

Sujet de thèse

Synthèse de parole en temps-réel pour la production d'une personnalité vocale

Contenu

Pour le cinéma la post-production permet de proposer des corrections au jeu d'un acteur. Pour le jeu vidéo, on cherche à créer des personnages au réalisme toujours plus poussé et personnalisé.

Ce projet se propose de développer des méthodes de synthèse de la parole à partir du texte temps réel afin d'aborder la partie vocale de cette problématique.

Du point de vue de la recherche, il s'agit de construire un synthétiseur vocal à partir du texte ou d'une description phonétique riche (Français/Anglais) temps réel, expressif et personnalisé.

Pour le jeu vidéo, cela permettra de produire des variations expressives en fonction du contexte de jeu et de l'état affectif du joueur. Pour le cinéma, de contribuer au doublage ou à la post-production.

La synthèse vocale est donc envisagée ici comme modèle et outil de « doublure numérique » réaliste d'un personnage (acteur, personnage de jeu vidéo, personnage virtuel associé à un joueur).

L'évaluation de la synthèse et de son réalisme, en regard de la personnalité visée, permettra de valider la « doublure numérique ».

Cette thèse, inscrite dans le cadre d'un projet du FUI, sera menée en collaboration étroite avec des partenaires industriels, dont la mission est d'intégrer des outils de post-production.

Compétences requises

Ce travail de recherche demande de solides connaissances en traitement du signal acoustique, ainsi que des connaissances en traitement automatique de la langue parlée. Des compétences en phonétique ou linguistique seraient appréciées, ainsi qu'un intérêt marqué pour la parole et l'expression sonore.

Une bonne capacité au travail en collaboration sur un projet lié à des échéances strictes est nécessaire. Une connaissance fonctionnelle du français et de l'anglais est importante pour la réalisation de cette thèse.

Des profils de type traitement du signal ou informatique linguistique seront considérés en priorité.

Contexte et équipe d'accueil

Cette thèse s'inscrit dans le contrat FUI ADN T-R. Elle se déroulera au LIMSI-CNRS (www.limsi.fr) dans le groupe Audio & Acoustique et en collaboration étroite avec les partenaires industriels du projet.

Le LIMSI-CNRS est un laboratoire propre du CNRS, situé sur le campus de l'université Paris Sud XI à Orsay. Le LIMSI est un laboratoire de réputation internationale en matière de recherches sur la parole. Le Groupe Audio & Acoustique s'intéresse en particulier à l'analyse et à la synthèse de parole, au traitement audio temps-réel, au son expressif.

Cette thèse commencera à partir d'octobre 2011, financée par le FUI pour une durée 3 ans. La rémunération est calculée sur la base du contrat doctoral CNRS (entre 1750 et 2100 € brut mensuel).

Encadrement – contact :

La thèse sera encadrée par Christophe d'Alessandro, directeur de recherche au CNRS.

Les candidatures seront adressées à :

Christophe d'Alessandro <cda@limsi.fr>