

Apprentissage par analyse linéaire discriminante des paramètres de fusion pour la recherche d'information multimédia texte-image

Christophe Moulin*, Christine Largeton*, Cécile Barat*,
Mathias Géry*, Christophe Ducottet*

*Université de Lyon, F-42023, Saint-Étienne, France ;
CNRS, UMR 5516, Laboratoire Hubert Curien, F-42023, Saint-Étienne, France ;
Université de Saint-Étienne, Jean-Monnet, F-42023, Saint-Étienne, France.

Résumé

Avec le développement du numérique, des quantités très importantes de documents composés de texte et d'images sont échangés, ce qui nécessite le développement de modèles permettant d'exploiter efficacement ces informations multimédias. Dans le contexte de la recherche d'information, un modèle possible consiste à représenter séparément les informations textuelles et visuelles et à combiner linéairement les scores issus de chaque représentation. Cette approche nécessite le paramétrage de poids afin d'équilibrer la contribution de chaque modalité. Le but de cet article est de présenter une nouvelle méthode permettant d'apprendre ces poids, basée sur l'analyse linéaire discriminante de Fisher (ALD). Des expérimentations réalisées sur la collection ImageCLEF montrent que l'apprentissage des poids grâce à l'ALD est pertinent et que la combinaison des scores correspondante améliore significativement les résultats par rapport à l'utilisation d'une seule modalité.

Summary

Due to the expansion of digital processes, large amount of documents composed of text and images are now shared. It raises the need for the development of models allowing to efficiently exploit such multimedia data. In the context of information retrieval, a common strategy consists in representing textual and visual information separately and fusing the scores obtained with each single media using a linear combination. A main question in this approach is how to calculate weights that balance the importance given to each modality. The aim of this article is to present a new approach for learning these weights thanks to the linear discriminant analysis of Fisher (LDA). Experimentations performed on the ImageCLEF collection show that learning the weights with the LDA is relevant and leads to an improvement when combining visual and textual scores against those obtained with a single modality.