

Rapport animation Alexandre Meyer

Barros Mikel-ange

I) TP

I.1) Présentation du TP

L'objectif de ce TP est de simuler l'animation d'un homme dans différents déplacements et mouvements. Pour cela nous créerons un personnage animé à partir d'un modèle défini dans des fichiers bvh. L'objectif de ce tp est de mettre en place des animations simples avec interruption possible.

I.1) Choix d'architecture

Afin de symboliser cela nous utilisons une machine à état composée de 12 états. En effet, nous avons manuellement défini les différentes animations et le graphe d'exécution plutôt que de laisser l'ordinateur le générer automatiquement. Pour cela je crée un bvh par animation voulue et je change de bvh en fonction de l'animation désirée.

Ainsi dans l'ordre de priorité nous avons :

L'ordre de priorité défini qu'une ligne bloque toutes les lignes du dessous

- si on appuie sur X on donne un coup de pied
- si on appuie sur I on fait un salto
- si on appuie sur U on salue en s'inclinant
- si on appuie sur M on s'incline
- si on appuie sur L on salue de la main
- si on appuie sur O on saute
- si on appuie sur P on tappe dans nos mains
- si on appuie sur T on donne deux coups de poings
- si on appuie sur R on boit un verre
- si on appuie sur Z ou S et qu'on appuie pas sur F on marche
- si on appuie sur Z ou S et qu'on appuie sur F on court
- si on appuie sur rien on reste en attente

Toutes les touches sauf F,Z,S,O bloquent les déplacements et les rotations.

Cette approche nous permet d'obtenir simplement une machine à état permettant d'exécuter différentes animations en fonction des désirs de l'utilisateur. Cependant la transition entre les animations est brusque et donne des résultats peu utilisables dans une vraie application ou un jeu.

II) Capture et interactions temps réel dans les jeux vidéos

II.1) Introduction

Initiée par la Wii la capture des mouvements des joueurs a atteint son paroxysme avec le kinect et le real sense. Ainsi de plus en plus de jeux utilisent ces techniques afin d'exploiter le corps du joueur dans des mécanismes de plus en plus inventifs. De plus la précision d'une telle détection permet de nos jours de reproduire le joueur dans différentes applications. De telles technologies possibles seulement en laboratoire il y a encore quinze ans sont donc aujourd'hui largement démocratisées et on peut donc se demander comment une telle prouesse a pu être réalisée et comment fonctionne le dit mécanisme. Pour étudier cela, nous reviendrons tout d'abord sur l'histoire de cette technologie, ensuite nous regarderons plus en détails son fonctionnement et son utilisation actuellement et enfin nous explorerons des pistes d'évolutions pour la technologie.

II.2) Histoire de la capture de mouvements et de corps

La capture de mouvement dans les jeux vidéo est quelque chose d'assez récent, elle a vu le jour véritablement en 2008 avec la sortie de la Wii. À cette époque, pas question de caméra, la technologie utilisée était basée sur un capteur (deux avec le Nunchuck) et permettait de détecter des mouvements simples des bras. Malgré une précision assez faible et des résultats pas toujours cohérents cette première approche a permis d'imaginer la possibilité de faire interagir le corps du joueur directement avec le jeu.

Cependant, comme dit plus haut la Wii n'était pas particulièrement précise et une première amélioration de cette technologie est venue de deux constructeurs à peu près en même temps. Nintendo avec le Wii motion plus et Sony avec son Playstation Move. Ces deux appareils corrigeaient le plus gros défaut de la Wii : la précision. Ainsi grâce à ces deux appareils nous pouvions maintenant détecter avec précision la position du joueur et les mouvements de ses bras. Parmi les jeux bénéficiant de cette technologie nous pouvons noter The legend of Zelda Skyward's sword sur Wii et LittleBigPlanet2 sur Playstation.

En parallèle de ces deux entreprises, une autre entreprise s'est lancée à peu près au même moment dans le marché de la détection de mouvements mais avec une tout autre technologie. En effet, en 2010, Microsoft lançait sa Kinect. Un accessoire pour sa console Xbox 360 qui permettait de détecter les mouvements de l'ensemble du corps à l'aide d'une caméra. Peinant à convaincre, notamment à cause d'une précision faible et de nombreux bugs, la Kinect c'est rapidement retrouvée reléguée au niveau de gadget et a peu été exploitée dans les jeux.

