

Suivi d'objets dans une séquence d'images

La vue est notre sens le plus utilisé mais aussi le plus complexe à reproduire. Notre cerveau est capable d'analyser en temps réel l'information procurée par nos yeux et d'en tirer des informations très diverses allant de la perception de la géométrie de la scène nous entourant à la lecture d'un texte. Ce qui est capital pour mener à bien cette analyse est de pouvoir reconnaître un même objet en toutes circonstances. Cela permet de le suivre et donc d'améliorer la connaissance que l'on en a ou au contraire de l'identifier à un objet que l'on a déjà rencontré auparavant. Lorsque l'on souhaite permettre à une machine d'analyser le contenu d'images, il est alors nécessaire de reproduire ce comportement. Je me suis donc intéressé à cet aspect du traitement d'image et à ses applications possibles.

Recherches

Les besoins en robustesse de la reconnaissance d'objets ne permettent pas une approche naïve du problème comme c'est parfois le cas en algorithmie. Un premier travail a donc consisté en la détermination des propriétés clefs que doit posséder l'algorithme. Bien que d'autres méthodes existent, les algorithmes les plus efficaces dans le cas général sont basés sur la recherche et la comparaison de points caractéristiques des images. Le besoin d'invariance de l'algorithme aux conditions de visualisation de l'objet possède donc deux aspects : l'invariance de la sélection des points et celle de leur caractérisation afin que leur comparaison ait un sens.

Mise en pratique

Après avoir identifié les étapes communes aux différents algorithmes employés dans le domaine de la reconnaissance d'images, j'ai implémenté l'algorithme SIFT (Scale Invariant Feature Transform) de D.Lowe. J'ai choisi cet algorithme car il est reconnu pour son efficacité et est à la base de beaucoup d'autres. De plus cet algorithme est complet car il s'intéresse non seulement à l'extraction des points remarquables mais aussi à leur comparaison. Cela m'a permis dans un second temps d'appliquer le suivi d'objets à un exemple concret, un aspect important de la réalité augmentée : le calcul du mouvement d'une caméra à partir de l'image filmée. J'ai cependant simplifié le problème en ne conservant pour seuls degrés de liberté que la rotation autour de deux axes afin de pouvoir comparer les résultats du programme à ceux donnés par un gyroscope solidaire de la caméra.

Résultats

Dans le cas simple de la comparaison de deux images liées par une transformation simple (similitude plane ou une transformation affine de faible déformation), l'algorithme est capable de retrouver cette transformation de façon très précise. Cependant pour des images réelles, les modifications supplémentaires comme la déformation de l'optique, le bruit apporté en grande partie par la compression de la vidéo ou le flou cinétique rendent les résultats beaucoup moins précis que ceux du gyroscope. Toutefois, l'enregistrement vidéo permet sur le long terme de parer à la dérive du gyroscope en donnant une estimation de la valeur attendue.

Plan de l'exposé

- I. Recherche de primitives dans une image
- II. Comparaison des primitives de deux images
- III. Applications : Extractions d'information de la séquence d'images

Bibliographie

- Martin A. FISCHLER and Robert C. BOLLES, Random Sample Consensus : A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. 1981
- David G. LOWE, Distinctive Image Features from Scale-Invariant Keypoints. 2004
- Krystian MIKOLAJCZYK and Cordelia SCHMID, Scale & Affine Invariant Interest Point Detectors. International Journal of Computer Vision 60, p.63-86, 2004
- Diane LINGRAND, *Introduction au traitement d'images*, 2^eédition. Vuibert, 2004
- Krystian MIKOLAJCZYK et al., A Comparison of Affine Region Detectors. 2006