

Titre : Analyse et caractérisation automatique de dialogues dans les contenus audiovisuels

Nature du projet : Réalisation d'un prototype (POC) d'indexation et de visualisation de contenu répondant à des requêtes utilisateurs

Description du sujet et du travail demandé :

Ce stage s'inscrit dans le cadre d'un projet de recherche à long terme sur l'accès et l'exploitation de contenus hétérogènes, au travers de méthodes d'**analyse multimodale de contenus audiovisuels**. L'approche privilégiée est centrée sur la détection automatique de zones d'interaction entre intervenants et sur la caractérisation de la situation de communication et du type d'interaction correspondants.

Ce stage vise à implémenter un ensemble de briques de base pour (i) **caractériser des zones de dialogue** dans des documents de fictions, enregistrés sur **support DVD**, (ii) pour les **indexer** suivant un ensemble de descripteurs pertinents (représentation multimodale et/ou hiérarchique du contenu) et (iii) permettre une **recherche par critères** qui soit adaptée aux besoins des utilisateurs et au niveau de compréhension du contenu qu'ils recherchent.

Le stage comportera donc un aspect recueil et **analyse des besoins de groupes d'utilisateurs**, **analyse automatique des composantes audio/vidéo/texte** et un aspect **IHM** pour gérer la partie recherche et visualisation :

- le premier aspect concernera l'implémentation de briques logicielles d'analyse et d'indexation multimodale (audio, vidéo et textuelles *via* l'exploitation des sous-titres disponibles) d'un DVD (extraction des différents flux) soit à partir d'outils existants à adapter (notamment pour l'analyse du signal audio) soit *via* de nouveaux outils à concevoir ;
- le second aspect doit permettre le visionnage de séquences particulières (résultat du processus de détection et d'indexation).

Ce stage se fera en collaboration avec l'entreprise **Archean Technologies (Montauban, 82)** qui souhaite à terme proposer le financement d'une **thèse de doctorat** (*via* une convention de type **CIFRE**) pour poursuivre le travail de recherche ainsi que pour développer un premier outil commercialisable (certainement sous forme de service web). Une partie du travail demandé concerne donc **l'établissement d'une problématique de recherche** en identifiant des verrous scientifiques et techniques nécessaires à la réalisation du projet. Des compléments d'informations seront apportés lors de la prise de contact.

Stage niveau M2 avec possibilité de poursuivre par une thèse (CIFRE)

Principales compétences techniques visées :

- **Développement logiciel** (conception, développement, tests, intégration - Conception et programmation orientée objet)
- **Traitement des média audio, vidéo, texte** (traiter, analyser et interpréter des signaux audio, image ou vidéo, exploiter des données textuelles pour en extraire de l'information multimodale pertinente en mobilisant les connaissances en traitement du signal, reconnaissance des formes/machine learning, ...)
- **Interaction multimodale/dialogue** (Analyser des contenus relativement aux séquences de dialogue)

Langage(s) : Python (ou possibilité Java), Shell, prototypage possible sous matlab, possibilité JavaScript pour l'IHM

Contacts : ARCHEAN : fontan.lionel@gmail.com & IRIT-SAMoVA : isabelle.ferrane@irit.fr