Cependant grâce à la Kinect une porte s'était ouverte, celle de la reconnaissance complète des mouvements et de la capture en jeu de ses mouvements. Cette technologie, originalement utilisée dans les studios d'animations et dans les laboratoires, faisait enfin sa sortie pour le grand public et avec un résultat convaincant. Et c'est dans cette porte ouverte par la Kinect que se sont engouffrés Intel et Microsoft avec deux technologies : le realSense sur ordinateur en 2015 et la Kinect 2 sur Xbox one en 2013. Bien plus fiable que la Kinect originale ces nouvelles technologies ont réussi à se démarquer et à se démocratiser dans le domaine du jeu vidéo. Ainsi, on retrouve l'usage de ces technologies dans de nombreux jeux sur ces différentes plateformes et la technologie semble encore avoir de nombreuses années devant elle.

II.3) Fonctionnement des appareils

Que ce soit la Kinect où le realSense l'objectif des deux appareils est de créer une carte des profondeurs de la zone capturée. Et bien que similaire dans le principe et dans le fonctionnement le RealSense et le kinect sont deux appareils à part et nous détaillerons le fonctionnement de chacun des deux appareils. Nous allons commencer par cela avant de revenir sur leur utilisation dans le monde du jeu vidéo.

II.3.a) Real Sense [RL17]

Le realSense est un accessoire Intel composé d'un module de profondeur, d'une caméra couleur et d'un processeur dédié nommé processeur d'imagerie de profondeur. Le module de profondeur est quand à lui composé d'une caméra couleur, de deux caméras détectant la profondeur et nommées « left » et « right » et d'un projecteur Infrarouge . (Figure 1)
Revenons rapidement sur le fonctionnement de ce module de profondeur :

Comme dit plus haut le module de profondeur est composé de quatre autres modules. Chacun de ces modules à un usage spécifique que je vais détailler ici :

-Le module le plus important de la caméra est le double capteur de profondeur. Chacun de ces capteurs va capturer une photo en nuances de gris de la scène. Chacune de ces images servira à créer une carte de profondeur que je détaillerais en présentant le processeur d'imagerie de profondeur.

-Le projecteur AR va quand à lui projeter un modèle infrarouge fixe sur les objets pour augmenter la texture sur les objets ayant une fiable texture.

-Le module de détection de couleurs va recueillir des informations sur les textures et notamment la superposition sur une image de profondeur et la superposition sur un modèle 3D pour la reconstruction.

-Le processeur d'imagerie de profondeur quand à lui va calculer les valeurs de profondeur pour chaque pixel en corrélant les points de l'image de gauche à l'image de droite et en échangeant un point de l'image de droite avec l'image de gauche. Cela permet donc de créer une carte de profondeur sur l'image et ensuite un flux de profondeur.

Figure 1-3: RealSense™ System Block Diagram

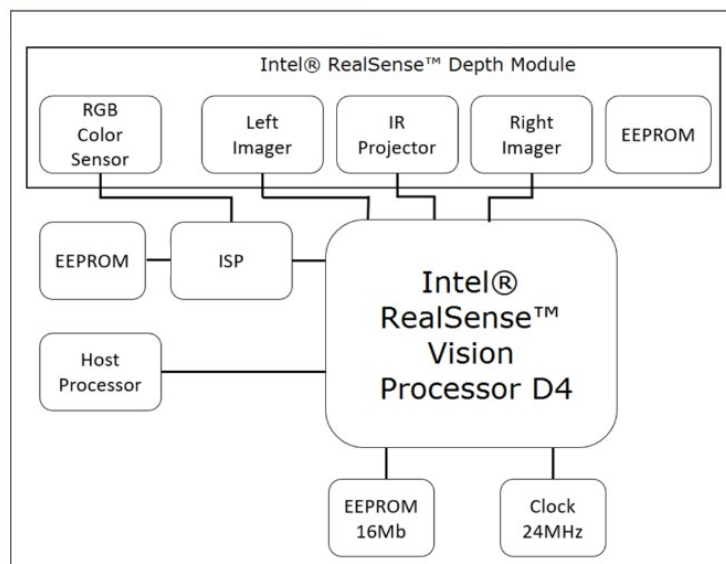


Figure 1 : Schéma de présentation du realSense

II.3.ab) Kinect [MK]

Tout comme pour le RealSense, la Kinect va créer une image de la profondeur de la scène et une image de texture. Mais pour cela la méthode va être légèrement différente sur la Kinect. Si les deux appareils utilisent bien un capteur couleur, la Kinect a remplacé les deux caméras de profondeur par un capteur infrarouge et le projecteur infrarouge par un illuminateur infrarouge. (Figure 2)

Comme au-dessus je vais détailler le fonctionnement du capteur infrarouge et l'illuminateur infrarouge mais je vais laisser de côté le capteur de couleur qui fonctionne comme sur le realSense.

