078	0.0168138	0.0156804
Yellow half-sphere	0.12617	0.0215568	-0.0215568
Magenta pyramid	0.12583	0.0232983	-0.0232983
Blue pyramid	0.136497	0.0254308	-0.0254308
Blue half-sphere	0.110225	0.0178525	-0.0178525

Ci-dessous le tableau résumant le taux d'erreur approximatif obtenu sur des distances différentes, et ce, en prenant en considération que la valeur RMSE calculée.

	Distances (m)							
object	1.936	1.843	1.748	1.639	1.550	1.445	1.359	1.259
Whole model	65%	69%	71%	76%	80%	84%	89%	94%
Blue piece	16%	17%	17%	18%	18%	21%	21%	22%
Black piece	14%	15%	15%	14%	17%	16%	18%	18%
Vertical magenta piece	30%	30%	30%	31%	31%	31%	32%	38%
Yellow piece 1	13%	13%	13%	13%	13%	13%	13%	13%
Yellow piece 2	12%	12%	12%	13%	13%	14%	12%	15%
Yellow half-sphere	13%	12%	12%	14%	15%	14%	13%	15%
Magenta pyramid	13%	14%	14%	14%	14%	15%	16%	16%
Blue pyramid	14%	14%	13%	13%	14%	16%	16%	17%
Blue half-sphere	11%	11%	12%	12%	13%	14%	15%	14%
Average rate	20.1%	20.7%	20.9%	21.8%	22.8%	23.8%	24.5%	26.2%

On peut observer aisément un taux d'erreur de plus en plus élevé pour le modèle en entier. Pour les autres éléments, le taux reste pratiquement constant.

3.3.2 Critiques

Comme nous doutons de la méthode employée pour extraire les données OpenGL correspondantes aux données de la depth de la caméra (données extraites aux mêmes pixels), nous pensons qu'il est tout à fait normal que le taux soit si élevé. Toutefois, comme après chaque changement de distance (ie : la caméra se rapproche du modèle observé), le modèle 3D subit un changement d'échelle afin qu'il soit superposé au modèle réel. Ainsi, la valeur de la coordonnée z de chaque pixel du modèle 3D subit pareillement ce changement (devient de plus en plus grande).

Chapitre 4

Conclusion et Perspectives

Dans ce rapport, nous avons présenté l'outil que nous avons commencé à développer pour la qualification de caméras RGB-D. Comme nous avons pu le constater, cet outil recense des erreurs au niveau des mesures. En effet, des données OpenGL qui ne sont pas forcément fiables, particulièrement lors de la mise à l'échelle, influent considérablement sur les résultats que nous avons vus lors du précédent chapitre.

En perspectives, nous souhaiterions vérifier de manière plus pointilleuse les données OpenGL extraites et améliorer la mise à l'échelle du modèle 3D de manière à ne pas impacter le taux d'erreur. Dans ce cas, la qualification pourra être effectuée dans des conditions permettant une meilleure comparaison entre les différents modèles de caméras.