

Proposition de TFE : interprétation des mouvements humains par une analyse anatomique

Ce travail de fin d'études s'inscrit dans la thématique de l'analyse automatique des mouvements humains. Il fait intervenir des méthodes d'analyse d'images, et d'apprentissage automatique.

Personnes de contact : M. Van Droogenbroeck - S. Piérard.

Introduction

Pourquoi analyser les mouvements humains ?

Ces dernières années, les caméras vidéo sont devenues omniprésentes dans de nombreux domaines : la vidéo-surveillance, la domotique, les jeux (par exemple la *Kinect* de *Microsoft*), etc. Les applications sont nombreuses, mais, le plus souvent, elles nécessitent l'analyse et l'interprétation automatique de la scène filmée. Or, dans la majorité des scènes, il y a des personnes. Des techniques permettant d'interpréter automatiquement leurs mouvements se révèlent donc indispensables. L'interprétation automatique des mouvements humains n'est pas un problème qui puisse être considéré comme résolu dans l'état actuel de la science.

Quelles données utiliser ?

Malheureusement, les personnes peuvent prendre une multitude d'apparences : leurs couleurs et leurs textures sont imprévisibles. Afin d'obtenir des techniques robustes, il est préférable de se baser sur des informations géométriques plutôt que sur l'apparence. Pour y arriver, deux techniques principales peuvent être utilisées :



FIGURE 1 – Deux méthodes peuvent être utilisées pour extraire des informations purement géométriques. (a) Application d’un algorithme de soustraction d’arrière plan avec une caméra “classique”. (b) Application d’un algorithme de soustraction d’arrière plan avec une caméra de profondeur.

1. Se servir de l’information temporelle présente dans le flux vidéo pour détecter les personnes et les objets en mouvement. Les silhouettes des utilisateurs et des objets avec lesquels ils interagissent sont ainsi mises en évidence. (voir figure 1a)
2. Utiliser une caméra de profondeur (aussi appelées caméras “3D”, comme par exemple la *Kinect* de *Microsoft*). L’information fournie en chaque pixel n’est pas une couleur, mais une estimation de la distance séparant la caméra du point de la scène correspondant au pixel. On obtient ainsi des silhouettes annotées avec une information de profondeur. (voir figure 1b)



FIGURE 2 – La *Kinect* de *Microsoft* est une caméra de profondeur.

Comment faire ?

Plusieurs approches différentes peuvent être envisagées. Cependant, dans de nombreux cas, il semble utile, dans une étape de pré-processing, de repérer précisément l’emplacement des différentes parties du corps (tête, mains, bras, jambes, etc) dans l’image. La figure 3 montre deux exemples de silhouettes synthétiques annotées avec une information anatomique (de telles données peuvent servir à constituer un *learning set*). La figure 4 illustre quelques exemples d’annotations anatomiques automatisées (c’est le résultat souhaité).

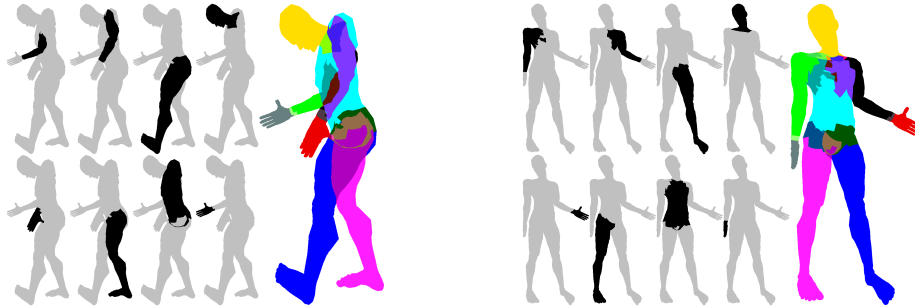


FIGURE 3 – Deux exemples de silhouettes annotées avec une information anatomique.