-Le capteur infrarouge va détecter la lumière reflétée par les différents objets de la scène au niveau infrarouge, de plus il va mesurer pour chaque pixel le temps qui a été parcouru par le photon avant de lui revenir. Pour effectuer cette mesure, il va diviser chaque pixel de la caméra en deux. Chacune de ces moitiés va changer constamment entre deux états :

- l'état On : capte la lumière
- l'état Off : ne capte pas la lumière

La différence entre les deux moitiés de pixels est que quand la première moitié est dans l'état on la deuxième est dans l'état off et inversement.

Ainsi il crée la carte de profondeur à partir la distance allé retour parcourue par chaque photon. Afin de limiter l'impact lié à la différence de position entre le capteur infrarouge et l'émetteur infrarouge, on triangule la position de chaque pixel par rapport à la position des deux capteurs. Cette méthode augmente grandement la précision de la Kinect par rapport au realSense. [RLMK]. En effet, là où le realSense calcul les distances et peut donc avoir des erreurs, le kinect lui les mesures.

- L'illuminateur infrarouge quand à lui va comme pour le realSense projeter un motif infrarouge mais il ne va le faire que quand la première moitié des pixels est dans l'état ON.



Figure 2 :Présentation de l'intérieur d'une kinect

II.3.c) Utilisations dans le jeu vidéos

Maintenant que nous avons présenté les deux appareils et leurs composition, revenons sur son utilisation dans le jeu vidéo. Comme dit au-dessus, la caméra realSense tout comme le kinect va créer un flux de profondeur et une image de couleur pouvant se superposer dessus. C'est à partir de ces deux images que l'ordinateur va pouvoir suivre le mouvement et modéliser le personnage en 3D.

Suivi de mouvement :

La caméra doit identifier l'objet dont elle doit suivre le mouvement, pour cela elle utilise la carte de profondeur pour identifier les éléments au premier plan. Une fois cela fait, de nombreux algorithmes tels que celui de Lucas Kanade [LKW] permettent de suivre le mouvement d'un élément.

Reconstruction 3D :

Tout d'abord, la caméra doit déterminer les contours et la forme de la personne à modéliser. Pour cela , on va se servir du flux de profondeur qui va nous permettre de déterminer quels sont les éléments en premier plans et les éléments à suivre. En effet, en fonction de la profondeur du pixel sur l'image et du mouvement de l'image on peut déterminer précisément la forme du corps et ses mouvements. Une fois que l'on sait récupérer ces informations, il faut les recréer. Pour cela on utilise la combinaison entre la carte de profondeur et l'image texturée pour recréer un corps aux bonnes proportions et à la bonne forme.

II.4) Evolution de la technologie

La technologie, bien que très aboutie et permettant de détecter un personnage ou un mouvement, est encore loin de la précision que l'on pourrait retrouver dans un studio ou dans un laboratoire. Afin d'améliorer la précision on pourrait imaginer un ajout de capteur comme la Wiimote de la Wii ou les touch de l'oculus rift. Ce type de manettes avec capteur permettrait de définir avec une bien meilleure précision les mouvements des bras et pourquoi pas des jambes (tout dépendra des manettes et du nombre de capteurs). Avec cette méthode on se rapprocherait beaucoup de la technique de motion capture utilisée dans le jeu vidéo ou le dessin animé d'animation bien que très simplifié.

Une autre méthode d'amélioration viendrait de l'ajout de nouvelles caméras autour de la pièce. On retrouve déjà cela dans certains escape game en réalité virtuelle afin de situer le joueur dans l'espace. En augmentant le nombre de caméras et d'émetteurs IR il serait ainsi plus simple de détecter la perspective ainsi que les dimensions exact des objets. Ainsi on se rapprocherait de la méthode de capture optique : Ascension reactor, sans pour autant devenir aussi complexe. Cependant cette méthode a un défaut majeur, elle demanderait d'avoir une pièce dédiée aux jeux vidéo, là où la kinect donne un résultat très performant et bien plus adaptable.

II.5) Conclusion

Les interactions temps réel dans les jeux vidéo initiées par la Wii ont été grandement amélioré depuis l'arrivée de la kinect et le realSense. Véritable référence en la matière ces deux appareils ont révolutionné le genre et ont permis une reconnaissance très efficace des mouvements et de l'être humain. Et bien que des pistes d'améliorations soient explorés on peut se demander si vu la qualité

des résultats obtenus cela est bien nécessaire. Souvent peu pratique ou complexe à mettre en place, ces améliorations ne seraient pas forcément pratiques pour le grand public et se limiterait sans doute à un public très restreint ou à une utilisation très spécifique.

II.6) Références

[RL17] https://www.mouser.com/pdfdocs/Intel_D400_Series_Datasheet.pdf

[MK] https://anyline.com/wp-content/uploads/2017/10/peter_kinect_presentation.pdf

[LKW] https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Lucas%E2%80%93Kanade

[RLMK] <https://www.youtube.com/watch?v=O0hSzay3fns>