

JOURNEE SCIENTIFIQUE LIMA (Loisirs et IMAgés)

**Projet du Cluster n° 2 ISLE de la région Rhône-Alpes
8 juillet 2008**

**Grenoble INP, 46 avenue Félix Viallet
Salle R**

(http://www.grenoble-inp.fr/12657183/0/fiche___pagelibre/)

Programme de la journée

Cette journée comporte deux exposés "senior", le matin, quatre exposés "doctorant", quelques démonstrations et une discussion sur les actions à mener pour la suite du projet.

9h30 - 10h : Accueil, introduction

10h - 10h50 : Emmanuel Dellandréa, LIRIS / IMAGINE - Lyon
Classification de l'émotion dans les signaux de parole et musique

10h50 - 11h40 : Christophe Ducottet, LHC- Hubert Konic, LIGIV - Saint Etienne
Mise sous contrainte de segmentations pour le suivi d'objets dans les séquences d'images

11h40 - 12h : Claude Cadoz (Directeur de l'ACROE)
Présentation du logiciel GENESIS pour la synthèse de son et la synthèse musicale par modèles physiques

12h - 13h30 : Buffet

13h30 - 14h : Annie Luciani
Démonstrations du labo ICA (systèmes temps réel et retour d'effort)

14h - 14h40 : Matthieu Evrard - ICA - Grenoble
MIMESIS : un environnement de conception interactif pour la création de mouvements 3D par modélisation et simulation de réseaux masses-interactions

14h40 - 15h10 : Hichem BARKI, LIRIS / M2DisCo - Lyon
FIMS : algorithme de calcul exact de la somme de Minkowski de maillages 3D

15h10 - 15h40 : Sophie Marat et Tien Ho, GIPSA - Grenoble
Un modèle de saillance spatio-temporelle pour la prédiction des mouvements oculaires lors du visionnage d'une vidéo

15h40 - 16h10 : Ludovic Dutrève, LIRIS / SAARA - Lyon (introduction Alexandre Meyer)
Transfert d'Animation Faciale en Temps Réel à Base de Points Caractéristiques

16h10 - 17h : Discussion sur les actions à mener pour la poursuite du projet

Résumés des présentations

10h - 10h50 : Emmanuel Dellandréa, LIRIS / IMAGINE - Lyon

Titre : Classification de l'émotion dans les signaux de parole et musique

Résumé : Nous présentons lors de cet exposé nos travaux de recherche portant sur la reconnaissance automatique de l'émotion et de l'humeur au sein de signaux audio. L'émotion portée par les signaux audio constitue en effet une information sémantique particulièrement importante dont l'analyse automatique offre de nombreuses possibilités en termes d'applications, telles que les interactions homme-machine intelligentes et l'indexation multimédia. Afin d'analyser l'émotion audio tant pour la parole que pour les signaux musicaux, nous utilisons dans notre travail un modèle émotionnel discret combiné à un modèle dimensionnel, en nous appuyant sur des études existantes sur les corrélations entre les propriétés acoustiques et l'émotion dans la parole ainsi que l'humeur dans les signaux de musique. Les principales contributions de nos travaux sont les suivantes. Tout d'abord, nous avons proposé, en complément des caractéristiques audio basées sur les propriétés fréquentielles et d'énergie, de nouvelles caractéristiques harmoniques et Zipf, afin d'améliorer la caractérisation des propriétés des signaux de parole en terme de timbre et de prosodie. Deuxièmement, dans la mesure où très peu de ressources pour l'étude de l'émotion dans la parole et dans la musique sont disponibles par rapport au nombre important de caractéristiques audio qu'il est envisageable d'extraire, une méthode de sélection de caractéristiques, ESFS, basée sur la théorie de l'évidence est proposée afin de simplifier le modèle de classification et d'en améliorer les performances. De plus, nous avons montré que l'utilisation d'un classifieur hiérarchique basé sur un modèle dimensionnel de l'émotion, permet d'obtenir de meilleurs résultats de classification qu'un unique classifieur global, souvent utilisé dans la littérature. Par ailleurs, puisqu'il n'existe pas d'accord universel sur la définition des émotions de base, et parce que les états émotionnels considérés sont très dépendant des applications, nous avons également proposé un algorithme basé sur ESFS, permettant de construire automatiquement un classifieur hiérarchique adapté à un ensemble spécifique d'états émotionnels dans le cadre d'une application particulière. Cette classification hiérarchique procède en divisant un problème de classification complexe en un ensemble de problèmes plus petits et plus simples grâce à la combinaison d'un ensemble de sous-classifieurs binaires organisés sous forme d'un arbre binaire. Enfin, les émotions étant par nature des notions subjectives, nous avons également proposé un classifieur ambigu, basé sur la théorie de l'évidence, permettant l'association d'un signal audio à de multiples émotions, comme le font souvent les êtres humains.

L'efficacité de ces techniques de reconnaissance a été évaluée sur les ensembles de données Berlin et DES pour la reconnaissance de l'émotion dans la parole, et sur un ensemble de données construit au sein de notre laboratoire pour l'humeur dans les signaux de musique, dans la mesure où il n'existe pour le moment aucun jeu de données public.

10h50 - 11h40 : Christophe Ducottet, LHC- Hubert Konic, LIGIV - Saint Etienne

Titre : Mise sous contrainte de segmentations pour le suivi d'objets dans les séquences d'images.

Résumé :

This talk presents a clustering-based color segmentation method where the desired object is focused on. As classical methods suffer from a lack of robustness, salient colors appearing in the object are used to intuitively tune the algorithm. These salient colors are extracted according to a psychovisual scheme and a peak-finding step. Results on various test sequences, covering a representative set of outdoor real videos, show the improvement when compared to a simple implementation of the same K-means oriented segmentation algorithm with ad hoc parameter setting strategy and with the well-known mean-shift algorithm.

