

# ALGORITHMES POUR L'IMPLÉMENTATION D'UN SYSTÈME D'ÉTIQUETAGE AUTOMATIQUE DES IMAGES

Gilles Daigle, Shengrui Wang, Djemel Ziou  
Laboratoire de vision et traitement d'image,  
Université de Sherbrooke,  
Sherbrooke, Québec, Canada, J1K 2R1

## Résumé

La révolution informatique a mis à notre disposition un nombre phénoménal de bases de données de toutes sortes. Il est facile de rechercher un texte quelconque à l'aide de mots-clés, mais le domaine émergent de l'audiovisuel pose un problème de taille. Une solution à ce problème est d'attacher une description textuelle à l'item problématique.

La présente recherche vise à la création d'un système pour l'étiquetage automatique des images. Ce domaine n'est pas entièrement nouveau, mais les systèmes proposés jusqu'ici n'en sont qu'à l'état expérimental. Nous nous sommes donc donnés comme critère de base que le système à développer devait paraître professionnel, être rapide et donner de bons résultats.

## Abstract

Content-based image retrieval is emerging as an important research area with potential applications in many domains, notably multimedia databases and digital libraries. The goal of the present research is to create a usable system that performs automatic picture labeling. Recent studies have shown that some improvements can be achieved with a system which selects the best from amongst many models of representation. The focus of our research is thus twofold: creation of models and selection of the best ones for each label.

## 1 Introduction

La recherche d'images basée sur leur contenu émerge aujourd'hui comme un thème de recherche important ayant de nombreuses applications, notamment dans les bases de données multimédia et les bibliothèques numériques. En effet, avec la vaste quantité d'informations disponibles aujourd'hui, l'utilisation d'index et même d'outils de recherche devient cruciale. L'internet nous en offre un exemple typique. Nous n'avons jamais eu autant

d'informations si facilement disponibles, mais trouver l'information *utile* devient un défi. De nouveaux outils de recherche et d'indexage apparaissent à tous les jours. Ces outils manipulent l'information textuelle pour faire leur travail. Cependant, l'internet devient de plus en plus un médium audiovisuel et indexer ce genre d'information est tout un problème.

Il y a beaucoup de recherche dans les systèmes d'extraction de données visuelles, mais beaucoup de ces systèmes exigent que la base de données visuelles et l'outil de recherche se retrouvent tout deux dans le même système informatique. Une meilleure approche serait de pouvoir utiliser les outils textuels que l'on retrouve partout sur l'internet dans le but d'intégrer la recherche d'images avec la recherche de textes. Pour cela, l'information visuelle doit être traduite en un type d'information textuelle utilisable par les outils de recherche.

Présentement, cette traduction se fait par une personne qui soit écrit une description de l'image, soit en retire quelques mots-clés. Si l'on restreint le domaine aux images, une approche pour créer cette information automatiquement est d'étiqueter tous les éléments de l'image. Ceci nous donne une série de mots-clés se rapportant à l'image et même un pourcentage de l'image se rapportant aux mots-clés. Cette méthode a aussi l'avantage d'être moins subjective qu'une description de l'image par une personne.

Il faut souligner ici un point important: il y a un degré de subjectivité dans la création de mots-clés. Loin d'être un inconvénient, nous croyons que ceci est un avantage pour le genre de système que nous proposons. En effet, ceci veut dire que nous pouvons tolérer, sans trop de problème, un certain pourcentage d'erreur dans un système automatique d'étiquetage.

Nous proposons un système d'étiquetage à deux phases. Dans la première phase, l'utilisateur instruit le système à l'aide d'une série d'images typiques (voir fig. 6 à la fin de l'article). C'est la phase d'apprentissage. Dans la deuxième phase, le système peut être utilisé pour faire l'étiquetage d'un nombre illimité d'images. Le résultat est, pour chaque image, un fichier contenant les étiquettes que le système propose