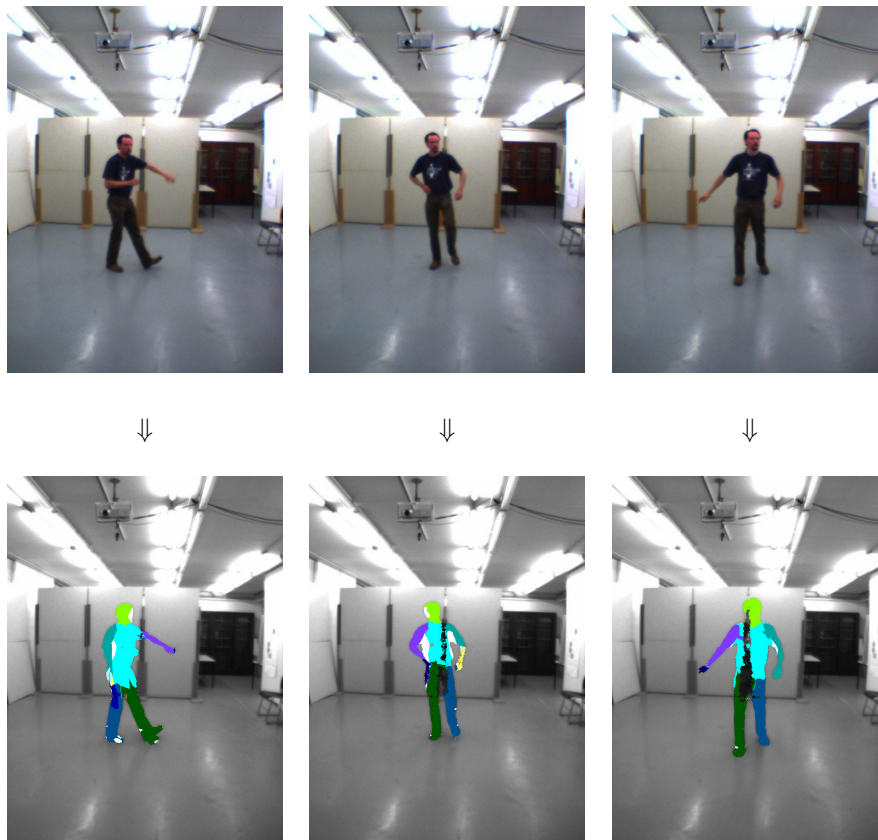


FIGURE 4 – Quelques exemples d'annotations anatomiques automatisées.

Description du travail

Il est demandé, dans un premier temps, de réaliser une étude bibliographique concernant les techniques permettant d’annoter anatomiquement les images avec une information anatomique, et d’évaluer les débouchés potentiels d’une telle technique.

Puis, il est demandé de concevoir et de réaliser un système original permettant d’annoter les différentes parties anatomiques du corps en temps-réel. Cette analyse se fera uniquement sur base d’informations géométriques, et par des méthodes d’apprentissage automatique. Il sera également demandé d’évaluer quantitativement la méthode implémentée, et de comparer les résultats obtenus à partir des silhouettes avec et sans information de profondeur.

Enfin, l’étudiant choisira une application concrète (identification des personnes par leur démarche [1], reconnaissance de la pose [2], ou une application de son choix) et utilisera l’annotation anatomique comme point de départ.

Le logiciel sera réalisé en C ou C++ sous Linux (version 32 bits). Pour réaliser ce travail, l’étudiant disposera des outils suivants :

1. une caméra de profondeur “kinect” (uniquement au labo) ;
2. un avatar humain réaliste déformable (MakeHuman), dont la pose peut être commandée par 105 degrés de libertés (afin de visualiser une pose si l’application choisie est la pose recovery) ;
3. un algorithme de soustraction d’arrière plan que l’on peut utiliser pour mettre en évidence la silhouette de la personne observée ;
4. un logiciel permettant de générer automatiquement des silhouettes humaines réalistes annotées avec une information de profondeur ou avec des informations anatomiques (les données ainsi produites sont destinées à constituer un *learning set*).

Références

- [1] N. Boulgouris and Z. Chi. Human gait recognition based on matching of body components. *Pattern Recognition*, 40(6) :1763–1770, June 2007.
- [2] M. Lee and R. Nevatia. Body part detection for human pose estimation and tracking. In *IEEE Workshop on Motion and Video Computing (WMVC)*, Austin, USA, February 2007.