Semi-supervised learning of multimedia concepts

Mots clés :

- Directeur de thèse : bernard MERALDO
- Co-encadrant(s) :
- Unité de recherche : Laboratoire de recherche d'EURECOM
- Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal : Divers

Résumé du projet de recherche (Langue 1)

I am doing my PhD in the multimedia communications department at EURECOM under the supervision of Prof Bernard Merialdo. At Eurecom I am working on indexing and learning of multimedia concepts under the umbrella of SuMACC project that is a collaboration among EURECOM, LIA (Laboratoire Informatique d'Avignon), Syllabs (http://www.syllabs.com/) and Wikio (http://www.wikio.com/). The project aims at proposing certain methods for automatic indexing of multimedia web documents, a task done by hand at the moment. Indexing using collaboration of different Medias is also among the deliverables of the project. I am working on semi-supervised clustering algorithms for improving semantic concepts video retrieval. I am also studying online and active learning techniques in order to incorporate the huge amount of un-labeled data with the few labeled training examples, on the fly.

Résumé du projet de recherche (Langue 2)

I am doing my PhD in the multimedia communications department at EURECOM under the supervision of Prof Bernard Merialdo. At Eurecom I am working on indexing and learning of multimedia concepts under the umbrella of SuMACC project that is a collaboration among EURECOM, LIA (Laboratoire Informatique d'Avignon), Syllabs (http://www.syllabs.com/) and Wikio (http://www.wikio.com/). The project aims at proposing certain methods for automatic indexing of multimedia web documents, a task done by hand at the moment. Indexing using collaboration of different Medias is also among the deliverables of the project. I am working on semi-supervised clustering algorithms for improving semantic concepts video retrieval. I am also studying online and active learning techniques in order to incorporate the huge amount of un-labeled data with the few labeled training examples, on the fly.