

Proposition d'un sujet de thèse

Sous-titrage automatique de films avec reconnaissance des locuteurs et de leurs émotions

LIRMM, ECOSM, Montpellier

Les équipes ICAR (Image & Interaction, <https://www.lirmm.fr/icar/>) et TEXTE (<https://www.lirmm.fr/recherche/equipes/texte>) du LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier, <https://www.lirmm.fr/>), Université de Montpellier (<https://www.umontpellier.fr/>) et CNRS (<http://www.cnrs.fr/>), en collaboration avec la société ECOSM (European Compagnie of Softwares for Medias, <https://ecosm.fr/>) localisée également à Montpellier, proposent conjointement un thèse CIFRE (co-financée par l'ANRT) dans le domaine du **sous-titrage automatique de films avec reconnaissance des locuteurs et de leurs émotions**. Les objectifs principaux de cette thèse consistent à partir d'un moteur de speech-to-text d'analyser le texte pour générer des sous-titres courts et cohérents, les intégrer de manière optimale dans des séquences continues de la vidéo en fonction de la longueur des sous-titres, reconnaître et identifier les locuteurs (noms des artistes ou des personnages) permettant de les suivre tout au long d'une même séquence, mais également de reconnaître leurs expressions faciales afin d'ajouter aux sous-titres les noms de locuteurs mais également leurs émotions (joie, tristesse, colère, ...). L'analyse du texte, couplée à la reconnaissance d'expression faciale) aidera à préciser les émotions. Le sous titrage pourra différer en fonction du public et de sa provenance. Il est à noter qu'à ce jour seulement 4% des films et vidéos au monde bénéficient de sous-titrages alors que c'est le seul moyen pour les personnes sourdes ou malentendantes d'accéder à ce type de média.

La sélection des candidats s'effectuera au plus tôt, pour un début de thèse en octobre de cette année.

Pour plus d'information et pour candidater : william.puech@lirmm.fr