14h - 14h40 : Matthieu Evrard - ICA - Grenoble

Titre : MIMESIS : un environnement de conception interactif pour la création de mouvements 3D par modélisation et simulation de réseaux masses-interactions

Résumé : Le sujet de cette présentation porte sur la conception d'un outil informatique de création d'animations par modèles de réseaux masse-interaction. L'animation par ordinateur a fait l'objet de développements intensifs durant ces dix dernières années. De nombreux modèles de mouvements existent : depuis la modélisation de mouvements de phénomènes naturels jusqu'à la représentation de mouvements humains. De nombreux logiciels sont disponibles sur le marché ou sur le net. Parmi les modèles de synthèse de mouvement, la modélisation physique et

en particulier la modélisation par réseaux masse-interaction ont donné lieu à de nombreux modèles de phénomènes ou d'effets au cas par cas, mettant souvent l'accent sur des problématiques de contrôle à offrir à un utilisateur non expert de ce type de modélisation et exploitant peu sa modularité conceptuelle inhérente, apte à permettre à un utilisateur formé de développer ses propres modèles de mouvements. Des algorithmes de modèles masse-interaction sont aujourd'hui couramment implantés comme fonctions de base des grands logiciels d'animation. Cependant, les fonctions proposées sont en général assez particulières et s'intègrent mal dans les concepts fondateurs de ces logiciels, basés essentiellement sur l'application cinématique de mouvements sur des formes 3D. Ils donnent lieu également à des logiciels jouets disponibles sur le net, offrant la modularité à l'utilisateur mais très limités dans les fonctions proposées. La présentation portera sur la conception d'un outil de création informatique complet basé sur le concept du modèle masse-interaction. La généricité ainsi que la modularité d'un tel système de représentation de mouvements étant acquise, la conception et le développement de cet outil devront affronter les autres problèmes théoriques et pratiques inhérents à ce formalisme de manière à intégrer toutes les fonctionnalités nécessaires pour la création interactive de modèles de mouvement et de modèles de formes associées à ces mouvements. La présentation se clôturera sur les expériences pédagogiques ayant eu lieu au cours de ce travail.

14h40 - 15h10 : Hichem BARKI, LIRIS / M2DisCo - Lyon

Titre : FIMS : algorithme de calcul exact de la somme de Minkowski de maillages 3D.

Résumé : Dans cet exposé, nous parlerons des travaux que nous effectuons en analyse de maillage par morphologie mathématique. Nous commencerons par un état de l'art des méthodes existantes et leurs limitations ainsi que des applications potentielles de la morphologie mathématique dans le domaine des maillages. Par la suite, nous exposerons notre algorithme FIMS qui permet de calculer la somme de Minkowski de polyèdres convexes d'une manière exacte et efficace ainsi que les travaux en cours afin de l'adapter aux maillages non convexes. Nous présenterons aussi une implantation exacte et simplifiée des diagrammes de pente pour le calcul de la somme de Minkowski de polyèdres non convexes. Enfin, nous discuterons et comparons FIMS avec d'autres algorithmes et nous parlerons des directions futures.

15h10 - 15h40 : Sophie Marat et Tien Ho, GIPSA - Grenoble

(S. MARAT, T. HO PHUOC, L. GRANJON, N. GUYADER, D. PELLERIN, A. GUERIN-DUGUE)

Titre : Un modèle de saillance spatio-temporelle pour la prédiction des mouvements oculaires lors du visionnage d'une vidéo

Résumé : This talk presents a spatio-temporal saliency model that predicts eye movements. This biologically inspired model separated a video frame into two signals corresponding to the two main outputs of the retina (parvocellular and magnocellular outputs). Both signals are then decomposed into elementary feature maps by cortical-like filters. These feature maps are then used to form two saliency maps: a static one and a dynamic one. These maps are fused into a spatio-temporal saliency map. The model is evaluated by comparing the salient areas of each frame predicted by these saliency maps (static, dynamic, spatio-temporal) to the eye positions of different subjects during a video free viewing experiment with a large database (17000 frames)

15h40- 16h10 : Ludovic Dutrève, LIRIS / SAARA - Lyon (introduction Alexandre Meyer)

Titre : Transfert d'Animation Faciale en Temps Réel à Base de Points Caractéristiques

Résumé : Nous présentons une méthode de transfert d'animation faciale en temps réel. Notre système offre la possibilité d'utiliser différents types d'animations sources : des animations existantes de visages 3D par exemple ou bien des animations 2D issues du suivi de points caractéristiques dans une vidéo ou d'un système de capture de mouvements. A partir de deux ensembles de points caractéristiques sélectionnés manuellement sur les visages source et cible, un réseau de fonctions à base radiale (Radial Basis Functions, RBF) est initialisé et fournit une transformation géométrique entre les deux visages. A chaque frame de l'animation, nous appliquons la transformation obtenue sur les nouvelles positions des points de contrôle du visage source, fournissant les positions des points de contrôle du visage cible adaptés à l'expression du visage source et à la morphologie du visage cible. Le déplacement de ces points caractéristiques permet de déformer le maillage directement sur le processeur graphique par une méthode appelée Linear Blend Skinning largement utilisée dans l'industrie des jeux-vidéos. Dans l'optique de proposer un outil d'animation accessible au plus grand nombre, nous proposons une technique procédurale pour générer automatiquement les poids associant chaque sommet aux points caractéristiques. Pour résumé, notre méthode offre la possibilité de transférer en temps réel une animation d'un visage à un autre, elle peut s'adapter facilement à différents types d'animations sources et elle ne nécessite qu'une intervention humaine minimale pendant la phase de configuration du visage cible.