

# **Proposition de mémoire de maîtrise**

Stéphane Drouin  
19 décembre 2001

Sujet: description 3D d'une personne en mouvement à partir d'images 2D.

## ***Le problème***

Afin de comprendre le comportement d'êtres humains et leurs interactions avec l'environnement, une description adéquate des personnes observées doit être disponible. Le problème étudié dans ce mémoire est la description 3D d'êtres humains en mouvement à partir d'information extraite d'images 2D.

## ***Motivation***

Les systèmes intelligents sont appelés à devenir omniprésents dans notre environnement. Leurs interactions avec les humains demanderont de leur part une capacité d'acquérir des connaissances nouvelles par un processus d'apprentissage. C'est dans ce contexte qu'a été mis sur pied le projet COGNOIS (Communication and Observation toward a Generic Natural Ontogeny for Intelligent Systems) au Laboratoire de vision et systèmes numériques. Un des objectifs de ce projet est l'observation des êtres humains et la compréhension de leur comportement.

Plusieurs couches de traitement sont nécessaires afin d'arriver à ce résultat.

1. Des images de la scène doivent être acquises par des caméras calibrées et synchronisées.
2. L'information de chaque image doit être extraite et combinée à celle obtenue des autres images.
3. L'information tridimensionnelle doit être décrite et analysée afin d'y reconnaître des humains.
4. L'évolution temporelle de chaque humain doit être analysée pour reconnaître son comportement, ses mouvements.

La première étape ne dépend pas de l'application. Elle permet de transformer le problème à quatre dimensions d'entités dynamiques en un ensemble de problèmes à deux dimensions. Les trois autres couches correspondent aux aspects cognitifs dont on veut doter le système: analyser l'information disponible afin de reconnaître des modèles, faire le lien entre l'observation et les connaissances. Ces connaissances peuvent être données au système ou encore être construites par apprentissage.

## ***Travaux reliés***

Les questions de synchronisation ont été étudiées dans le projet 3D room [Kanade]. Dans ce cas, la synchronisation des acquisitions est faite électroniquement, en ajoutant l'information temporelle aux images.

Le calibrage d'un système de vision se fait généralement en deux étapes; les paramètres intrinsèques sont calibrés, puis les paramètres extrinsèques sont obtenus séparément. Le calibrage des paramètres intrinsèques d'une caméra peut se faire avec une cible planaire. [Zhang] propose une méthode flexible utilisant une cible planaire connue. [Heikkilä] tient compte des erreurs inhérentes à l'utilisation de cercles comme points de calibrage et ajoute un modèle de distorsion inverse. La méthode de calibrage extrinsèque d'un système de vision à grand volume proposée par [Chen] utilise une seule LED déplacée dans tout le volume.

Une revue de récente des travaux concernant l'observation d'êtres humains a été faite par [Moeslund]. [Wren] utilise des modèles cinématiques et statistiques afin d'obtenir une pose vraisemblable à partir d'information limitée à quelques points d'intérêt. [Bottino] ajuste un modèle détaillé d'être humain sur un ensemble de voxels obtenus par intersection de volumes. [Bregler] retrouve les déplacements d'un modèle 3D directement à partir des déplacements 2D d'amas de pixels. [Delamarre] suggère une méthode d'ajustement du modèle utilisant la notion de force physique et qui s'applique tant à des systèmes à petit et à grand baseline. Pour ce qui est de la reconnaissance des mouvements, [Aggarwal] classe les approches basées ou non sur un modèle précis d'être humain selon qu'elles utilisent la concordance matricielle (template matching) ou l'espace des états (state space) pour reconnaître le mouvement.

## ***Objectifs de recherche***

Objectifs expérimentaux:

1. Développer une architecture expérimentale pour le projet Cognitois.
2. Extraire et valider un ensemble de paramètres de mesure du mouvement d'un humain dans une séquence d'images obtenues à partir d'un réseau de caméras.

Objectifs scientifiques:

Nous faisons l'hypothèse que l'information 3D obtenue de plus d'une caméra, sera avantageusement exploitée dès le début du processus de mesure dans les images. Cette hypothèse est en opposition avec l'approche où l'information est complètement extraite en 2D puis fusionnée à la toute fin. Notre **objectif principal** sera donc de valider cette hypothèse en la confrontant à l'approche 2D. Il serait éventuellement intéressant de vérifier comment l'être humain exploite ces informations par un survol de la littérature.

