

Sujet de thèse équipe Magrit

Reconnaissance de lieux et d'objets pour la Réalité Augmentée en milieu complexe

Encadrants : Gilles Simon et Marie-Odile Berger

Mots-clés : réalité augmentée, vision par ordinateur, reconnaissance, apprentissage.

1 Sujet

Cette recherche a pour objectif de permettre l'utilisation de la Réalité Augmentée (RA) dans des environnements industriels de grande taille (usines, centrales, navires, ...). Les applications visées sont, entre autres, l'aide à la fabrication, l'aide à la maintenance, la documentation et la formation. Les systèmes de localisation basés sur l'émission / réception d'ondes physiques (magnétiques, acoustiques, optiques, wifi...) ou sur la détection de tags dans les images ne sont pas adaptés aux milieux considérés car ils limitent les interactions avec l'environnement et ont une faible portée. La solution développée devra donc utiliser l'analyse d'image uniquement, et exploiter les caractéristiques visuelles des objets présents dans l'environnement.

Nous considérons des environnements de grande taille, éventuellement composés de plusieurs lieux. Nous chercherons dans un premier temps à identifier le lieu dans lequel se trouve l'opérateur, puis dans un second temps à reconnaître les objets présents dans ce lieu qui feront l'objet d'annotation ou d'augmentation. Les données disponibles sur la scène sont constituées d'un ensemble d'images prises dans les différents lieux et d'un ensemble d'images des objets à reconnaître dans cette scène.

1.1 Reconnaissance de lieux

Les méthodes de reconnaissance usuelles risquent d'être peu performantes dans notre contexte en raison notamment de la nature des objets considérés (peu texturés), de la présence de nombreuses surfaces spéculaires et de l'apparition d'objets identiques dans un même lieu ou dans des lieux différents (valves, tuyaux, ...). De plus, l'utilisateur pourra se trouver dans un lieu non représenté dans la base, le système devant alors détecter que le lieu est inconnu. Enfin, l'utilisateur sera libre de ses déplacements et les lieux pourront être observés depuis des points de vue éloignés de ceux choisis au cours de la phase d'acquisition. Les pistes de recherche envisagées pour contrer ces difficultés sont les suivantes : 1. Générer un vocabulaire visuel utilisant des indices denses (dense SIFT, HOG, ...) ; 2. Proposer une méthode d'indexation permettant de tenir compte des objets répétés [3]. 3. À partir de plusieurs photos, reconstruire localement des points clés en 3D afin de générer des indices correspondant à des points de vue éloignés de ceux représentés dans la base [5].

1.2 Reconnaissance d'objets

Différents descripteurs ont été utilisés en classification d'images : mots visuels, histogrammes de gradients, shape contexts, caractéristiques pseudo Haar, etc. Les objets rencontrés en environ-

nement industriel étant de natures très diverses, notamment en terme de forme, de texture et de matériau, il est probable qu'un seul de ces critères ne convienne pas à l'ensemble des objets considérés. Nous envisageons donc d'intégrer à la phase d'apprentissage la détermination du ou des critères les plus pertinents à utiliser pour chaque objet d'intérêt. De plus, l'organisation spatiale des objets proches sera exploitée afin de faciliter leur reconnaissance. Cette organisation étant dépendante du point de vue d'observation, nous souhaitons utiliser un modèle flexible dans l'esprit de [1, 2, 6]. L'objectif est que les arrangements entre objets soient appris statistiquement à partir de photos ou de vidéos et intégrés au framework de reconnaissance [2]. L'arrière-plan des objets (le *scene context*) pourra aussi être pris en compte. Dans [6], une description holistique du contexte est utilisée, qui repose sur une représentation spectrale de l'image [4]. Toutefois, d'autres descripteurs, tels que ceux basés sur des bag-of-words, pourraient aussi être utilisés. Nous considérerons ces différentes options et élaborerons une représentation susceptible de discriminer au mieux les différents contextes présents dans les environnements industriels.

2 Compétences requises

Des compétences sont attendues dans l'un ou plusieurs des domaines suivants : vision par ordinateur, reconnaissance, apprentissage.

3 Encadrement et contact

Gilles Simon (gilles.simon@loria.fr) et Marie-Odile Berger (marie-odile.berger@loria.fr). Des articles et démos des activités de l'équipe Magrit sont disponibles sur notre site web <http://magrit.loria.fr>.

References

- [1] David Crandall, Pedro Felzenszwalb and Daniel Huttenlocher. Object Recognition by Combining Appearance and Geometry. In *Conference on Computer Vision and Pattern Recognition*, CVPR, 2005.
- [2] C. Desai, D. Ramanan and C. Fowlkes. Discriminative models for multi-class object layout. In *IEEE 12th International Conference on Computer Vision*, Kyoto, 2009, pp. 229-236.
- [3] Nicolas Noury. Mise en correspondance a contrario de points d'intérêt sous contraintes géométrique et photométrique. Thèse. Université Henri Poincaré - Nancy I, 2011.
- [4] Aude Oliva and Antonio B. Torralba. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision - IJCV*, vol. 42, no. 3, pp. 145-175, 2001.
- [5] Pierre Rolin, Marie-Odile Berger and Frédéric Sur. Viewpoint simulation for camera pose estimation from an unstructured scene model. *International Conference on Robotics and Automation*, May 2015, Seattle, United States.
- [6] Yukun Zhu, Jun Zhu and Rui Zhang. Contextual Object Detection. In *IEEE Transactions on Multimedia*, vol. 16, no. 6, pp. 1585-1596, Oct. 2014.