



Apprentissage de modèle d'attention pour le suivi de conversations multipartites par un robot humanoïde

Contexte

L'objectif général de ces recherches est de développer un robot humanoïde capable de copier le comportement humain : le robot doit montrer qu'il est attentif à la scène dans laquelle il est plongé par un comportement verbal et gestuel approprié, imitant le comportement d'un tuteur humain placé dans des situations similaires.

Dans le cadre de ce travail, on s'intéresse au suivi de conversations. On cherche à ce que le robot se comporte de manière adéquate par des mouvements de tête et des yeux, afin qu'il semble suivre une conversation entre interlocuteurs placés à divers endroits en face de lui.

Sujet

Le sujet de stage comporte trois tâches essentielles :

1. *Recueil de données de comportement humain.* On place le tuteur humain, équipé de capteurs, en face d'un écran où sont placés une dizaine de visages (cf. ci-contre), dont un seul parle à la fois (ici le deuxième en partant du bas). Ces visages se relaient pour lire à haute voix un texte, chacun prononçant successivement chaque phrase. Les prises de parole sont déterminées de manière aléatoire parmi les visages. Comme il s'agit toujours du même locuteur et de la même voix, l'attention du tuteur n'est influencée que par les mouvements de bouche en cohérence avec la parole et la spatialisation sonore. On demande au tuteur humain de porter attention au visage qui parle, et on enregistre les mouvements de sa tête (avec un système de capture de mouvement) et de son regard (avec un oculomètre portable).
2. *Modélisation.* On cherche à développer un modèle capable de prédire les mouvements de la tête et du regard, tels qu'ils ont été enregistrés à partir de la séquence de prise de parole des visages. Une première piste est de construire des automates probabilistes type chaînes de Markov cachées (Hidden Markov Models) décrivant la succession des divers états de la scrutation et les comportements associés. Les Trajectory-HMM (Zen, Nose et al. 2007) ont la capacité de générer des comportements optimaux à partir d'une suite d'états. Le progiciel HTS développé par Zen et al est bien maîtrisé au laboratoire. D'autres approches seront examinées.
3. *Evaluation.* On évaluera la capacité du modèle à imiter le mouvement de la tête et des yeux du tuteur humain. On cherchera à reproduire notamment les propriétés de coordination tête-regard dans les aspects temporels et spatiaux. Une analyse critique des performances sur cette tâche simple permettra de conditionner les développements futurs.

L'objectif final est d'implémenter un modèle d'attention audiovisuelle sur un robot humanoïde (ICUB2 ci-contre) que le laboratoire accueillera à la fin du printemps 2013.



Thématiques abordées dans le stage

- Capture de mouvement, oculométrie
- Modélisation statistique

Compétences requises

- Notions de statistique, maîtrise de Matlab

Contacts

Gérard Bailly	GIPSA-lab	04 76 57 47 11
Denis Pellerin	GIPSA-lab	04 76 57 43 69
Michèle Rombaut	GIPSA-lab	04 76 57 43 68

Gerard.Bailly@gipsa-lab.grenoble-inp.fr
Denis.Pellerin@gipsa-lab.grenoble-inp.fr
Michele.Rombaut@gipsa-lab.grenoble-inp.fr

Indemnités de stage

Ce stage fait l'objet d'une indemnité fixée annuellement par le conseil de laboratoire, de l'ordre de 400€ mensuels.

Référence

Zen, H., T. Nose, et al. (2007). The HMM-based speech synthesis system version 2.0. Speech Synthesis Workshop, Bonn, Germany, pp. 294-299.