Un objectif plus spécifique sera de stabiliser (filtrer) la description du modèle humain par une observation répétée dans le temps de certains invariants comme la longueur des membres. On devrait alors observer l'animation d'un modèle stable malgré le mouvement.

## ***Méthodologie***

Ce projet de recherche est divisé en deux parties. L'acquisition d'images calibrées et synchronisées fait l'objet de la première partie. La description 3D d'un être humain est le deuxième objectif.

1. Acquisition d'images calibrées et synchronisées.
  - a. Implanter une méthode de synchronisation efficace et robuste.
  - b. Proposer une méthode de calibrage flexible pour accommoder un large éventail de configuration des caméras. Les méthodes utilisant une cible planaire librement déplacée dans l'espace proposées par [Heikkilä] et [Zhang] sont étudiées à cette fin.
2. Description 3D d'un être humain à partir d'images 2D.
  - a. Segmenter les images 2D indépendamment pour y identifier un être humain. On fait des hypothèses simplificatrices, mais sans supposer que la segmentation est parfaite.
  - b. Extraire des primitives 3D (voxels) en combinant l'information 2D. Une approche par intersection de volumes est étudiée afin d'arriver à ce résultat [Kutulakos], [Laurentini].
  - c. Ajuster un modèle adéquat sur les primitives 3D [Moeslund], [Bottino], [Delamarre].

## ***Échéancier***

L'échéancier suivant est prévu pour mener le projet à terme :

- Mai à septembre 2001: mise en marche du système, développement des outils d'acquisitions, développement des méthodes de calibrage et de synchronisation.
- Septembre à novembre 2001: expérience sur le calibrage et la synchronisation, mesure de l'erreur pour la préparation d'un article de conférence.
- Novembre 2001 à mars 2002: reconstruction d'entités 3D par l'intersection de volume: segmentation en 2D, intersection de volumes, choix de primitives 3D.
- Mai à juillet 2002: expérience sur la reconstruction d'objets 3D, mesure de l'erreur qualitativement et/ou quantitativement pour la préparation d'un article.
- Juillet et août 2002: rédaction du mémoire.

## **Références**

[Aggarwal] J.K. Aggarwal and Q. Cai. Human motion analysis: a review. *CVIU*, 73(3):428-440, March 1999.

[Bottino] A. Bottino and A. Laurentini. *A silhouette based technique for the reconstruction of human movement*. *CVIU*, 83(1):79-95, July 2001.

[Bregler] C. Bregler and J. Malik. *Video motion capture*. UCB//CDS-97-973.

[Chen] X. Chen, J. Davis and P. Slusallek. *Wide area camera calibration using virtual calibration objects*. In *CVPR00*, pages II:520-527, 2000.

[Delamarre] Q. Delamarre and O. Faugeras. 3D Articulated models and multiview tracking with physical forces. *CVIU*, 81(3): 328-357, March 2001.

[Heikkilä] J. Heikkilä. *Geometric camera calibration using circular control points*. *PAMI*, 22(10):1066-1077, October 2000.

[Kanade] T. Kanade, H. Saito and S. Vedula. *The 3D room: digitizing time-varying 3D events by synchronized multiple video streams*. CMU-RI-TR-98-34, December 1998.

[Kutulakos] K.N. Kutulakos and S.M. Seitz. *A theory of shape by space carving*. University of Rochester CS Technical Report #692, May 1998.

[Laurentini] A. Laurentini. *The visual hull concept for silhouette-based image understanding*. *PAMI*, 16(2):150-162, February 1994.

[Moeslund] T.B. Moeslund and E. Granum. *A survey of computer vision-based human motion capture*. *CVIU*, 81(3):231-268, March 2001.

[Wren] C.R. Wren and A.P. Pentland. *Dynamic model of human motion*. AFGR, Japan, 1998.

[Zhang] Z. Zhang. *A flexible new technique for camera calibration*. *PAMI*, 22(11):1330-1334, November 2000.