



Université Claude Bernard



DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **13 décembre 2021**

Nom de famille et prénom de l'auteur : **Madame FULOP-BALOGH Beatrix-Emoke**

Titre de la thèse : « *Acquisition multi-vues et rendu de scènes animées* »

Résumé



Les récentes percées technologiques ont conduit à une abondance d'appareils d'enregistrement vidéo conviviaux. De nos jours, les nouveaux modèles de smartphones, par exemple, sont équipés non seulement de plusieurs caméras, mais également de capteurs de profondeur. Cela signifie que tout événement peut facilement être capturé par plusieurs appareils et technologies différents en même temps, et cela soulève des questions sur la façon dont on peut traiter les données afin de restituer une scène 3D significative. La plupart des solutions actuelles se concentrent uniquement sur les scènes statiques, les scanners LiDaR produisent des cartes de profondeur extrêmement précises et les algorithmes stéréo multi-vues peuvent reconstruire une scène en 3D à partir d'une poignée d'images. Cependant, ces idées ne sont pas directement applicables en cas de scènes dynamiques. Les capteurs de profondeur échangent la précision contre la vitesse, ou vice versa, et les méthodes basées sur des images couleur souffrent d'incohérences temporelles ou sont trop exigeantes en termes de calcul.

Dans cette thèse, nous visons à fournir des solutions conviviales pour fusionner des technologies multiples, éventuellement hétérogènes, pour reconstruire et rendre des scènes dynamiques 3D.

Premièrement, nous introduisons un algorithme qui corrige les distorsions produites par de petits mouvements dans les acquisitions de temps de vol et produit une séquence animée corrigée. Pour ce faire, nous combinons un système LiDAR à temps de vol lent mais haute résolution et un capteur de profondeur consommateur rapide mais basse résolution. Nous avons présenté le problème comme un recalage courbevolumique, en voyant le nuage de points LiDAR comme une courbe dans l'espace-temps à 4 dimensions et la vidéo de profondeur à basse résolution capturée comme un volume d'espace-temps à 4 dimensions. Nous convoyons ensuite les détails du nuage de points haute résolution à la vidéo de profondeur en utilisant son flux optique.

Deuxièmement, nous abordons le cas de la reconstruction et du rendu de scènes dynamiques capturées par plusieurs caméras RVB. Dans des contextes occasionnels, les deux problèmes sont difficiles à fusionner : la structure à partir du mouvement (SfM) produit des nuages de points spatio-temporellement instables et parcimonieux, tandis que les algorithmes de rendu qui reposent sur la reconstruction doivent produire des vidéos temporellement cohérentes. Pour relever le défi, nous considérons les deux étapes conjointement. Tout d'abord, pour SfM, nous récupérons des poses de caméra stables, puis nous différons l'exigence de points cohérents dans le temps sur la scène et ne reconstruisons qu'un nuage de points éparés par pas de temps qui est bruité dans l'espace-temps. Deuxièmement, pour le rendu, nous présentons une formulation de diffusion variationnelle sur les profondeurs et les couleurs qui nous permet de faire face de manière robuste

au bruit en appliquant une cohérence spatio-temporelle via des poids de reprojection par pixel dérivés des vues d'entrée.

Dans l'ensemble, nous montrons que notre travail a contribué à la compréhension de l'acquisition et du rendu de scènes dynamiques capturées simplement.