

Titre : Indexation spatio-temporelle avec les descripteurs locaux structurés des flux vidéo du monitoring embarqué des patients

Encadrant : Dr. Aurélie Bugeau, Pr. Jenny Benois-Pineau (benois-p@labri.fr)

Equipe : Image et Son

Thème : Analyse et Indexation vidéo

Le thème AIV de l'équipe Image et Son mène une recherche pionnière en matière d'indexation vidéos acquis par des caméras portées par des patients dans le contexte de diagnostic et suivi des traitements thérapeutiques des démences liées à l'âge (maladie d'Alzheimer). Les vidéos embarquées présentent de nombreuses difficultés : flou lié aux forts mouvements et changements d'illumination brutaux complexité des scènes observés.

Dans ce contexte se pose le problème d'identification des objets d'intérêt le long des vidéos acquises pour constituer un flux de concepts pouvant être utilisé dans le système global d'indexation développé à des fins d'interprétation de ces contenus par des chercheurs en médecine.

Les méthodes d'état de l'art en indexation des objets s'appuient sur des descripteurs locaux (points d'intérêt SIFT, SURF.[1]..). Pour ces derniers, la signature de la vidéo ou des objets dans la vidéo est construite à partir des caractéristiques du voisinage local des points.

Cependant, ces descripteurs sont peu répétables et ne tiennent pas compte de la répartition spatiale des points d'intérêt.

La recherche d'un objet particulier dans un tel contexte requiert la mise en place de nouveaux descripteurs robustes.

L'information spatiale peut être ajoutée en regroupant ces points à l'aide de graphes. En effet, les avancées récentes sur le "graph embedding" donnent de nouvelles possibilités pour la mise en correspondance et codage de graphes.

Il est possible donc de décrire les objets à l'aide de ces nouveaux descripteurs locaux structurés que nous appelons « graphwords »[2,3].

L'objectif du stage est alors de représenter un objet par un ensemble de graphes et de descripteurs associés, et de mettre en place des mesures de similarité entre ces signatures. Il sera alors nécessaire d'étudier la construction de tels descripteurs, leur stabilité, et l'apprentissage dans l'espace des graphes.

Le stage sera effectué dans le cadre du projet européen Dem@care et sera suivi par la proposition d'une bourse de thèse.

Financement : 100% Dem@care

[1] J. Sivic et A. Zisserman, "Video google: a text retrieval approach to object matching in videos," *ICCV'2003*, vol. 2, pp. 1470-1477, 2003.

[2] Svebor Karaman¹, Jenny Benois-Pineau¹, Rémi Mégret² et Aurélie Bugeau¹, « Mots visuels issus de graphes locaux multi-niveaux pour la reconnaissance d'objets », RFIA'2012, Lyon, France

[3] . Svebor Karaman, Jenny Benois-Pineau, Rémi Mégret and Aurélie Bugeau, Multi-Layer Local Graph Words for Object Recognition, MMM2012, Klagenfurt, Austria, January 2012