



UNIVERSITÉ PARIS 8 - VINCENNES À SAINT-DENIS

Master 1 Informatique

Mémoire projet tuteuré
Qualification de caméras RGB-D

Yasmine BOUDJEMAÏ
Mélanie DE JESUS CORREIA

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

Dédicaces

Remerciements

Résumé

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 Taux d’erreur (RMSE)	4
1.4 Théoriquement (RMSE)	4
2 Modèle logiciel pour la qualification	5
2.1 Description de l’application	5
2.1.1 Première partie : Affichage de la scène captée	5
2.1.2 Seconde partie : Interface graphique	5
3 Cas pratiques de qualification	7
3.1 Conception du modèle réel et du modèle virtuel	7
3.1.1 Conception du modèle réel	7
3.1.2 Conception du modèle virtuel	7
3.2 Preuve de concept (Proof of concept - POC)	7
3.3 Résultats et critique	8
3.3.1 Résultats obtenus	8
3.3.2 Critiques	9
4 Conclusion et Perspectives	11

Table des figures

Introduction

Introduction a faire à la fin.

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. Elles offrent 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, elles offrent 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 telles 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 derniers, 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 Taux d'erreur (RMSE)

1.4 Théoriquement (RMSE)

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>.

Chapitre 2

Modèle logiciel pour la qualification de caméra RGB-D

Dans ce passage, nous allons décrire notre outil conçu pour qualifier les caméras.

2.1 Description de l'application

2.1.1 Première partie : Affichage de la scène captée

2.1.2 Seconde partie : Interface graphique

Chapitre 3

Cas pratiques de qualification

3.1 Conception du modèle réel et du modèle virtuel

3.1.1 Conception du modèle réel

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.

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