

Sujet de thèse LORIA

Equipe Magrit

Calcul de pose par simulation de points de vue pour la réalité augmentée

Simulation-based viewpoint computation for augmented reality.

Encadrants: Marie-Odile Berger (CR INRIA, HDR) et Frédéric Sur (MdC)

Mots clés: simulation, mise en correspondance, pose

Key-words: simulation, matching, pose computation.

1 Contexte

La réalité augmentée vise à intégrer des objets virtuels dans un flux vidéo du monde réel de manière réaliste. Il est donc nécessaire de calculer avec précision la position de la caméra pour pouvoir projeter les objets virtuels de manière cohérente dans la scène. La pose d'une caméra est le plus souvent calculée en mettant en correspondance des éléments extraits de l'image courante avec un modèle tri-dimensionnel texturé de la scène acquis antérieurement. Beaucoup de méthodes permettant la mise en correspondance modèle-images utilisent des techniques de reconnaissance. Lors d'une phase d'apprentissage, la scène est d'abord parcourue par une caméra. Des points sont extraits et mis en correspondance dans ces images puis reconstruits en 3D. Chaque point du modèle est défini par ses coordonnées 3D mais aussi par une classe contenant l'ensemble des patchs autour des points images ayant contribué à la reconstruction de ce point 3D [2]. Lors d'une application, les points extraits dans l'image courante sont mis en correspondance avec le modèle en identifiant le patch image du modèle le plus semblable à celui présent dans l'image courante. Ceci permet d'obtenir des correspondances 2D/3D puis de calculer la pose de la caméra.

2 Sujet

Lorsque le point de vue courant est éloigné de ceux adoptés pendant l'apprentissage, la mise en correspondance est plus difficile car l'apparence d'un point peut être très différente de celles présentes dans le modèle. Pour remédier à ce problème, une possibilité est d'utiliser les techniques de simulation de points de vue telles que ASIFT [1] dont le but est de calculer des transformations affines des images afin de simuler des images de la scène avec des points de vue différents de ceux présents dans l'image. Cette méthode a montré son efficacité mais au prix d'un fort coût de calcul.

L'objectif de cette thèse d'utiliser les techniques de simulation pour améliorer le calcul de pose lorsque l'image courante est éloignée de la séquence d'apprentissage, ce qui est particulièrement le cas lors de l'initialisation d'une application de réalité augmentée, aucune localisation initiale grossière n'étant a priori disponible. Etant donné un modèle de scène reconstruit, l'objectif est d'ajouter des vues simulées pertinentes qui permettent d'élargir les possibilités de mise en correspondance sans trop augmenter le nombre de patchs dans le modèle et donc la complexité de la phase de reconnaissance. Les deux points d'étude principaux de ce travail seront:

- **Ajout de vues simulées et proposition d'une représentation compacte des classes:** l'objectif est de représenter un point 3D par un ensemble de patchs le plus petit possible pour des raisons d'efficacité, mais qui permette un processus de mise en correspondance fiable. Le processus d'ajout de vues simulées s'appuiera dans un premier temps sur la donnée des normales au modèle produites par la reconstruction. Idéalement, la représentation doit éliminer des vues réelles redondantes et simuler de nouveaux patchs de façon à fournir des vues locales du modèle avec une couverture relativement équiprobable des directions de vue. Outre cet aspect spatial, la représentation doit aussi tenir compte des descripteurs photométriques et paver davantage l'espace des vues là où les descripteurs sont différents.

- **Elaboration d'une stratégie de mise en correspondance entre image et modèle:** l'utilisation de la simulation peut introduire dans la classe des patches incorrects en nombre éventuellement élevé et augmenter ainsi la combinatoire et le nombre de correspondances 2D/3D aberrantes. Il est donc fondamental de disposer de stratégies permettant de limiter ces hypothèses incorrectes, même si par ailleurs des estimations robustes sont utilisées (RANSAC) et permettent au calcul de pose de rester insensible à la présence d'une certaine fraction d'appariements aberrants. Plusieurs pistes de recherche sont possibles. Parmi celles ci, on pourra prendre en compte des contraintes de visibilité permettant d'assurer la cohérence des appariements retenus, dans l'esprit de ce qui est fait en stéréovision [4]. Une autre possibilité est de diminuer la probabilité d'apparitions d'incohérences en utilisant des contraintes spatiales dans la génération de vues simulées, une idée étant par exemple d'appliquer les simulations à des groupes de points voisins (la notion adéquate de groupe restant à définir) et non plus des points indépendants.

Un prototype implantant ces stratégies sera développé et comparé aux autres méthodes classiques de relocalisation [3].

3 Autres informations

L'équipe Magrit travaille depuis plusieurs années sur le problème du positionnement par des méthodes de reconnaissance à partir d'une seule carte ou d'un modèle structuré. Plus de renseignements sont disponibles sur le site web de l'équipe <http://magrit.loria.fr>

Contact: Les candidatures sont à adresser à Marie-Odile.berger@loria.fr et frederic.sur@loria.fr. Des connaissances en vision par ordinateur et des notions de mathématiques appliquées sont souhaitables.

References

- [1] J. Morel and Y. Guoshen SIFT: A New Framework for Fully Affine Invariant Image Comparison In *SIAM J. Img. Sci.*, 2009.
- [2] S. Bhat and M.O. Berger and F. Sur Visual words for 3D reconstruction and pose computation In *The First Joint 3DIM/3DPVT Conference, Hangzhou, China*, 2011.
- [3] G. Klein and D. Murray Improving the agility of keyframe-based SLAM. In *ECCV 08*, 2008.
- [4] S.N. Sinha, D. Steedly, and R. Szeliski. Piecewise planar stereo for image-based rendering. In *Proc. Int. Conf. Computer Vision (ICCV'09)*, pages 1881–1888, 2009.